UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN BIBLIOTECOLOGÍA Y ESTUDIOS DE LA INFORMACIÓN

FACULTAD DE FILOSOFÍA Y LETRAS
INSTITUTO DE INVESTIGACIONES BIBLIOTECOLÓGICAS Y DE LA INFORMACIÓN

El comportamiento informativo y los procesos de valor agregado a la información de los reporteros que ejercen el periodismo de datos

TESIS

QUE PARA OPTAR POR EL GRADO DE:
DOCTOR EN BIBLIOTECOLOGÍA Y ESTUDIOS DE LA INFORMACIÓN
PRESENTA:

José Valentín Ortiz Reyes

TUTOR PRINCIPAL:
Dra. María Guadalupe Vega Díaz
El Colegio de México

MIEMBROS DEL COMITÉ TUTOR:

Dra. Patricia Hernández Salazar, Dr. Egbert John Sánchez Vanderkast
Instituto de Investigaciones Bibliotecológicas y de la Información

Ciudad Universitaria, Ciudad de México, agosto 2019

Universidad Nacional Autónoma de México
UNAM
Biblioteca Central
Dirección General de Bibliotecas de la UNAM

UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Para Angélica Antonio, por su solidaridad, por cada palabra de aliento, por su amor e inspiración, por darle alegría a nuestra vida juntos y por el tiempo que siempre fue (es) suyo.

Lo hicimos, corazón.

A Carmen y Valentino, por su amor y ejemplo, y por no dejar de esforzarse por nosotros, sus hijos.

A mi familia, pequeña y extendida.

Agradecimientos

Esta investigación es el resultado de un conjunto de esfuerzos. Desde su formulación conté con la colaboración de la Dra. Guadalupe Vega Díaz, a quien agradezco su generosa apertura para discutir con rigurosidad mis planteamientos, así como su esmero, acompañamiento, lectura y dirección durante todo este proceso.

Agradezco también al Dr. Álvaro Quijano por la exigencia metodológica, el escepticismo crítico y por confiar en el desarrollo de esta tesis. A la Dra. Patricia Hernández por creer en la pertinencia de mi investigación, por arriesgarse a colaborar en su desarrollo y mostrarme los puntos ciegos durante su elaboración. Al Dr. Egbert Sánchez Vanderkast por la minuciosa lectura y las atinadas sugerencias que llevaron a buen término esta investigación. Agradezco también las observaciones de la Dra. Judith Licea, las cuales abonaron a la forma definitiva de mi tesis.

Mi agradecimiento también a todas las personas que, por razones de confidencialidad, no puedo mencionar, pero que participaron generosamente en esta investigación compartiendo su tiempo y experiencia. Su profesionalismo y dedicación son motivo de mi reconocimiento, y en buena parte mi admiración a su oficio es el origen de esta investigación.

Gracias también a El Colegio de México y en particular a Micaela Chávez, directora de la Biblioteca Daniel Cosío Villegas, por propiciar el desarrollo profesional de quienes sumamos nuestro esfuerzo a esta noble institución.

Índice

INTRODUCCIÓN ..... **X**

EL ESTUDIO DEL COMPORTAMIENTO INFORMATIVO DE PERIODISTAS **1**

EL PERIODISMO DE DATOS 1

Antecedentes *1*

El concepto de periodismo de datos *5*

El periodismo de datos en México *8*

El periodismo de datos y el comportamiento informativo *9*

LA TEORÍA FUNDAMENTADA EN EL ESTUDIO DEL COMPORTAMIENTO INFORMATIVO **18**

TEORÍA FUNDAMENTADA: CONCEPTO E IMPLICACIONES 18

APLICACIÓN DE LA TEORÍA FUNDAMENTADA EN EL ESTUDIO DEL COMPORTAMIENTO INFORMATIVO 20

VALOR AGREGADO Y PROCESOS COGNITIVOS EN EL ESTUDIO DEL COMPORTAMIENTO INFORMATIVO 27

MÉTODO **34**

ESCENARIOS 36

INSTRUMENTO 37

Entrevista semiestructurada *37*

Obtención de datos *38*

PROCEDIMIENTO PARA EL ANÁLISIS Y LA SISTEMATIZACIÓN DE LOS DATOS 40

Etapa I. Lectura e interpretación del texto *42*

Etapa II. Construcción de códigos y categorías *44*

Etapa III. Codificación y segmentación del texto *47*

Etapa IV Análisis *52*

Visión general de categorías y códigos *59*

RESULTADOS **63**

PARTE I CONDICIONES CAUSALES 63

La brújula interna del periodista de datos *63*

Los espacios de trabajo *77*

PARTE II ESTRATEGIAS DE ACCIÓN 88

Del dato a la nota *88*
Trabajar con datos *107*
PARTE III CONDICIONES INTERVINIENTES 121
El tiempo y otros obstáculos *121*
Los elementos de un todo: visualizar, programar y editar *132*
Una nota que entiendan todos: la idea del lector *143*
Los formatos de salida *147*
PARTE IV CONSECUENCIAS (RESULTADO DE LA ESTRATEGIA) 150
El periodismo de datos visto desde su práctica *150*
PROCESOS COGNITIVOS 158

DISCUSIÓN 191

CONCLUSIONES 203

REFERENCIAS 211

ANEXO 1 GUIÓN DE ENTREVISTA SEMIESTRUCTURADA 221

ANEXO 2 ESTRUCTURA DE CATEGORÍAS Y CÓDIGOS Y SU INTERRELACIÓN CON PROCESOS COGNITIVOS 223

CUADROS

Cuadro 1: Categorías teóricas de conducta informativa identificadas a partir de la teoría fundamentada 25
Cuadro 2: Criterios de los usuarios y valor agregado 31
Cuadro 3: Entrevista semiestructurada 38
Cuadro 4: Instrumentos y participantes 40
Cuadro 5: Definiciones de código y categoría 43
Cuadro 6: Esquema basado en el propuesto por Kuckartz y McWhertor (2014) 53
Cuadro 7: Estructura del estudio 55
Cuadro 8: Presencia de códigos de procesos cognitivos invivo en escenarios y por tipo de participantes 166
Cuadro 9: Presencia de códigos de procesos cognitivos de aplicación en escenarios y por tipo de participantes 168

Cuadro 10: Presencia de códigos de procesos cognitivos de conocimiento en escenarios y por tipo de participantes....169
Cuadro 11: Presencia de códigos de procesos cognitivos de comprensión en escenarios y por tipo de participantes....171
Cuadro 12: Presencia de códigos de procesos cognitivos de evaluación en escenarios y por tipo de participantes....172
Cuadro 13: Presencia de códigos de procesos cognitivos de síntesis en escenarios y por tipo de participantes....173
Cuadro 14: Presencia de códigos de procesos cognitivos de síntesis en escenarios y por tipo de participantes....173
Cuadro 15: Códigos de procesos cognitivos con mayor frecuencia....174

FIGURAS

Figura 1: John Snow's cholera map (Cairo s/a)....3
Figura 2: Procesos de investigación periodística y escritura....14
Figura 3: Espectro de valor agregado....28
Figura 4: Proceso de análisis de la información recopilada. basada en el esquema de Kuckartz y Mcwhertor (2014)....41
Figura 5: Ejemplo de grupo de documentos en MAXQDA....45
Figura 6: Ejemplo de codificación en MAXQDA....45
Figura 7: Ejemplo de agrupación de categorías en MAXQDA....46
Figura 8: Ejemplo de segmentos codificados....48
Figura 9: Vista general de categorías y códigos....51
Figura 10: Ejemplo de tabla de resumen en MAXQDA....52
Figura 11: Procesos de acopio y análisis de la información recopilada en campo....58
Figura 12: Frecuencia de segmentos codificados por escenario....59
Figura 13: Frecuencia de códigos aglomerados por categoría....60

Figura 14: Frecuencia de códigos aglomerados en categorías por escenario 62
Figura 15: Participante, estructura y frecuencia de códigos 65
Figura 16: Espacio de trabajo, estructura y frecuencia de códigos 79
Figura 17: Desarrollo de la nota, estructura y frecuencia de códigos 90
Figura 18: Interacción con los datos, frecuencia 109
Figura 19: Obstáculos, estructura y frecuencia de códigos 123
Figura 20: Visualización, programación y producción multimedia 134
Figura 21: Audiencia, frecuencia de códigos 144
Figura 22: Entornos de publicación 147
Figura 23: Periodismo de datos, estructura y frecuencia de códigos 151
Figura 24: Profesionales que agregan valor a los datos 158
Figura 25: Procesos cognitivos y su relación en torno a las relaciones lógicas axiales 160
Figura 26: Frecuencia de verbos segmentados por proceso cognitivo 163
Figura 27: Frecuencia de códigos relacionados con Procesos cognitivos y su agrupación en categorías 165
Figura 28: Procesos cognitivos por tipo de participante 178
Figura 29: Procesos cognitivos por escenario 181
Figura 30: Procesos cognitivos presentes en el origen de una nota 183
Figura 31: Procesos cognitivos presentes en la construcción de una nota 184
Figura 32: Procesos cognitivos presentes en Contar una historia 184
Figura 33: Procesos cognitivos presentes en obstáculos 186
Figura 34: Procesos cognitivos presentes en la identificación de necesidad de datos 187
Figura 35: Procesos cognitivos presentes en trabajar con datos 188
Figura 36: Procesos cognitivos presentes en visualización 189
Figura 37: Procesos cognitivos presentes en programación 189
Figura 38: Procesos cognitivos presentes en producción multimedia 190

Figura 39: Desarrollo de la nota, caracterización de procesos 198

Así como no hay un registro perfectamente neutral,
tampoco existe una pregunta neutral.

Pierre Bourdieu, *El oficio de sociólogo*, 1975.

INTRODUCCIÓN

El estudio del comportamiento informativo en comunidades específicas es un tema que va ganando cada vez más importancia en las ciencias de la información debido al interés en comprender con mayor profundidad los procesos de uso de la información y, más recientemente, de los datos de investigación. El uso de la información con distintos propósitos es un tema que no se limita a una esfera o disciplina en concreto y es una aportación de las ciencias de la información para profundizar en la forma en qué se trabaja en entornos académicos, laborales y cotidianos. En este contexto se insertan los estudios sobre el comportamiento de profesionales en diferentes comunidades o disciplinas, por ejemplo: Ellis y Haugan (1997) centran su estudio en un grupo de ingenieros; Kuhlthau y Tama (2001) analizan el comportamiento de abogados; Leckie, Pettigrew y Sylvain (1996) proponen un modelo de búsqueda de información para profesionales basándose en los resultados de un estudio aplicado a ingenieros, profesionales de la salud y abogados.

En el caso de profesionales del periodismo, autores como Campbell (1997) han resaltado el nivel de singularidad que representa su forma de trabajo y su relación con la información. Debido a estas características son vistos como una comunidad susceptible de ser analizada por la forma en que buscan y utilizan la información, por las limitaciones de tiempo con que trabajan, por la cantidad de canales de comunicación a la que están expuestos, por la agenda que deben cubrir,

las tareas que deben ejercer y el nivel de exposición de los productos que deben generar.

Cabe señalar que en los últimos años el ejercicio periodístico se ha caracterizado por un ambiente de constante cambio, producido no sólo por los nuevos medios o canales de comunicación que han surgido a partir del desarrollo de las nuevas tecnologías, sino también por la competencia de un mercado que, sin dejar atrás el discurso de ofrecer información inmediata u oportuna, busca brindar productos que los distingan de otros medios de información. A esto se suma además la apertura de información gubernamental, el periodismo ciudadano que se emprende principalmente en la red y la proliferación de herramientas para el análisis de datos.

Este marco nos ayuda a explicar el surgimiento de un tipo de periodismo que se basa en la interpretación de grandes cantidades de datos que, tras un proceso de análisis, hace pública esta información en un contexto significativo para la ciudadanía. Esto es el periodismo de datos, que como lo define Egawhary y O'Murchu (2010), es la capacidad de analizar y examinar números, de manejar conjuntos de datos, interpretarlos y contar una historia en forma de artículos, visualizaciones o infografías. El profesional que practica este tipo de periodismo realiza una serie de procesos que se asocian no sólo con la búsqueda de información, sino también con la integración, interpretación y síntesis de ésta. Dentro de este entorno, la presente investigación busca identificar el comportamiento informativo de los periodistas (en este caso los involucrados en hacer este tipo de periodismo) durante el desarrollo o elaboración de productos noticiosos que incluyan datos y elementos contextuales que propicien la reflexión de sus lectores. Por esta razón, el comportamiento informativo se concibe, en este caso, como un proceso global en el que es necesario aplicar habilidades cognitivas tales como la evaluación, la síntesis y la integración de datos e información para generar un producto o resolver un problema, para este caso una noticia con valor agregado.

Se tiene claro el debate sobre el término adecuado para describir este tipo de procesos relacionados con la información (ver, Montesi, 2011), pues en otros

estudios se emplea, incluso de forma indistinta, el concepto de comportamiento informacional para referirse también al término en inglés “information behavior”. En nuestro caso empleamos el término de comportamiento informativo, por su asociación con los procesos globales, activos y pasivos de interacción con la información (ver, Hernández et. al, 2007)

Identificar cómo los profesionales involucrados en la elaboración de una nota de periodismo de datos resuelven sus necesidades de información y toman decisiones es un tema que nos abre interrogantes acerca de la forma en que interactúan con los datos, los ubican, sintetizan, interpretan y los dotan de sentido para conformar un producto noticioso. A diferencia de las investigaciones dirigidas a caracterizar patrones de uso o a identificar los principales recursos utilizados, el presente estudio propone centrarse en el análisis y la caracterización de los procesos que intervienen en la elaboración de un producto en concreto.

Con este propósito, se ha decidido emplear una perspectiva similar a la utilizada por Attfield, Blandford y Dowell (2003), quienes además de analizar las necesidades de información, también analizan el proceso de diseño y prospectiva; es decir, la forma en que se integra sintetiza y va dando forma a un producto de información.

Como ya se mencionó, el tema del comportamiento informativo plantea varias interrogantes, pero la pregunta que conduce a esta investigación se enfoca en identificar *¿Cuál es el comportamiento que siguen los reporteros que se dedican al periodismo de datos durante la elaboración de un producto noticioso?* Bajo los supuestos de que:

- La identificación de los procesos cognitivos, prácticas y disposiciones asociados con la elaboración de un producto noticioso de esta naturaleza, nos permitirá definir patrones de uso de la información.
- El desarrollo de productos noticiosos relacionados con el periodismo de datos requiere de un conjunto de tareas y actividades complejas susceptibles de caracterizarse.

De manera puntual, el objetivo de la investigación es *analizar y caracterizar el comportamiento informativo de profesionales que hacen periodismo de datos durante la elaboración de un producto noticioso de esta naturaleza*. Con relación a los objetivos específicos, el estudio se propone:

- Identificar los procesos cognitivos, prácticas y disposiciones que realizan los profesionales en: la resolución de sus necesidades de información, la interacción con los datos, la interpretación y síntesis de la información y la forma en que dota de significado a la información durante la conformación de un producto noticioso asociado con el periodismo de datos.
- Identificar los patrones de comportamiento que siguen las personas que ejercen el periodismo de datos para delinear una caracterización de las etapas o procesos que se desarrollan durante la elaboración de un producto noticioso.
- Analizar al periodismo de datos bajo la perspectiva del valor agregado, el cual, en el contexto de este trabajo, se entiende como un sistema de información en el que los insumos son datos que, tras procesos de sistematización, integración, interpretación y síntesis, van cambiando gradualmente de dato a información y de información a conocimiento.

El planteamiento de este tipo de objetivos supone aplicar un método cualitativo que permita explicar el fenómeno de interés a partir de la interacción directa con los sujetos de estudio, con el fin de que la interpretación de los fenómenos sea en su contexto y del modo más cercano a como ellos los sienten o viven (Blaxter, Hughes y Tight, 2000). Es el caso de lo que nos proponemos analizar y comprender ampliamente, de modo que este método permite identificar las situaciones y los contextos en los que se lleva a cabo el proceso y los criterios que se utilizan para la toma de decisiones. Es por ello que se consideró pertinente el uso de la Teoría Fundamentada (TF) para poder caracterizar el comportamiento informativo a partir de un cuerpo de categorías con una coherencia lógica que derive de la comparación constante de los datos recolectados y de los códigos generados de forma progresiva

durante su estudio, con el fin de analizar los procesos de valor agregado y la forma en que se desarrollan en contextos específicos, evitando en la medida de lo posible categorías y conceptos predefinidos. La vigencia de la TF radica en la capacidad de ofrecer hallazgos respaldados en datos sistematizados de forma coherente, además de su capacidad explicativa frente a nuevos fenómenos. El hecho de trabajar con el significado de los datos y hacer una lectura abierta y profunda de los mismos, permite ampliar las posibilidades de encontrar nuevas explicaciones, causas o significados, dejando a un lado las fórmulas preestablecidas que limitan la interpretación e incluso los propios resultados. En ese sentido la teoría que se genera se centra en la forma en que los individuos interactúan con el fenómeno bajo estudio y se conforma una estructura lógica de conceptos emergentes (Urquhart, 2013). En sentido estricto, la TF puede ser vista como una manera de aproximarse a los datos de investigación y cuestionarlos, de forma rigurosa, coherente y sistemática, para establecer una interpretación abierta que da pie a nuevos hallazgos.

El estudio abre con una panorámica del periodismo de datos en el que se ofrecen antecedentes de este tipo de periodismo, las habilidades técnicas que requiere, así como su resonancia en la esfera social. También se ofrece un compendio de las principales investigaciones en torno a su definición, alcance, evolución y práctica en el contexto nacional e internacional. De la misma forma se brinda un análisis de las investigaciones que centran su atención en el comportamiento informativo de periodistas, en este caso se describe sus métodos y principales hallazgos.

Más adelante se encuentra el apartado sobre la TF en el que se habla de los trabajos seminales de Glaser y Strauss y se trata el tema de la inhibición teórica temporal y sus implicaciones al momento de desarrollar una investigación empleando este método. También se ofrece una discusión sobre los riesgos y prejuicios que hay en torno a la TF y el papel que tiene el investigador en este terreno. Por otro lado, se habla sobre la aplicación de la TF en el estudio del comportamiento informativo a principio de la década de los 80 y cuya mayor aportación, en ese momento, fue cambiar la óptica de que las personas no deben

ser vistas sólo como objetos de estudio, sino como participantes (Wilson, 1990). Esto dio pie a estudios más realistas, construidos a partir de cero y con énfasis en el entorno (Ellis, 1993). El apartado cierra con una descripción del concepto de valor agregado de Taylor (1986) que busca clarificar y contextualizar el término a la luz de la proliferación de datos y de procesos para su transformación. Por último, se hace repaso de los principales estudios que analizan el componente cognitivo en el estudio del comportamiento informativo.

En cuanto al método empleado se ofrece a detalle información sobre cada uno de los tres escenarios en los que se interactuó con los participantes, así como de los procesos de acopio y sistematización de la información. En cuanto al procedimiento para el análisis y sistematización de los datos se describe la forma en que se trabajó en la lectura e interpretación de los datos, la construcción de códigos y categorías, el proceso de codificación, segmentación y análisis y las implicaciones de trabajar con el programa de MAXQDA.

El análisis de resultados se divide en tres partes: Parte I. Condiciones causales: en el que se analizan los resultados relacionados con los participantes, su formación, áreas de interés, motivaciones y disposiciones. Parte II. Estrategias de acción: se describen los procesos que se ponen en juego para la elaboración de una nota de periodismo de datos, desde el proceso inicial, el desarrollo de la nota, los cambios de ruta y estrategias de validación. Parte III Condiciones intervinientes: se trata de los principales obstáculos que enfrentan los participantes en el desarrollo de una nota de periodismo de datos. Parte IV. Consecuencias: en este caso se habla de las prácticas y estrategias emprendidas, así como de las características del producto final.

El apartado sobre los procesos cognitivos centra el análisis en los procesos que ponen en juego los participantes en cada uno de los escenarios durante el origen y la construcción de una nota, así como en la interacción con los datos.

Finalmente, se ofrece una discusión de resultados en la que se recogen los principales hallazgos de cada uno de los apartados y se sintetizan en una representación gráfica que caracteriza los procesos que intervienen el desarrollo de la nota. Las conclusiones abstraen los resultados obtenidos y ofrecen un testimonio

de lo que implicó trabajar con datos cualitativos para responder los objetivos iniciales y en particular el empleo de la TF a lo largo de todo este proceso. Incluye, además, una reflexión sobre una veta de investigación que puede resultar de interés para aquellos interesados en el estudio del comportamiento informativo en el contexto del uso de datos.

EL ESTUDIO DEL COMPORTAMIENTO INFORMATIVO DE PERIODISTAS

El periodismo de datos

Antecedentes

A principios de los años setenta, cuando el uso de la computadora en una redacción era visto como un lujo, Philip Meyer escribió su libro *Precision Journalism: a reporter introduction to Social Science Method* (1973) para visualizar la necesidad de que el ejercicio del periodismo incorporara herramientas de investigación propias de los científicos sociales. De acuerdo con Parasie y Dagiral (2013), el uso de la computadora y la estadística han potencializado el trabajo periodístico y las bases de datos son consideradas como el elemento clave en un momento de amplia oferta informativa.

Fue hasta mediados de la primera década del 2000 cuando comenzaron a incorporarse a las redacciones de diarios como el *New York Times* y *The Guardian*, personal capaz de armar una historia basada en datos y programadores capaces de conjugar elementos de diseño y programación para ofrecer elementos gráficos de fácil comprensión. Otros periódicos que cuentan con secciones de periodismo

de datos es el *Chicago Tribune*. A nivel Latinoamérica resalta el trabajo de *La Nación* con su blog NACIÓN data.

En relación con los estudios sobre el periodismo de datos, se puede consultar a Parasie y a Dagiral (2013) quienes analizan la integración de los programadores y el personal de la redacción en un diario norteamericano. Este estudio se vale de la metodología de la entrevista semiestructurada y el análisis de contenido para evidenciar la forma en que los reporteros trabajan con los datos y logran extraer, por medio del acceso, combinación y el procesamiento de datos, información relevante y clara para una sociedad determinada. Este es uno de los primeros estudios en advertir la necesidad de utilizar la perspectiva etnográfica para analizar el proceso de producción de una noticia y los cambios en el *ethos* del periodista.

De acuerdo con Chaparro (2014) este tipo de periodismo se vale de herramientas estadísticas y de programas para la visualización que facilitan la comprensión de las historias de un modo más claro para el público. Aunque este tipo de periodismo nace formalmente en el siglo SI, es posible encontrar antecedentes de periodistas que se han valido de datos para visualizar situaciones y obtener conclusiones. Alberto Cairo (s/a) comparte uno de estos primeros ejemplos:

> “Existe un mapa que fue hecho en el siglo XIX por el médico John Snow, quien investigaba qué era lo que causaba el cólera en Londres en ese siglo, ahora sabemos que se debe a beber agua contaminada por una bacteria. En el siglo XIX se pensaba que el cólera se transmitía al inhalar aire contaminado debido a que muchos doctores observaron que los lugares que apestaban eran donde se presentaban más casos de cólera, incluso establecieron una relación entre las dos variantes "entre más apeste el lugar, existen más casos de cólera". Sin embargo, John Snow sabía que el cólera no podía ser causado por aire contaminado, él sabía que debía ser por algo más. Su hipótesis era que el cólera era causado por beber agua contaminada, pero ¿cómo comprobarlo en esa época sin los avances de hoy? Lo hizo mediante un mapa.“(Cairo, s/a)

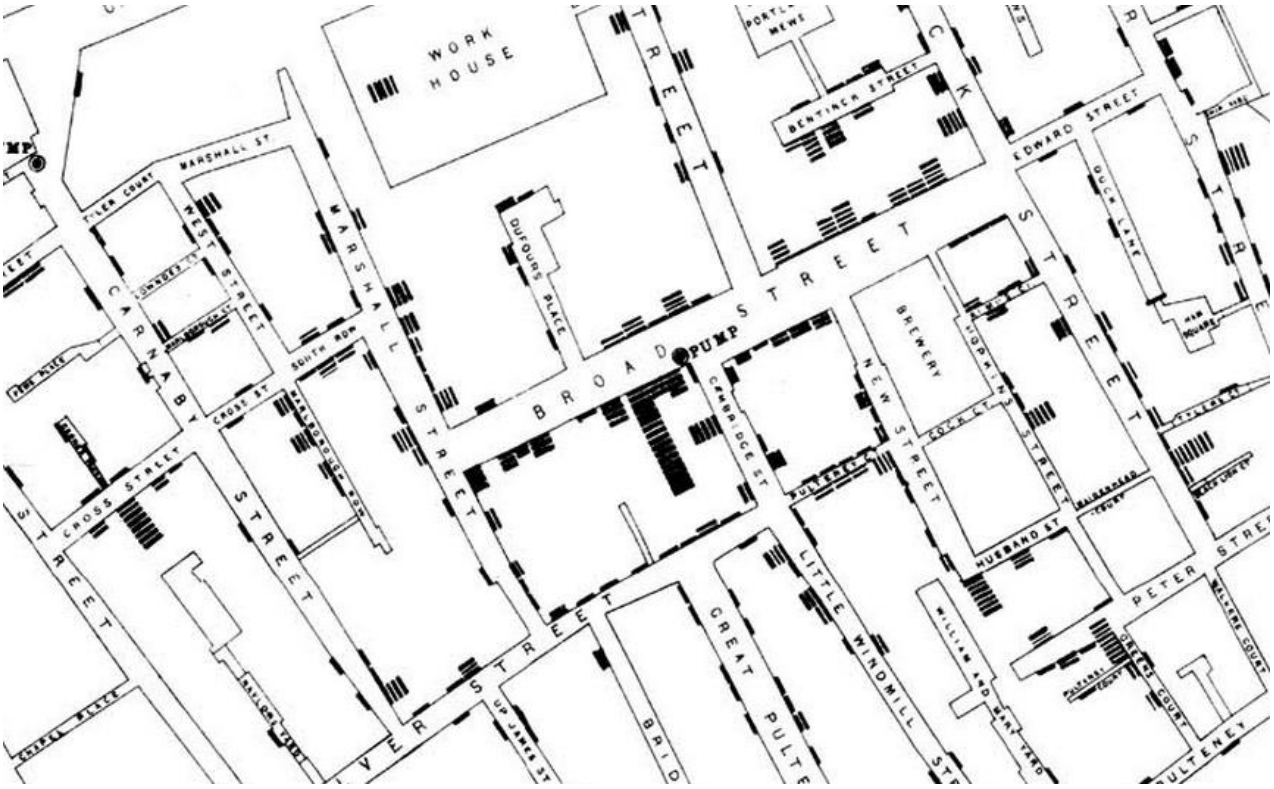


FIGURA 1: JOHN SNOW'S CHOLERA MAP (CAIRO S/A)

Continúa la cita:

> "Las rayas más oscuras representan el número de muertes, entre más se aproximen al centro del mapa, más muertes registradas. Cerca del centro hay un pozo de agua y ahí hay más casos de cólera. Es muy claro. Estos patrones de concentración comprueban la hipótesis, pero también tenemos aquí ciertas cosas que la podrían refutar, ¿cómo cuáles? También hay gente que murió de cólera, muy lejos del pozo de agua, ¿por qué murieron? Hay muchos edificios alrededor del pozo de agua en donde no se registraron muertes, eso podría refutar potencialmente la hipótesis.
>
> John Snow hizo algo que nosotros los periodistas deberíamos hacer más, puso atención a estas excepciones. Fue a esos lugares y descubrió que personas que murieron de cólera, que vivían lejos del pozo de agua, cada día camino a sus trabajos y a sus escuelas pasaban y bebían agua del pozo, por eso se contagiaron con la enfermedad" (Cairo, s/a).

Ana María Ávila hace referencia a Clay Shirky (Bounegru, et. al, 2011), profesor de la Universidad de Nueva York, para enfatizar que el periodismo actual es visto como un ecosistema de esfuerzos de investigación que reúne a diversas disciplinas en una red de colaboradores no jerárquica para responder a las necesidades de rendición de cuentas de los ciudadanos. En ese sentido, el periodismo de datos

puede entenderse también como un intento de los medios por adaptarse a los cambios en la audiencia y en los nuevos entornos informacionales que exigen formas más interactivas y multidimensionales.

Hoy en día el periodismo de datos se ha beneficiado del aprovechamiento de un volumen de datos inmenso que aumenta continuamente. Nos referimos al Big data o grandes datos, en referencia a los sistemas de información que permiten la gestión de grandes conjuntos de datos (Antón Bravo, 2013). De igual forma la vinculación de los datos por medio de iniciativas como “Linked Open Data” (Datos Vinculados Abiertos) que permite utilizar grandes cantidades de datos, relacionarlos y analizarlos como una unidad. Se vale también de la visualización de los datos como una herramienta heurística capaz de provocar múltiples lecturas e interpretaciones, así como de las herramientas de mapeo de los Sistemas de Información Geográfica (GIS, por sus siglas en inglés) las artes de visualización de las estadísticas y el diseño gráfico (Rodríguez Brito y García Chico, 2013). Más allá de obtener el dato, una de las claves básicas de este periodismo radica en la contextualización de la información, en divulgar su significado y de ahí contar una historia que se respalde con datos confiables.

La mayor parte de los estudios que se han realizado en España sobre el periodismo de datos se ciñe a la aproximación del perfil del periodista que genera este tipo de productos y a evidenciar el estado de este tipo de periodismo en sociedades específicas. En el caso de Eva María Ferreras (2013) hace un análisis del perfil profesional que emerge de esta práctica y se delinea un estado de la cuestión en España para luego proponer investigaciones que atiendan temáticas como: la introducción de informaciones originales a partir de las nuevas posibilidades tecnológicas; la necesidad de que los periodistas adquieran nuevas competencias y habilidades en el campo del análisis de grandes bases de datos; la programación y la generación de visualizaciones; las nuevas narrativas utilizadas en las visualizaciones; los aspectos económicos; su vínculo con la transparencia y la información gubernamental.

En esa línea encontramos el trabajo de López-García, Toural-Bran, Pereira-Fariña y Barbosa (2009) que analiza la irrupción de un nuevo perfil de periodista

provocado por la evolución de los formatos y la aparición de nuevos medios de comunicación basados en Internet. En este caso los autores trazan un comparativo entre habilidades y responsabilidades surgidas en el contexto digital y las tareas del periodista tradicional. Finalmente resaltan el papel de las bases de datos en la estructura y documentación de las informaciones, la contextualización, la creación de nuevos modelos narrativos y la visualización de los contenidos periodísticos.

Por otro lado, ubicamos el trabajo de Chaparro (2014) que hace una descripción teórica del término periodismo de datos y analiza la oferta educativa sobre esta modalidad en España y en el extranjero. En ambos casos utiliza la investigación documental.

En el caso de Flores Vivar y Salinas Aguilar (2012) se analiza la forma de trabajar el periodismo de datos en los medios españoles y se muestra cómo la narrativa transmedia converge con este tipo de periodismo. La metodología utilizada se basa en la investigación documental sobre los antecedentes del periodismo de investigación, la evolución de los equipos de trabajo en España y el conocimiento tecnológico de los periodistas.

Una de las tesis que ubicamos sobre periodismo de datos es la de Martina Cavanna en Italia (2013). El objetivo de este trabajo es analizar el origen y evolución del periodismo de datos para proporcionar un marco amplio que lo describa en su evolución histórica y estructural. Para ello hace un repaso de los trabajos más relevantes en este ámbito para luego discutir sobre las habilidades que requiere el periodista para recopilar, filtrar y contar una historia a partir de datos. Ofrece también un panorama de las principales herramientas de filtrado y visualización de datos. Finalmente se enumeran los diarios más importantes a nivel internacional que han incursionado en esta nueva forma de hacer periodismo.

El concepto de periodismo de datos

De acuerdo con Nikki Usher (2016) el periodismo tradicional enfrenta uno de los momentos más críticos y de mayor transformación desde el punto de vista tecnológico. Entre ellos está la forma en que las personas utilizan la tecnología y la

diversidad de usos sociales que ésta ha tenido. Por otro lado, se encuentra la autoridad del periodismo para contar la verdad en un momento en que la información se genera y transmite a gran velocidad. Todo ello exige a quienes practican el periodismo adaptarse a un entorno en constante cambio.

El periodismo de datos es una vertiente de lo que se denomina en el terreno profesional como periodismo interactivo: una forma de presentar o narrar visualmente una historia por medio de la codificación multicapa y el control táctil del usuario, con el propósito de informar (Usher, 2016). En su libro *Interactive Journalism: hackers, data, and code* Nikki Usher (2016) retoma la definición de periodismo de datos que utiliza Alexander Howard (2014), para señalar que es una de las más claras y completas para referirse a este término que suele ser confuso incluso para quienes lo practican. La definición es la siguiente:

> "[El] acopio, limpieza, organización, análisis, visualización y publicación de datos para respaldar la creación de productos noticiosos […] en su forma más elemental el periodismo de datos combina el tratamiento de datos, su acopio y validación. La aplicación de la estadística para interrogar a los datos y la visualización para su presentación" (Howard, 2014).

Normalmente en este tipo de periodismo conviven por igual la ciencia de datos, el rigor científico, la aplicación de metodologías, la visualización[1], las narrativas visuales y el periodismo de precisión. Además, su práctica se caracteriza por la obtención de datos en grandes cantidades, su procesamiento, análisis e interpretación por medio de computadoras y métodos estadísticos, para luego mostrarla de una forma comprensible al lector. Es decir, a partir de datos se cuentan historias.

Al hablar del periodismo de datos, autores como Parasie y Dagiral (2013) han resaltado su importancia en la mejora de la democracia, particularmente en un momento en que los gobiernos ponen a disposición de la ciudadanía grandes cantidades de información. En ese sentido, Ferreras (2013) destaca el papel que

[1] El término visualización "es empleado en situaciones en donde determinados datos que en principio no son visuales, son cuantificados y transformados en representaciones visuales; por ejemplo, la información meteorológica, las variaciones del mercado de valores" (Bárcenas, 2012).

desempeñan instituciones como *Nieman Foundation* y *Knight Foundation* en la promoción de este tipo de periodismo que, a su juicio, representa una tendencia actual y de futuro para la profesión periodística. El tema del periodismo de datos, tal como señala Ferreras, es una disciplina incipiente y centrada más en circuitos profesionales que académicos.

En cuanto a la concepción del periodismo de datos que se asume en esta investigación, es necesario precisar que más allá de definirlo o delimitarlo nos interesa analizarlo con la perspectiva del valor agregado, esto a raíz de la perspectiva de Ortiz (2007) en el que al hablar de periodismo de investigación los participantes aludían a que había una exigencia institucional por documentar un texto, ofrecer contexto y propiciar la reflexión de los lectores. Esta exigencia en forma de valor agregado nos remite necesariamente al esquema propuesto por Taylor (1982) en el que los datos se convierten en información y conforme esta información es seleccionada, validada o interpretada hay un cambio gradual de la información hacia el conocimiento, posteriormente este conocimiento nos ayuda a comprender mejor las opciones, las ventajas o desventajas y finalmente nos ayuda a tomar decisiones. Estos atributos que se le añaden a los datos al momento de sistematizarlos, integrarlos y sintetizarlos van conformando un valor agregado.

Acorde con esta óptica el periodismo de datos puede ser visto como un proceso complejo, en el que se habla de agrupar, clasificar, relacionar y formatear los datos para luego analizar la información, por medio de procesos de síntesis, interpretación, comparación, validación, evaluación y separación de los propios datos.

En esa lógica el análisis del periodismo de datos se centra en un sistema de información en el que el principal insumo son datos a los que se les dota de sentido y significación. En otras palabras, esta perspectiva nos permitirá evidenciar o hacer más ilustrativos los elementos que dan forma al proceso de valor agregado en el desarrollo o ejercicio del periodismo de datos.

En este concepto la información más allá de verse como un elemento estático es visto como un elemento que al ser interpretado por el periodista se torna dinámico y orientado a una tarea, propósito y contexto específico. En ese sentido, el

comportamiento informativo involucra también los cambios en la estructura cognitiva del periodista a partir de la incorporación de nuevos datos.

El periodismo de datos en México

En el caso de México, ubicamos esfuerzos incipientes en diarios como *Reforma* y *El Universal*. En el *Reforma* el principal referente es el *Ejecutómetro*, que es un gráfico que integra datos por entidad federativa sobre homicidios violentos en México, durante los sexenios de Calderón y Peña Nieto. Por su parte *El Universal* de acuerdo con Jesús Ibarra (s/a), es de los pocos diarios que cuentan con un área de periodismo de datos creada al interior de la empresa para trabajar, desde esta vertiente, temas de investigación. Este periódico incluye en sus contenidos digitales la sección “Periodismo de datos”.

Fuera del ámbito periodístico algunas asociaciones como *Fundar* han utilizado la perspectiva del periodismo de datos para generar micrositios sobre: la forma en que será invertido el presupuesto en 2014; el uso de la publicidad oficial en las entidades federativas; el estatus de las iniciativas de ley por tema; el destino de los subsidios destinados al campo y sus beneficiados, entre otros. En cuanto a los medios electrónicos podemos encontrar el trabajo de *Reporte Índigo*, *Animal Político* (con colaboraciones de Data4) y *Sin embargo* que cuenta con una sección denominada Unidad de datos en la que se publican notas desarrolladas a partir de datos gubernamentales.

También se encuentra el trabajo de Data4, un colectivo de investigadores, desarrolladores y diseñadores que trabajan en el procesamiento, análisis y gestión de datos para generar contenidos como textos editoriales, reportes, micrositios y plataformas. De igual forma está el portal de noticias *La Data* un medio digital que integra a periodistas, físicos, sociólogos, matemáticos, diseñadores, economistas, programadores que aprovechan sus capacidades en visualización de datos, gestión de datos abiertos y nuevas narrativas para enriquecer el análisis político y social.

Las investigaciones en México sobre este tema son limitadas, por no decir que nulas. Al respecto ubicamos el trabajo de Bárcenas Aboites (2012) que de forma tangencial trata el tema del periodismo de datos en su estudio sobre la definición de infografía. En este análisis el autor habla sobre las posibilidades que los medios digitales ofrecen al proceso de visualización y discute las posibilidades de considerar a la infografía como un medio periodístico. Finalmente, analiza las propiedades semióticas de este medio.

Por último, cabe destacar la labor del Centro de Investigación y Docencia Económicas (CIDE) que ha difundido las capacidades del periodismo de datos en México por medio de su proyecto *Data periodismo CIDE* que promueve la investigación periodística a partir de datos gubernamentales disponibles.

El periodismo de datos y el comportamiento informativo

El estudio del periodismo de datos es escaso. Hasta el momento no se han encontrado estudios que analicen las fuentes de información que utilizan, las estrategias que emplean los periodistas para buscar y evaluar información, y más aún, la forma en la que integran y sintetizan la información para crear un producto noticioso. Asumimos que el periodista, al estar involucrado con el uso y el tratamiento de la información en contextos en que la entrega de productos se acota a tiempos reducidos y con especificaciones concretas, constituye un caso pertinente para analizar elementos del comportamiento informativo.

Por ello, se considera que el periodismo de datos es un terreno inexplorado en el campo de los estudios de la información. Así como también son pocos los medios que comienzan a ejercer este tipo de periodismo que se asemeja al periodismo de precisión o asistido por computadora (CAR), una modalidad que surge con la apertura de bases de datos de sectores públicos y privados que planteaban para el periodista una exigencia de formación para el uso y aprovechamiento de programas que permitieran sistematizar grandes volúmenes de información, así como examinar las bases de datos, ubicar patrones, cruces, asociaciones y correlaciones entre los datos (Ortiz, 2007).

Modelos como el de Leckie, Pettigrew y Sylvain (1996) han resaltado el papel que juega el rol, las tareas asociadas, las características de las necesidades de información, el conocimiento de las fuentes y los resultados en los procesos asociados con el comportamiento que emprende un profesional para cubrir una necesidad de información relacionada con sus tareas rutinarias. Esta propuesta se formuló con el propósito de ofrecer un modelo general susceptible de aplicarse a todo tipo de profesional. La base principal de dicho modelo es una investigación documental sobre el comportamiento informativo de tres grupos de profesionales: los ingenieros, los médicos y los abogados. En ese sentido, como se puede deducir, las características de las necesidades de información dependen de la profesión que desempeñan, la urgencia con que se requieren y qué fuentes se tienen a su disposición. Uno de los méritos del estudio de Leckie, Pettigrew y Sylvain es la caracterización del comportamiento informativo en términos o categorías amplias susceptibles de aplicarse para una mejor comprensión de un profesional. Sin embargo, al ampliar el alcance de su estudio le resta profundidad, en relación con la identificación de características específicas de cada grupo.

Otro modelo que se ha centrado en el aspecto situacional es el de Dervin (1983) que hace énfasis en las tareas de búsqueda y el logro de objetivos. Una de sus finalidades es la comprensión de los procesos que desarrollan los participantes para dotar de sentido sus tareas y necesidades, en específico, la identificación de una necesidad de información. A decir de Attfield (2005) este modelo se basa en la estructura de ‘Situación-Vacío-Uso’; es decir, una necesidad de información es vista como un vacío en la comprensión de un hecho o fenómeno y este vacío oscila entre una situación actual (contexto en el que surge la necesidad) y el uso que se le dará a la información. De acuerdo con Attfield, la información más allá de ser una representación estática es vista como un elemento dinámico o subjetivo que puede reconfigurar el estado de conocimiento. De tal modo que el comportamiento informativo es concebido como una actividad dinámica en el que el periodista actualiza, y pone en juego, su estructura interna a partir de la incorporación de nuevos datos. Bajo esa lógica, se deduce que el comportamiento informativo requiere también de una perspectiva cognitiva para comprender cómo opera el

participante en su toma de decisiones a partir de la información que incorpora. En el caso de los periodistas se trata de comprender la serie de procesos que desarrolla el participante frente a una situación problemática, como lo es la creación de un producto noticioso.

La mayor parte de los estudios que se han realizado sobre el comportamiento informativo de la figura del periodista tradicional, en el sentido más extenso, han sido desarrollados por investigadores ubicados en Inglaterra y Estados Unidos. Entre los primeros destaca el trabajo de David Nicholas, Peter Williams, Helen Martin y Peter Cole (2000) en el que se analiza el impacto que ha tenido la introducción de Internet en las redacciones, la caracterización de las necesidades de información de los distintos tipos de periodista, las restricciones en términos de tiempo y las estrategias de verificación. Para su desarrollo se utilizó la técnica de entrevista y el estudio de caso. Se centró en periodistas de los diarios *The Independent*, *The Sunday Times* y *The Guardian*. Una de las conclusiones que arroja la investigación de Nicholas y su grupo de colaboradores es que los periodistas requieren de información validada, actual y de acceso rápido. Sin embargo, advierten que el principal obstáculo sigue siendo la disponibilidad de tiempo. Este estudio que puede considerarse como fundacional, es retomado por Charlotte Wien (2000) para el desarrollo de un curso de recuperación de información para estudiantes de periodismo; por Chinn (2001) quien retoma la caracterización de necesidades de información que propone Nicholas para analizar el comportamiento informativo de tres periodistas y por Poteet (2000) que con base en los resultados de una encuesta realizada a periodistas de los diarios *News & Observer* y *Charlotte Observer* caracteriza las necesidades de información de este grupo.

Entre los trabajos que se han escrito en Inglaterra, resaltamos el de Attfield, Blandford y Dowell (2003), quienes realizan un estudio de corte cualitativo que se basa en una serie de entrevistas a periodistas del diario *The Times*, para identificar las actividades que realizan los periodistas para cubrir sus necesidades de información. El estudio provee una descripción amplia y profunda de las motivaciones que están detrás de las prácticas de búsqueda de información y su

relación con los requerimientos del producto noticioso con el que trabajan, el contexto en el que se produce y de los recursos que tienen disponibles para su producción. Una de las principales aportaciones de este estudio es la idea de que el comportamiento informativo no sólo incluye la localización de la información, sino que también la conforma la subsecuente gestión, organización y uso de la misma en un proceso amplio de construcción de un producto noticioso con especificaciones propias que se va conformando a partir de la perspectiva o ángulo que el reportero le quiere dar en el tratamiento de los datos y la información que de forma paralela localiza y va reconfigurando el producto. Esta aproximación al producto final se da en un contexto iterativo entre la información localizada y la integración de la misma en un entorno de incertidumbre. Attfield reconoce que muchos de los estudios que analizan el comportamiento informativo tienen como objeto de estudio a personas, que entre sus tareas se encuentra la escritura (2003). En ese sentido, la escritura se convierte en una situación de resolución de problema o *problem solving*, pero además involucra un proceso de diseño. Es decir, bajo la óptica de esta perspectiva, el autor es un diseñador de textos que afronta las siguientes situaciones (ver, Attfield, Blandford, y Dowell, 2003):

- Especificación incompleta: se refiere a la carencia de información en tres componentes de los problemas para el diseño: el estado de inicio, el estado del objetivo por cumplir y la función de transformación que va del comienzo al cumplimiento del objetivo. Es decir, no hay una noción clara de lo que se pretende desarrollar en un texto. Hay una indeterminación o incertidumbre durante el comienzo y el desarrollo del mismo.
- Generadores primarios: en el caso de los periodistas, se trata de requerimientos adoptados en el desarrollo de un producto noticioso. El más común, es la idea de dotar de un ángulo o perspectiva a un texto. Normalmente, esta solicitud la hace el editor durante las primeras fases. En ese sentido, el ángulo se vuelve una especulación de lo que puede ser un texto, de tal modo que se convierte en una hipótesis o conjetura que puede confirmarse o refutarse de acuerdo con los resultados de la búsqueda de información. De acuerdo con Attfield, Blandford y Dowell (2003) varios

periodistas hablaron de que en ocasiones durante los procesos de escritura y búsqueda de información surge un mejor ángulo que el que habían imaginado al comenzar con el texto y esto desestabiliza el proceso, pues los obliga a modificar continuamente sus planteamientos.

- Análisis y síntesis dinámica: el autor, visto como el diseñador de un texto, se involucra en un proceso continuo de lectura y relectura. Durante este proceso se va formando en él una nueva apreciación del problema. Durante las primeras fases el autor se avoca a la tarea de integrar ideas en el texto. Durante la fase de reflexión, el autor revisa su texto (*reviewing*) y genera nuevas ideas (*contemplation*) y finalmente toma decisiones sobre que integrar y cómo organizarlo (*planning*).
- Obstáculos: la escritura está llena de restricciones que operan en distintos niveles: de manera superficial o más profundos, sobre todo tratándose de un periodista. Algunos ejemplos son las implicaciones sociales y políticas; el evitar la difamación; la veracidad de los hechos conocidos; organizar la información; presentar argumentos válidos; usar el humor; asegurar la estética de la prosa; utilizar la gramática y ortografía de forma apropiada; apegarse a las limitaciones de espacio o tiempo que estipule el medio en el que publican; el ángulo o perspectiva que se le quiere imprimir a un texto; la originalidad del tema o la perspectiva; la estructura y las características del género periodístico que desea desarrollar, entre otros.

Lo que se observa, tras la descripción de estos elementos es la constante exposición a la indeterminación y a la incertidumbre en el proceso de búsqueda/escritura. No sólo eso, la formulación de un producto noticioso es un proceso dinámico que exige una revisión y evaluación constante de lo que se pretende comunicar, de lo que se está diciendo, de cómo se está diciendo, de lo que falta por decir y de lo que nos hace falta para comunicarlo mejor. Esto último, en relación con el comportamiento que emprende el periodista para cubrir sus necesidades de información. La Figura 2 sintetiza los principales hallazgos de Attfield y Dowell (2003):

FIGURA 2: PROCESOS DE INVESTIGACIÓN PERIODÍSTICA Y ESCRITURA

Fuente: Attfield, S., & Dowell, J. (2003). Information seeking and use by newspaper journalists. *Journal of Documentation*, 59(2), 187–204. (Traducción propia.)

En esta figura se observan tres etapas centrales: el inicio, la preparación y la producción. También se pueden observar las restricciones del propio producto en términos de relevancia, correspondencia con los hechos, fechas límites, entre otras. También se aprecian las restricciones de los recursos en relación con la disposición de archivos, contactos personales, repositorios y programas. En este rubro se incluyen también los recursos internos como la memoria del autor, las estructuras cognitivas y el conocimiento acumulado sobre un tema. En resumen, al principio de la tarea el periodista se enfrenta a un conjunto de restricciones como un límite de palabras y una fecha de entrega. En esta etapa se prefigura un ángulo o perspectiva por utilizar, éste hará las veces de hipótesis o conjetura que se confronta con las nociones de originalidad, correspondencia con los hechos y relevancia.

En relación con la fase de preparación, se formulan objetivos operativos, se confirma la originalidad del ángulo, se desarrolla una comprensión personal del tema y se reúne información potencial. En su conjunto, estas actividades se relacionan con la actividad de buscar información.

En la literatura anglosajona, ubicamos, el trabajo de Julia Tylor (2014) quien no hace un estudio sobre comportamiento propiamente, sino que se centra en la evaluación de la credibilidad de las fuentes. Ella analiza la forma en que los estudiantes de periodismo de la Southwestern University en Texas, buscan información en línea y centra su atención en la evaluación de la credibilidad de los recursos. Para la recolección de los datos utiliza el análisis de contenido de videos que registran las búsquedas de los estudiantes y el análisis de las entrevistas semiestructuradas. Uno de los hallazgos de esta investigación es que, si bien los estudiantes de periodismo comprenden la importancia de la verificación, se basan en gran medida en los motores de búsqueda y confían en la credibilidad de los resultados (particularmente Google). En varios casos se observó que los estudiantes confían en la primera página de resultados, incluso en los primeros registros de esa página.

En el caso de latitudes como Kuwait y Pakistan encontramos los trabajos de Anwar, Ansari y Abdullah (2004); Anwar y Asghar (2004) y Chaudhry y Al-Sagheer (2011), respectivamente. En el primer caso se analiza el comportamiento informativo de los periodistas valiéndose de una encuesta. Los resultados evidencian la importancia de la verificación de los datos y la obtención de información para darle contexto a una nota. En gran medida esta investigación se centra en las fuentes de información que utiliza el periodista en su quehacer cotidiano. Se habla de un uso frecuente de Internet, boletines de prensa y entrevistas directas. Al final se revaloran las competencias para la búsqueda, sistematización, uso y análisis de los datos. En el segundo caso, se evidencia la importancia de la verificación de datos y la preferencia de recurrir a la fuente directa (humana), otros diarios y las agencias de noticias para la conformación de un producto noticioso. Del mismo modo, se hace énfasis en la necesidad de mejorar las competencias de búsqueda de información. Por último, la investigación de

Chaudhry y Al-Sagheer (2011) tiene como objetivo analizar el comportamiento informativo de los periodistas en Kuwait en los procesos iniciales de búsqueda de información. Se utiliza la perspectiva del incidente crítico para recopilar información y se discuten los resultados con los participantes por medio de un grupo enfoquehan centrado. En la mayor parte de los casos se descubrió que el recurso más utilizado es Internet, pero se reconoce que hay sitios que pueden no ser objetivos para el trabajo periodístico.

En Bangladesh Hossain y Shariful (2012) estudian el comportamiento informativo de periodistas de medios impresos. Sus objetivos fueron identificar las necesidades de información de los periodistas de medios impresos; identificar los recursos de información que utilizan para satisfacer sus necesidades y explorar los servicios que una biblioteca o centro de información le puede ofrecer a esta comunidad. En este caso el estudio es relevante por el contexto pues, si bien el gobierno de Bangladesh señala que hay una libertad total de prensa, sólo el 5 por ciento de los participantes está de acuerdo con esta visión. Como se observa, en este caso el ejercicio periodístico se ve limitado por las condiciones de trabajo. Este estudio revela la preferencia de trabajar con recursos en Internet, así como boletines, televisión y radio. Del mismo modo, señala limitaciones de tiempo, de competencias, de acceso a recursos de información internacionales y la falta de capacitación de bibliotecarios que atienden los centros de información a los que recurren los periodistas en este país.

En Iberoamérica hay trabajos como los de Calderín (2009) que analiza los procesos de búsqueda de información de los periodistas con la perspectiva teórica del *Sense-Making* lo que le permite analizar las necesidades físicas, cognitivas, sociales o afectivas del periodista durante este proceso. En el caso de autores como López Yepes (2006), Guallar Delgado y Cornet Casals (2004) han centrado sus estudios en las características y gestión de los servicios documentales en los medios impresos y audiovisuales. Por otro lado, hay quienes centran sus estudios en la transformación de la figura del periodista tras la integración de recursos electrónicos en su quehacer profesional (Marcos Recio, 1998; Aquesolo Vegas, 1996). En ese sentido, se ha resaltado la reconfiguración del rol de periodista frente

a la digitalización de los medios, pues se advierte que deben contar con conocimientos de documentalista o bien, el documentalista debe tener conocimientos sobre el quehacer periodístico (ver, Micó-Sanz, *et. al.* 2009). Por último, encontramos un trabajo que se centra en el uso de bases de datos automatizadas en las redacciones y analiza el tránsito del periodismo tradicional al periodismo digital (López-García, et. al. 2009).

Tras la revisión de los estudios que se han realizado sobre el comportamiento informativo de periodistas observamos que una de las técnicas más utilizadas es la entrevista; estudios como el de David Nicholas, Peter Williams, Helen Martin y Peter Cole (2000) utilizan esta técnica, junto con la perspectiva del estudio de caso, para analizar la introducción de Internet en tres diarios *The Independent*, *The Sunday Times* y *The Guardian*. En el caso de Poteet (2000), su objetivo es reunir información descriptiva para analizar los patrones del comportamiento informativo de los periodistas frente a los recursos en línea. Para ello diseña una encuesta basada en los indicadores de Nicholas y Williams (1997) y la aplica a periodistas de dos diarios *News & Observer* y *Charlotte Observer*. Por otro lado, el estudio realizado por Attfield, Blandford y Dowell (2003) se basa en una serie de entrevistas a periodistas del diario *The Times*, para identificar las actividades que realizan los periodistas para cubrir sus necesidades de información. Por último, Tylor (2014) utiliza el método de entrevista, la observación y la videograbación para analizar los procesos de validación de información por parte de periodistas. Como se observa en este último caso, el empleo de varias técnicas de recolección de datos permite validar la información y sus interpretaciones, lo que da pie a contar con un panorama multidimensional del comportamiento de los participantes.

LA TEORÍA FUNDAMENTADA EN EL ESTUDIO DEL COMPORTAMIENTO INFORMATIVO

Teoría Fundamentada: concepto e implicaciones

El descubrimiento de la teoría desde los datos es, en esencia, la definición de Teoría Fundamentada (TF) que formularon Glaser y Strauss (1968) en su trabajo seminal *The Discovery of Grounded Theory*. En su momento, este texto representó una invitación a generar nuevas teorías, en lugar de poner a prueba las prescripciones previas, y a privilegiar los datos en lugar de las teorías. Más adelante, en su trabajo *Theoretical Sensitivity*, Glaser (1978) consolida las bases conceptuales de la TF, habla de la necesidad de inhibir teorías en el momento de analizar datos empíricos, analiza el rol de la literatura previa y de la inducción durante el proceso de investigación e introduce los fundamentos para el análisis relacional de códigos de datos, entre otras aportaciones.

La TF representa un reto constante para el investigador, pues busca inhibir de forma temporal los prejuicios y antecedentes teóricos para valerse de una mirada nueva para recabar y analizar los datos de la investigación en curso. Muchas veces, al referirse al uso de la TF en una investigación, se suele pensar que el investigador debe dejar a un lado todo lo que sabe de un tema, desaprender lo aprendido, olvidar

sus lecturas y desvincularse de cualquier modelo que explique una realidad concreta. En contraste, autores como Soneira (2006) señalan que incluso la revisión progresiva de literatura relevante para la investigación y la revisión paralela de datos emergentes que surgen de la investigación es un procedimiento que se incorpora a la recolección de datos, la codificación y la redacción de memos. Mantener esa distancia, esa inhibición temporal de teorías latentes, esa sensibilidad para ubicar nuevos hallazgos, realizar comparaciones, realizar memos teóricos o conceptuales, esa posición crítica de lo construido y la forma de construir puede ser vista como un nivel de ignorancia óptima y vital durante todo el desarrollo de la investigación, sobre todo en el proceso de análisis, de construcción de códigos o categorías. Fases que, como se advierte, operan a lo largo de toda la investigación. En otras palabras, la TF requiere de un involucramiento activo y un posicionamiento crítico del papel que se tiene como investigador en la recolección y el análisis de datos.

Lo anterior se integra en esta investigación como un proceso de autorreflexión permanente en el que se reconoce la carga subjetiva que se permea desde el momento de la compilación de datos hasta su interpretación. Esta reflexión representa un esfuerzo intelectual por desprenderse o inhibir preceptos para dar pie al diálogo con los datos.

Todo este contexto toma relevancia en un campo de estudios tan fértil como el del comportamiento informativo, donde hay comunidades nuevas por explorar que tienen particularidades que no se asemejan a ningún otro caso, pues además están ligadas a un contexto y a un universo de creencias muy particulares. En ese sentido, la TF es una metodología que nos permitirá focalizar y ubicar nuevos hallazgos más que comprobar si un precepto teórico se cumple o no. Si consideramos las áreas que trabajan directamente con información, no limitándonos sólo a contextos laborales o académicos, encontramos que hay un terreno amplio para el desarrollo de nueva teoría alrededor de esta área de investigación.

Aplicación de la Teoría Fundamentada en el estudio del comportamiento informativo

Cuando se trata de resaltar las bondades de los métodos cualitativos, se habla de su pertinencia para analizar temas o vetas poco exploradas, así como del enriquecimiento de la teoría desde lo particular y su contexto. El terreno de los estudios de comportamiento informativo es un campo fértil para este tipo de métodos, no sólo por la naturaleza de su objeto de estudio sino también por el enfoque y los resultados que se persiguen. Una de las publicaciones fundacionales del método cualitativo en los estudios de la información es la revista *Social Science Information Studies* que en 1981 dedicó un número especial en el que se resalta el potencial de los métodos cualitativos en los estudios de la información y que, de acuerdo con Ellis (1993), representó una plataforma para que otros investigadores exploraran alternativas metodológicas para analizar el uso de la información y generar estudios cualitativos en esta área. Esta publicación es considerada como el pilar que allanó el camino para la introducción de la TF en los estudios de la información, pues instó a generar estudios con perspectiva cualitativa y bases fenomenológicas. Este tipo de estudios tuvo mayor eco en la Universidad de Sheffield (Wilson, 1980), y el mayor cambio de perspectiva que supuso fue persuadir a que las personas dejaran de ser un objeto de estudio en las investigaciones y se convirtieran en participantes (Wilson, 1990).

En varios de los modelos que se compilan en el libro *Theories of Information Behavior,* (2005), se utilizan aproximaciones cualitativas; un ejemplo de ello es el modelo de Ellis (1993) que emplea la TF para contrastar los patrones de búsqueda de información de investigadores académicos en Ciencias Sociales, Ciencias y Humanidades. En su análisis, Ellis (1993) asume que, aunque el uso de métodos cuantitativos ofrece la posibilidad de obtener una instantánea del uso de la información por parte de los usuarios, no representa la mejor forma de aproximarse a la percepción del usuario en las actividades que realiza durante la búsqueda de información. En su argumentación se opone al modelo de búsqueda de información en el que el usuario reconoce una necesidad y acude al sistema para obtener una respuesta basada en dicha necesidad. Ellis utiliza el término "*information man*" para

referirse a los modelos de comportamiento informativo que él considera primitivos y simplistas; es decir, aquellos que sólo analizan la respuesta que reciben de un sistema. En este sentido, el autor plantea la construcción de un modelo más realista de la búsqueda de información que, como tal, tendría que construirse desde cero y a partir del entorno (Ellis, 1993).

Es importante mencionar que uno de los primeros trabajos que utilizó la TF para el estudio del comportamiento informativo fue la tesis realizada por Ellis (1987), la cual analiza los patrones de búsqueda de información de científicos sociales y tras establecer procesos de comparación constante, ofrece un modelo útil para el diseño de sistemas de recuperación de información que se caracterizan en los siguientes pasos: comenzar, entrelazar, navegar, distinguir, monitorear y extraer. El propio Ellis (1987) reconoce que fue el análisis de las actividades de búsqueda de información lo que le permitió generar y ampliar nuevas categorías, en lugar de tomar un conjunto de supuestos *a priori* relacionados con estas actividades.

En el 2006 Mansourian realizó una revisión exhaustiva de la penetración de la TF en los estudios de la información. En su recuento resalta los trabajos de Seldén (2005), Mellon (1986), Mansourian y Ford (2006) y abunda en el trabajo de tesis de Pace (2003). Este último desarrolla una teoría que caracteriza elementos contextuales, como la curiosidad y la urgencia, presentes durante la interacción de los usuarios con la web en un contexto de búsqueda de información. Pace (2003) advierte que la mayor parte de los hallazgos provienen del análisis de los datos recabados y analizados a diferencia de lo que ocurre en otros estudios donde los hallazgos tienen su origen en el contraste de teorías preestablecidas. Asimismo, profundiza en el concepto de “flujo” o “*flow”* para referirse a un estado que experimentan los usuarios cuando se adentran de forma profunda en una actividad y tienen clara la tarea y el objetivo de su búsqueda, así como la capacidad de control, pérdida de autoconciencia, tiempo distorsionado y experiencia autotélica, es decir la recompensa o incentivo de realizar determinada tarea deriva del mismo acto de hacerla.

Uno de los textos más críticos con relación al uso de la TF en los estudios de la información es el trabajo de Seldén (2005), quien a partir de su propia experiencia

y de la lectura directa de los textos de Glaser y Strauss (1968) ofrece un balance de las ventajas y desventajas de utilizar esta teoría en este campo. Su atención se centra en tres aspectos:

- La posición del investigador en relación con las preconcepciones, con los cuestionamientos: ¿Es posible planear un proyecto e ignorar la literatura en cuestión?, ¿Cómo se puede lidiar con un deseo natural de avanzar en los hallazgos de los demás? Si, contra las expectativas, se logra mantener una distancia de las preconcepciones obvias en la literatura, aún existe el riesgo de asimilar inconscientemente otras preconcepciones más elusivas.
- Los problemas que rodean el tipo de conocimiento producido, en el caso de Seldén, ponen en evidencia que la teoría de búsqueda de información que él formuló en su investigación se originó durante esta indagación en que trató de apegarse a los preceptos de la TF.
- Problemas en la codificación, como el fanatismo en el procedimiento y la desconexión del contexto. Durante el proceso de codificación se corre el riesgo de desvincular los datos del contexto y desatender la creatividad, esto ocurre, según Seldén, al utilizar programas para el análisis cualitativo.

En el caso de Mansourian y Ford (2006) analizaron la percepción del concepto “web invisible” entre usuarios académicos, así como sus experiencias normales y fallidas en la búsqueda de información en internet. Este estudio se vale de la técnica de incidente crítico y de entrevistas semiestructuradas a quince académicos de la Universidad de Sheffield. Sin hacerlo explícito, utilizan la TF para el análisis de los datos, pues no trabajan con un conjunto de ideas preestablecidas, sino que usan la comparación constante y la interpretación emergente para establecer y relacionar las categorías que dan forma a un modelo.

Por su parte Mellon (2015) utiliza la TF para explorar los sentimientos que experimentan un grupo de estudiantes al utilizar por primera vez una biblioteca académica para sus investigaciones. Uno de los instrumentos que empleó para el

acopio de la información fue la elaboración de un ensayo alrededor de cuatro preguntas que solicitó a los alumnos al final de un curso de redacción: 1) ¿Cuál fue tu experiencia al utilizar la biblioteca para buscar información para tu estudio?, 2) ¿Cómo te sientes con respecto a la biblioteca y tus competencias para utilizarla?, 3) ¿Estos sentimientos cambiaron a lo largo del semestre? y 4) ¿Cómo te sientes al utilizar la biblioteca ahora? Los resultados se analizaron por medio del método de comparación constante para ubicar temas recurrentes. Encontró que de un 75 a 85 por ciento de los participantes utilizaban términos como ansiedad, miedo, fobia y confusión para referirse a su primera experiencia de búsqueda de información en la biblioteca.

Otra aplicación de la TF en el estudio de usuarios es la tesis realizada por Brahme (2010), que analiza las diferencias de comportamiento informativo entre estudiantes de doctorado residentes y estudiantes a distancia. En este caso, el estudio recaba información por medio de entrevistas semiestructuradas a dos grupos de alumnos que trabajaban de forma activa en la búsqueda de información para sus respectivos temas. A partir del análisis de resultados se identificó que los estudiantes a distancia tienen un mayor contacto con los bibliotecarios, confían más en sus propias habilidades y experimentan aislamiento y soledad durante el proceso de tesis.

Una perspectiva novedosa para analizar la forma en que las personas enmarcan un problema y definen sus necesidades de información es el estudio de Naumer (2012), que utiliza el concepto de "*framing*" o "encuadre" para aplicarlo a los conceptos inherentes de necesidad, búsqueda y uso de la información. Esta teoría (frame-based theory) explica la conducta informativa en relación con cuatro conceptos: marcos, necesidades de información, información y contexto. Los marcos son vistos como una representación del conocimiento que ofrece una forma de entender las necesidades de información en un contexto y dotar de sentido a problemas complejos. El estudio basa sus premisas en la TF para el desarrollo de una teoría generada a partir de los datos. Naumer trabajó con distintos grupos, como activistas políticos, urbanistas planificadores de jubilación y personas que trabajan con pandillas. Para la descripción de resultados, utilizó la técnica de viñetas, en las

cuales provee un panorama de cada participante y sus circunstancias particulares, lo que permite explicar la perspectiva de cada uno en sus propias palabras. El concepto de “frame” o “marco” con el que trabaja este autor representa la forma en que una persona concibe un problema y sus diferentes aspectos, éstos a su vez están relacionados con otros elementos, como sentimientos, prejuicios e interpretaciones. Es, en palabras del propio Naumer (2012), una forma de entender lo que una persona procesa para dar sentido a un problema en situaciones o contextos específicos.

Por su parte, Hsu (2014) busca caracterizar las razones por las que se selecciona una fuente de información en internet con el fin de ganar ventajas competitivas para el diseño de páginas web. Con este objetivo, el autor crea un modelo de selección de fuentes de información desarrollado a partir de los resultados que obtiene de encuestas en línea y recomendaciones de páginas web de información local sobre salud y productos. El análisis de los datos basado en la TF arrojó que la herramienta más popular para la información local son los motores de búsqueda y los factores que determinan esta preferencia se justifican por la experiencia directa, la costumbre de utilizar motores de búsqueda, la facilidad de uso percibida y la confiabilidad.

De acuerdo con González-Teruel y Abad-García (2012), la TF ha sido empleada para la construcción de teorías sobre el comportamiento informativo en las últimas décadas por dos factores principales: la naturaleza inductiva de la TF, pues (1) se privilegia la observación del usuario en su contexto sin depender de categorías preestablecidas y (2) su orientación a generar teorías a partir de datos, evitando descripciones superficiales de la interacción entre el usuario y la información. Estos autores también advierten que si bien hay un cierto grado de flexibilidad en la forma en que se aplican los procedimientos de la TF, se debe acompañar de documentación precisa que incluya los procedimientos utilizados para la recopilación, el análisis y la interpretación de la información proporcionada por los usuarios.

En resumen, y de acuerdo con la revisión de la literatura que se ha expuesto, es posible identificar que el estudio del comportamiento informativo ha requerido de

enfoques metodológicos cualitativos para analizar la complejidad de las situaciones y los contextos en los que se originan y desarrollan estos tipos de comportamientos. Esto es perceptible en los hallazgos reportados en los estudios analizados (ver, Cuadro 1).

CUADRO 1: CATEGORÍAS TEÓRICAS DE CONDUCTA INFORMATIVA IDENTIFICADAS A PARTIR DE LA TEORÍA FUNDAMENTADA

Autores	Comunidad	Categorías teóricas identificadas
Ellis, 1987	Científicos de sociales, ciencias y humanidades.	Comenzar, entrelazar, navegar, distinguir, monitorear y extraer.
Pace, 2003	Usuarios de la web.	Curiosidad y urgencia, capacidad de control, pérdida de autoconciencia, sentido del tiempo distorsionado y experiencia autotélica.
Brahme, 2010	Estudiantes de doctorado residentes y a distancia.	Confianza en sus habilidades, sensación de aislamiento y soledad.
Naumer, 2012	Activistas políticos, urbanistas, planificadores de jubilación y personas que trabajan con pandillas.	Sentimientos, prejuicios e interpretaciones.
Hsu, 2014	Usuarios de páginas web.	Preferencias, costumbres, facilidad de uso y confiabilidad.
Mellón, 2015	Estudiantes que usan por primera vez bibliotecas universitarias.	Ansiedad, miedo, fobia y confusión.

Como se puede observar en el Cuadro 1, las categorías teóricas identificadas se refieren a aspectos que requieren un análisis profundo de las expresiones, sentimientos, dudas, conocimientos de las personas en relación con la actividad y la tarea a desarrollar dentro de una comunidad. Estos aspectos son difíciles de

identificar por medio de metodologías tradicionales, pues se requiere de la interacción y el entendimiento personal. Además de ello, exige que el investigador adopte una postura crítica ante los hallazgos previos a fin de dar paso a nuevos patrones derivados de los cambios sustantivos y complejos que suceden en las organizaciones.

La presencia de la TF, como lo señala Ellis (1987, 1993), puede ser una respuesta a los modelos primitivos y simplistas de CI que no integran los procesos psicológicos y la naturaleza de los comportamientos que ocurren en contextos reales. Como observa Naumer (2012), la TF permite la identificación de los marcos que son necesarios para comprender la complejidad de un problema y dar sentido a las representaciones del conocimiento de las personas cuando están resolviendo problemas.

De acuerdo con Seldén (2005), el investigador que aplique la TF debe estar consciente y preparado para afrontar la dificultad para diferenciar sus aportaciones y hallazgos de los que han sido presentados previamente por otros investigadores. Esto es más difícil si se considera que por lo regular los investigadores ya han realizado trabajos preliminares, leído la literatura sobre el tema, tomado clases o acudido a congresos; es decir que cuentan con antecedentes, conocimientos y experiencias previas. Pero, por otro lado, la carencia de éstos puede llevar a un análisis deficiente de los datos; entonces, el reto para el investigador es lograr autorregularse, evitar prejuicios y estar abierto para dar paso a los nuevos hallazgos.

Otro problema que identifica Seldén es sobre la consistencia y validez a los hallazgos derivados de la TF, y cómo se les vincula con otros estudios. Por último, están los métodos de análisis comparativo, los cuales deben ser rigurosos y contextualizados, pero sin caer en el fanatismo, es decir en el apasionamiento desmedido hacia los datos.

A partir del análisis de cómo ha sido utilizada la TF en los casos expuestos y de las consideraciones ya mencionadas, es que se juzga pertinente emplear este método para el estudio del comportamiento informativo de los periodistas de datos, en el entendido de que sus actividades y tareas se generan en contextos y

situaciones específicas, las cuales son dinámicas y complejas. Asimismo, se observa también la necesidad de encontrar nuevos conceptos que nos permitan abonar a favor de una teoría sobre el comportamiento informativo de esta comunidad, la cual, al ser tan reciente, aún sigue construyéndose.

Valor agregado y procesos cognitivos en el estudio del comportamiento informativo

Este apartado tiene por objetivo clarificar el concepto de valor agregado en los términos en los que fue originalmente concebido por Taylor, para luego revalorizarlo en el contexto actual de los estudios de la información. La obra seminal del autor que aborda este tema es *Value-added processes in information systems*. En la obra se definen los procesos de valor agregado como “las actividades en los sistemas de información que proveen de mecanismos para identificar su potencialidad para relacionarse con un problema o ambiente específico” (Taylor, 1986: 17).

Para Taylor varias interpretaciones de valor agregado, una de ellas es la que responde a la pregunta “¿Qué características o atributos deben ser adicionados a los datos y a los ítems de información mediante procesos que se realizan para que sean más útiles (con más beneficios) para los usuarios?” (1986:18). Otras preguntas que se añaden a esta interpretación son: ¿cuánto valor es necesario agregar?, ¿a qué costo?, ¿con qué beneficios?, ¿en qué nivel y si este valor se depreciará con el tiempo? Taylor diseña un “espectro” de valor agregado el cual puede apreciarse en la Figura 3.

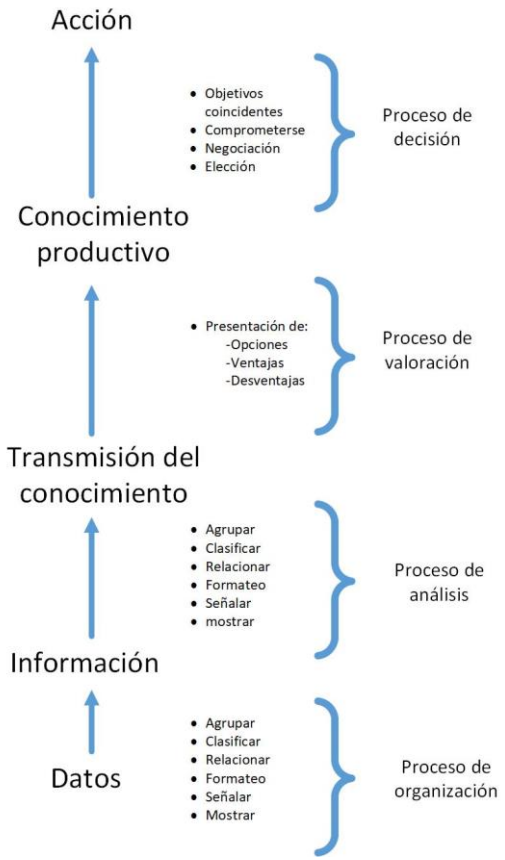


FIGURA 3: ESPECTRO DE VALOR AGREGADO
Fuente: Taylor, R. (1986). *Value-added processes in information systems*. Norwood, Jersey. (Traducción propia.)

El espectro se lee de abajo hacia arriba, cuenta con tres columnas, la primera detalla los productos de valor agregado que van del dato a la acción. En la tercera columna se presentan las categorías de procesos de valor agregado (PVA), que incluyen la organización de la información, el análisis, enjuiciamiento (valoración) y la toma de decisiones. En la columna central se pueden diferenciar los PVA específicos para transformarse de datos a acciones.

Los datos son considerados como símbolos (números, números en tablas y hechos no numéricos) que se almacenan en memorias de computadoras; sin embargo, éstos no son útiles si no se procesan para ser seleccionados. Lo que no significa que no tengan valor *per se*, si no que mediante los procesos de valor agregado las características y el contenido de los datos y de los documentos son más perceptibles a los usuarios.

Esto es lo que, en parte explica el alto costo de las bases de datos ya que prometen, como lo señala Taylor, aumentar el éxito en la recuperación de la información al proporcionar los mejores resultados para facilitar la evaluación y la selección de la información en la toma decisiones y en la clarificación de las opciones para solucionar un problema. En este caso hay varios procesos previos que garantizan que esta recuperación o ubicación de información se realice con éxito.

No obstante, la evaluación final de los beneficios del valor agregado es subjetivo y depende en buena medida de la percepción de los usuarios (Taylor, 1986: 13):

> "El valor de la información que contiene el mensaje y lo que significa, y si influye o no en su decisión (tarea o actividad).
>
> El valor de los recursos de información (servicios, tecnologías y sistemas), que almacenan, procesan, analizan, empacan y proveen de los mensajes".

Por lo anterior Taylor (1986:7) apunta que el costo real de un producto podría medirse en función de la energía, el tiempo, la pérdida de oportunidades y de dinero que invierte el usuario para recuperar información relevante y seleccionarla para los contextos y situaciones en los que serán usados o combinados para crear nuevos patrones (ver, Hall, 1981)

De acuerdo con lo anterior, los procesos de valor agregado (PVA´s) a la información retoman los flujos y los usos de la información que realiza un ser humano o una organización en un contexto específico (Taylor, 1986). La organización provee el contexto, las tareas y las responsabilidades en la solución de problemas; por ello es necesario identificar:

> 1. La estructura, propósito y la dirección de la organización y como afecta el comportamiento informativo.
>
> 2. La variedad de personas que trabajan en la organización quienes, de acuerdo con sus capacidades y sus responsabilidades, requieren diferentes tipos de información.
>
> 3. Las dimensiones de los problemas los cuales son generados por esas personas en las cuales reflejan los aspectos específicos en los que ellos trabajan" (Taylor, 1986: 35).

A partir de lo anterior, podemos afirmar que las organizaciones son las que determinan los fines y usos de la información y las actividades y metas de ésta. Es debido a esto que los flujos formales de la información y las estructuras de las relaciones humanas del trabajo afectan el comportamiento informativo (Taylor, 1986: 36):

> "Dentro de la colección interna de información, el flujo de la información debe responder a la siguiente pregunta: ¿Qué es lo que nosotros hacemos?, y ¿cómo lo hacemos? Un punto crítico es quién decide y qué es lo que se colecta y cómo se estructura. En otras palabras "¿quién es el propietario de la información?" (Taylor, 1986: 37).

Taylor señala que las organizaciones pueden referirse a comunidades que se agrupan por disciplinas o profesiones, como es el caso de los científicos. Para ejemplificar lo anterior, el autor describe cómo es que usan la información los científicos y los ingenieros. A partir de lo anterior se puede identificar que hay dos elementos esenciales en los PVA´s: el usuario y el contexto. El tercer elemento corresponde a la tecnología, la cual tiene dos funciones básicas, una es la de proveer de una interfaz para la negociación de significados entre el usuario y el sistema; y la segunda es ser la plataforma para el procesamiento de los datos.

En el caso de los usuarios los factores más importantes de análisis son los criterios que inciden en la selección de la información, y cómo estos se pueden retomar para agregar valor a la interfaz y al sistema de procesamiento de información (ver, Cuadro 2).

CUADRO 2: CRITERIOS DE LOS USUARIOS Y VALOR AGREGADO

Criterios de selección del usuario	Interfaz (valor agregado)	Sistema de procesamiento (ejemplos de valor agregado)
Facilidad de uso	Hojear Formateado Interfaz 1 (mediación) Interfaz 2 (orientación) Ordenamiento Accesibilidad física	Alfabetización Términos relevantes
Reducción del ruido	Acceso (identificación del ítem) Acceso II (resumen de materia) Enlazamiento Precisión Selectividad	Indexación Control de vocabulario Filtrado
Calidad	Actualidad Comprehensividad Vigencia Confiabilidad Validez	Control de calidad Edición Actualización Análisis y comparación de datos
Adaptabilidad	Cercanía con el problema Flexibilidad Simplicidad Estimulante	Previsión de la capacidad para manipular datos Clasificación de salida por relevancia
Tiempo salvado	Rapidez de respuesta	Reducción del tiempo de procesamiento
Ahorro de costo	Ahorro de costo	Baja conexión-precio del tiempo

Fuente: Taylor, R. (1986). *Value-added processes in information systems*. Norwood, Jersey, p. 50.

En el Cuadro 2 se observa cómo un criterio de selección se retoma como parte de la interfaz y de los procesos internos que realiza el sistema. Como lo propone Taylor, el análisis de los PVA´s puede darse en organizaciones como las bibliotecas (escolares, académicas, de investigación o públicas). En la tabla titulada “Values added by library types” (Taylor, 1986: 95), el autor resume los procesos de valor agregado que son aplicables a cada tipo de biblioteca. En los siguientes capítulos de su obra aplica los PVA’s a los servicios de indexación y abstracción y presenta unas tablas en las que muestra los PVA’s dirigidos a la adquisición, control, citación, indexación y servicios de abstracción, mismos que detalla en el apéndice B de su libro.

El mérito de un sistema de información, de acuerdo con Taylor es que incremente el éxito en la recuperación de la información que provea de los mejores resultados para facilitar la evaluación y la elección de los sets o conjuntos de datos, que faciliten la toma de decisiones, las soluciones y la clarificación de las opciones.

El modelo de Taylor es también uno de los primeros en tratar aspectos relacionados con componentes cognitivos centrales, los cuales son la base del comportamiento informativo.

El término cognitivo se aplica de diversas formas en las ciencias de la información, sin embargo, hay, según (Allen, 1991) una aceptación general de concebir este término como “objetos” mentales, es decir: conceptos, ideas, conocimientos y eventos que no son directamente observables. El término también ha funcionado de acuerdo con Todd (1999) como una metateoría que concibe al usuario como una persona que busca e interpreta la información.

En su momento, Brookes (1975) resaltó la importancia de comprender las interacciones entre las personas y la información, sentó además las bases para una conceptualización dinámica de la perspectiva cognitiva en las ciencias de la información y un marco para su investigación empírica. Otros estudios que han retomado esta perspectiva son los de De Mey (1982), quien se refería a la existencia de un marco, una imagen o una visión del mundo. Asimismo, los estudios sobre el uso de la información de Guthrie (1988) son de los primeros en involucrar a los procesos mentales.

Pettigrew, Raya y Harry (2001) definen este enfoque como un conjunto de constructos teóricos para la comprensión del comportamiento informativo que se centra en los atributos de un individuo y que tiene como eje principal el concepto de estructura del conocimiento. Éste consiste en un conjunto de conceptos relacionados que constituyen el modelo o visión del mundo de cada individuo, conceptos que se ponen en juego durante la percepción, interpretación, modificación y transferencia de información. Esta estructura del conocimiento es determinada, de acuerdo con De May (1982), por las experiencias colectivas, la educación o la capacitación.

En ese sentido, Pettigrew y Raya (2001) señalan que las investigaciones que trabajan con este enfoque cognitivo están encaminadas a estudiar la forma en que un individuo aplica su modelo o visión del mundo para procesar, definir, buscar, ofrecer y utilizar información y que ha sido útil para identificar patrones de comportamiento aplicables al desarrollo de sistemas de recuperación de

información. De hecho, trabajos como los de Ellis (1987) y Dervin (1983) son buenos ejemplos de cómo se ha aplicado este enfoque en estudios sobre comportamiento informativo, pues identifican procesos cognitivos claramente diferenciados que intervienen en la interacción con la información. Otro estudio considerado como un referente obligado en esta materia es el realizado por Kuhlthau (1991), el cual integra la perspectiva de niveles de necesidad planteada por Taylor (1991) y la anomalía en el estado del conocimiento de Belkin, Oddy y Brooks (1982) para dar forma a un modelo de búsqueda de información que sintetiza elementos físicos, afectivos y cognitivos.

Asimismo, se ubica el caso de Dinet (2014), quien señala que en la recuperación de la información hay aspectos psicológicos dados que involucran: a) actividades humanas, b) actividades cognitivas (memorización, tomar decisiones, entre otras) y c) procesos iterativos derivados de la reflexión sobre lo que hacen.

Por su parte, Wilson (1984) señala que el objetivo de la perspectiva cognitiva en el análisis del comportamiento informativo es descubrir cómo las imágenes o los marcos de referencia de las personas que se encuentran en procesos de disponibilidad y elección de información está determinada por el propio marco de referencia y cómo la incorporación de nueva información puede afectarlo o modificarlo. Una parte importante del discurso de la perspectiva cognitiva en el comportamiento informativo está asociada con los preceptos del construccionismo social (Savolainen, 2007). Lo que supone que los procesos de búsqueda de información se relacionan con el contexto del cual derivan prácticas que pueden ser vistas como naturales y dadas para quienes las realiza. Desde otra perspectiva se pueden definir como un conjunto de disposiciones del campo que se reproducen en formas de hechos o acciones por hacer, en un contexto particular (ver, Bourdieu, 2016).

MÉTODO

En este capítulo se describe el proceso realizado para el acopio y sistematización de la información, así como la forma en que se han construido los códigos que emergen del análisis de los datos. La perspectiva metodológica desde la cual se plantea la investigación es la cualitativa, pues nos permite analizar la riqueza de datos y comprender de una forma explicativa la complejidad de los procesos asociados. Asimismo, proporciona mejores garantes en la aproximación al objeto de estudio y en la comprensión de los procesos de transformación de la información. El periodismo de datos implica la toma de decisiones para resolver una serie de requerimientos durante la elaboración del producto. En ese sentido, nos interesa tener, hasta donde sea posible, una comprensión profunda de las motivaciones que están detrás de las decisiones en los procesos de búsqueda de información, selección de fuentes, evaluación de los datos y su uso.

A decir de Maxwell (1996) la investigación cualitativa se puede emplear para: 1) comprender los significados que los actores dan a sus acciones, vidas y experiencias y a los sucesos y situaciones en los que participan; 2) comprender un contexto particular en que los participantes actúan y la influencia que ese contexto ejerce sobre sus acciones; 3) identificar fenómenos e influencias no previstos y

generar nuevas teorías fundamentadas en ellos; 4) comprender los procesos por los cuales los sucesos y acciones tienen lugar; 5) desarrollar explicaciones causales válidas analizando cómo determinados sucesos influyen sobre otros, comprendiendo los procesos causales de forma local, contextual o situada. Wilson (1981) resalta la importancia de la investigación cualitativa en los estudios de usuarios, pues apoyan en la comprensión de las necesidades que motivan a una persona a realizar una búsqueda de información y con ello ampliar la comprensión del significado de la información en la vida cotidiana de las personas.

De esta manera es posible comprender el punto de vista y los elementos que pone en juego el periodista durante el desarrollo de un producto noticioso; el significado de sus acciones; la influencia del contexto en sus acciones y toma de decisiones. Para, en otras palabras, reconstruir una situación de uso de la información, desde la perspectiva de un tipo de usuario en particular.

En concreto, la perspectiva cualitativa que se aplica es la Teoría Fundamentada (TF), y una de sus ventajas es que nos permite un diseño flexible de la investigación. Esto posibilita también que, durante el desarrollo de la investigación, se incorporen los métodos y las técnicas necesarias para hacer frente a situaciones nuevas asociadas con el tema de estudio. El criterio empleado para la selección de grupos de participantes obedece al concepto de “muestra teórica” de Glaser y Strauss (1967), que supone que los participantes no se eligen con métodos estadísticos sino a partir de las necesidades que derivan de la construcción de los conceptos y con el objetivo de formarnos múltiples perspectivas sobre un fenómeno. Esto es para validar las categorías y relaciones emergentes de una muestra comparándola con una nueva con el fin de elaborar y reafirmar categorías (Charmaz, 2006) así como confirmar o rebatir las proposiciones anteriores tal como lo realiza Randall y Mello (2012).

El propósito de trabajar de esta forma es atender al diálogo e interacción de los propios datos con miras a ubicar hallazgos originales que emerjan de su análisis. La vocación inductiva planteada por la TF nos permitió generar conceptos e interrelacionarlos. Sin embargo, para habilitar este diálogo hay de por medio procesos exhaustivos de análisis y comparación que ajustan de forma granular el

conjunto de datos obtenidos, de tal modo que estos pasos permiten transitar a interpretaciones globales a partir de datos procesados, aglomerados y por tanto manejables.

Este proceso de transformación de los datos empíricos va del diseño de los instrumentos para el acopio de información, la interacción con los participantes en su contexto, el análisis de las transcripciones, las primeras lecturas abiertas, el establecimiento de códigos y categorías, la relectura y codificación de las entrevistas, la reclasificación de códigos y categorías. En este continuum el principal conductor de las decisiones tomadas a lo largo del estudio son el objetivo y pregunta de investigación del estudio.

En ese contexto, cabe hablar de una variable que supuso reajustes al planteamiento inicial de esta investigación, nos referimos al tiempo disponible de nuestros sujetos de estudio, pues el proceso de acopio, interpretación y análisis de la información recabada en las redacciones se vio limitada en su alcance por esta razón. Esta variable, como se verá más adelante, supuso decisiones que motivaron el replanteamiento de estrategias para el acopio de información.

Escenarios

La perspectiva cualitativa nos obliga a situarnos en el contexto de una realidad concreta con miras a registrar e interpretar las interacciones, prácticas y significados que se originan durante el desarrollo de una nota y que, en este caso, se vinculan con el comportamiento informativo de los periodistas de datos. Los primeros dos escenarios corresponden a dos diarios de circulación nacional y el tercero a una empresa dedicada a elaborar productos noticiosos para distintos medios. Como se puede advertir se trata de contextos diferentes. Nuestra intención es analizar los procesos que se siguen en entornos con condiciones y exigencias distintas, con el fin de maximizar sus similitudes y minimizar sus diferencias; y más adelante se procederá a la inversa, es decir se maximizarán sus diferencias y minimizarán sus similitudes. Con ello se busca enriquecer la caracterización de los procesos.

- Los dos primeros escenarios lo constituyen la redacción de dos diarios de circulación nacional que tuvieran en operación un departamento encargado de generar notas de periodismo de datos. Se identificarán en adelante como Escenario 1 (E1) y Escenario 2 (E2).
- El tercer escenario es una empresa que se dedica a generar contenidos de periodismo de datos para medios electrónicos. Esto supone el desarrollo de productos multimedia con distintas temáticas y trabajados por un equipo multidisciplinar. Se identificará en adelante como Escenario 3 (E3).

Las personas que participaron en la investigación son profesionales que colaboran de forma individual o grupal en la elaboración de notas de periodismo de datos en diarios o empresas de contenidos. Se trata de personas sin distingo de sexo con formación o experiencia en el campo de la comunicación y que han desarrollado notas de esta naturaleza con fines periodísticos, sin importar el rol que desempeñen, sea editor, diseñador, programador, analista o reportero.

Instrumento

Entrevista semiestructurada

El objetivo de este instrumento es reunir información sobre el contexto, la forma de trabajo de los participantes, su visión, motivaciones, flujos de trabajo, tareas, interacciones con la información; así como los elementos que ponen en juego durante el desarrollo de un producto noticioso; el significado de sus acciones; la influencia del contexto en sus acciones y los elementos que toma en cuenta para la toma de decisiones (ver, Cuadro 3). La guía de la entrevista (ver, Anexo 2) que fue empleada para obtener información de los participantes se utilizó en cada uno de los escenarios para el acopio de datos. Al igual que en el caso de Ellis (1993), se buscó que en las entrevistas los participantes hablaran de su trabajo, del tipo de cosas que hacían y de qué uso le dan a la información en su trabajo diario.

CUADRO 3: ENTREVISTA SEMIESTRUCTURADA

Instrumento	Tópicos	Variables de investigación
Entrevista semiestructurada	• Trayectoria personal -Formación -Forma en que se incorporan a las áreas -Tareas -Motivaciones • Contexto - (origen del área de trabajo) - Objetivo del área • Dinámica de trabajo - Flujo de trabajo - Condiciones y obstáculos - Roles • Reconstrucción del proceso de elaboración de un producto -Origen de un tema -Agenda y coyuntura -Desarrollo -Interacción con la información • Periodismo de datos -Conceptualización -Deber ser del periodismo de datos -Referentes de periodismo de datos	• Procesos cognitivos • Prácticas • Disposiciones • Resolución de necesidades de información • Interacción con la información • Interpretación y síntesis de la información • Valor agregado a la información

Obtención de datos

El ingreso a las redacciones fue precedido por una reunión inicial con los responsables de cada área, el contacto se dio por correo electrónico y en el caso del E1 se contactó a la editora telefónicamente por recomendación de la editora del E2. En cada escenario hubo una etapa de negociación para ingresar a las redacciones y realizar este estudio. En los escenarios se entregó una carta en la que se establecieron los términos del estudio en cuanto a manejo confidencial de la información proporcionada y el manejo de los resultados.

Una limitación del estudio fue la poca disponibilidad de tiempo por parte de los participantes. El tipo de trabajo que realizan además de ser demandante está

sujeto a la temporalidad propia de los medios de información, es decir se rige bajo valores de inmediatez y pertinencia. Lo cual supone una dinámica de trabajo que deja poco espacio para atender una investigación sobre los procesos que realizan durante el desarrollo de una nota. Esta circunstancia provocó el replanteamiento de realizar entrevistas individuales en el E2, pues no se contó con el tiempo necesario para su realización. La estrategia que se optó en este caso fue realizar una entrevista grupal para optimizar la economía de tiempo.

Muchos hechos marcaron este estudio y afectaron su seguimiento y resultados; entre ellos resaltan la desintegración del equipo del E2 a razón de una nota publicada en este diario. Esta situación marcó profundamente el curso de esta investigación junto con la salida temporal del principal informante del E3 que partió a estudiar al extranjero. El caso del E1 fue uno de los escenarios que brindó mayores facilidades para la realización de este estudio, muestra de ello fue la oportunidad de entrevistar dos veces con su editora y la facilitación de un reporte de actividades vinculado con una nota, la cual se integró como parte de la documentación de análisis de este escenario.

En total se realizaron 8 entrevistas, dos de ellas grupales. Participaron 16 personas (ver,

Cuadro 4). En aquellos casos en los que la entrevista fue grupal la razón de elegir esta dinámica, como ya se mencionó, obedece a limitaciones de tiempo por parte de los participantes. Se ofreció, en cada caso, una breve descripción de los propósitos de la investigación y las condiciones de su participación.

CUADRO 4: INSTRUMENTOS Y PARTICIPANTES

Participante	Escenario	Lugar	No. de participantes	Grabación	Duración
Entrevista semiestructurada:					
Analistas	E3	Redacción	3	Sí	42
Diseñadores	E3	Redacción	2	Sí	45
Programador	E3	Redacción	1	Sí	33
Editora	E2	Redacción	1	No	35
Equipo: analistas (3); reporteros (2); becaria (1); diseñador (1)	E2	Redacción	7	Sí	30
Editora	E1	Café	1	Sí	40
Editora (segunda entrevista)	E1	Redacción	1	Sí	28
Reportera	E1	Redacción	1	Sí	33

Procedimiento para el análisis y la sistematización de los datos

Para el proceso de codificación e interpretación de los datos se utilizaron elementos de la perspectiva de Teoría Fundamentada (Glaser y Strauss, 1967). En el caso de esta investigación el proceso inductivo fue orientado o motivado por la pregunta de investigación ¿Cuáles son los procesos que siguen los reporteros que se dedican al periodismo de datos, en relación con su comportamiento informativo, durante la elaboración de un producto noticioso de esa naturaleza?

Para facilitar la tarea de sistematización e interpretación de los datos se utilizó MAXQDA, un software comercial para la investigación cualitativa que permite sistematizar, integrar y analizar información proveniente de entrevistas, grupos focales, imágenes, audio, anotaciones de campo, fotografías y mapas.

Este tipo de herramientas, si bien facilita el proceso de sistematización y análisis, no exime las tareas de un investigador. Uno de los errores más comunes al momento de utilizar este tipo de herramientas es pensar que son programas que hacen el análisis que le corresponde al investigador. Como si se trataran de cajas negras en las que se introducen transcripciones y se recibe un análisis minucioso de los mismos. Autores como Kelle (1992) sostienen que los CAQDAS (del inglés Computer Assisted/Aided Qualitative Data Analysis Software) deberían llamarse “herramientas para el almacenamiento y recuperación”, pues si nos limitamos a las funciones más básicas de estos programas la definición resulta más fiel. Tal como

advierte Becker (1993) ningún programa de esta naturaleza es capaz, hoy en día, de ayudar al investigador a identificar las categorías centrales de su investigación.

Los pasos que se siguieron en esta fase de la investigación se apegan al modelo propuesto por Kuckartz y McWhertor (2014):

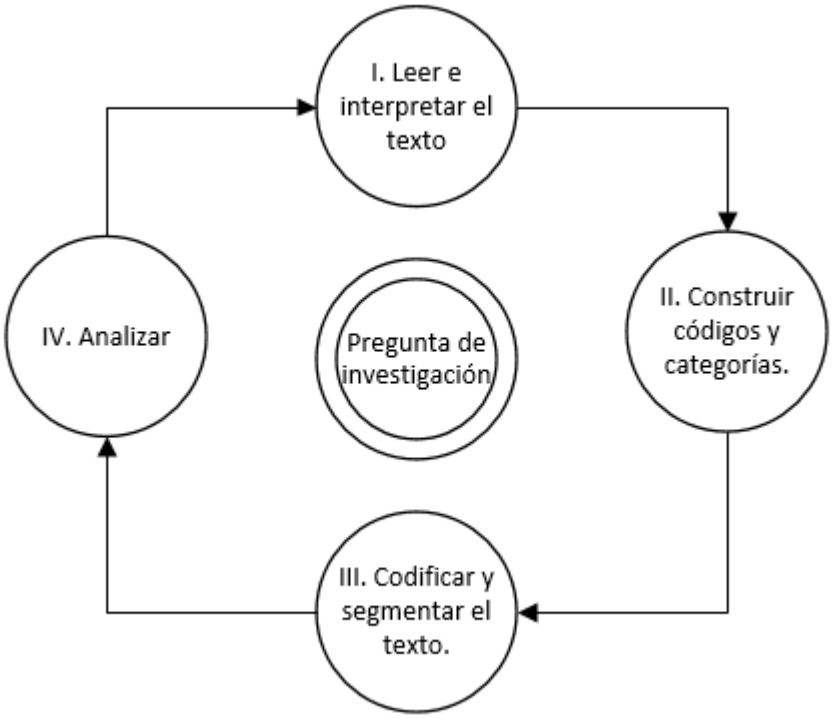


FIGURA 4: PROCESO DE ANÁLISIS DE LA INFORMACIÓN RECOPILADA. BASADA EN EL ESQUEMA DE KUCKARTZ Y MCWHERTOR (2014)

Como se observa en la Figura 4 el proceso no es lineal y supone una comparación constante de los datos recolectados, codificados, categorizados y analizados, mismos que se vinculan con la pregunta de investigación. De tal modo, que, a lo largo del estudio, estas fases siguen operando de forma simultánea.

Etapa I. Lectura e interpretación del texto

En este punto, la investigación sobre los datos obtenidos de los escenarios no se limita a las transcripciones de las entrevistas. Si bien representan el conjunto mayor de los datos de un contexto, no son exclusivos. Esto se explica por la interacción previa con los escenarios, las llamadas telefónicas, los correos e intercambios verbales no registrados en una grabación. Todo este panorama va marcando en buena parte la percepción del trabajo que se tiene como investigador sobre los participantes y su trabajo. Dicho esto, la tarea que se tiene no es sencilla, pues a lo largo de lo que ahora se presenta, hay un cuestionamiento constante sobre las percepciones y los hechos, la imposibilidad de la objetividad y la marca o huella del investigador en su estudio.

En este caso, el esbozo de los códigos comienza con la escucha activa de las transcripciones y el cotejo de las notas de campo. Este ejercicio permitió identificar palabras clave o sintagmas como un primer esbozo de códigos. Como se podrá advertir no todas las transcripciones se tienen al mismo tiempo, de hecho, se trabajó en cada una conforme se obtenía, lo cual permitió identificar patrones, dudas y conjeturas iniciales, mismas que se probaban en la siguiente entrevista acordada y trabajada. En este punto ya se tenían registrados códigos provisionales en las notas de campo.

De acuerdo con Kuckartz y McWhertor (2014), la primera aproximación a los datos es orientada por un proceso inductivo en el que el investigador inhibe ideas preconcebidas y da paso a sus propias apreciaciones sobre el proceso de análisis. Este primer acercamiento es conocido también como la fase de codificación abierta, la cual consiste en leer las entrevistas completas y responder preguntas asociadas con el objetivo de investigación.

Con esta perspectiva se realizaron varias lecturas a las entrevistas y notas de campo tratando dejar a un lado concepciones teóricas preconcebidas; se identificaron códigos que conforme se avanzaba en la revisión de nuevas entrevistas se confirmaban o desechaban.

En el Cuadro 5 se presenta una guía de cómo se emplearon los términos de código y categoría en el proceso de análisis de los datos obtenidos en las redacciones.

CUADRO 5: DEFINICIONES DE CÓDIGO Y CATEGORÍA

Tipo de elemento	**Definición**	**Ejemplo**
Código	Etiqueta utilizada para nombrar un hecho, un evento fenómeno u acción relevante para los propósitos de la investigación. Sirve para dar una denominación común a un conjunto de datos que comparten una misma idea. *Codificación abierta*: se crean y aplican códigos a los segmentos de las transcripciones a partir de sus relaciones y expresiones. *Códigos invivo*: son códigos que provienen de expresiones utilizadas por los participantes y que son tomadas por el investigador para conjuntar una idea o concepto. *Codificación axial*: búsqueda sistemática de propiedades y dimensiones de los códigos que dan pie a la creación de categorías y subcategorías.	-Falta de información -Base de datos sucia -Tiempo
Categoría	Se trata de una clasificación de códigos, hecha a partir de la comparación constante entre entrevistas, conceptos y su integración en un orden más elevado llamado categoría, también denominada en otros estudios como Familia de códigos.	Categoría: Obstáculos Agrupa a los códigos: -Falta de información -Base de datos sucia -Tiempo

Fuente: Soneira, A. (2006). La "teoría fundamentada en los datos" (grounded theory) de Glaser y Strauss. En Irene Vasilachis (ed). *Estrategias de investigación cualitativa*. Barcelona: Gedisa. pp.153-175.

Etapa II. Construcción de códigos y categorías

En palabras de Charmaz (2006) la elaboración de códigos es vista como un proceso complejo y como un vínculo esencial entre la recolección de datos y el desarrollo de teorías emergentes que explican esos datos. En esta etapa se define lo que está sucediendo con los datos y se comienza a trabajar con lo que significan. La construcción de códigos requiere de un trabajo de intuición, creatividad, coincidencia y suerte. La intención de la lectura se divide entre la identificación de códigos y su evaluación. En esta etapa hay de por medio un trabajo de lectura y relectura, que se focaliza por momentos en un aspecto específico para luego leerlo en el contexto global de lo recabado.

Se muestran los pasos que se siguieron para la construcción de los códigos. En los incisos D, E y F se retoman los pasos descritos por Kuckartz y McWhertor (2014):

A. Escucha activa de las grabaciones y lectura exploratoria de las transcripciones de las entrevistas realizadas junto con la anotación de códigos provisionales que a lo largo de la propia lectura se consolidaron o desecharon de acuerdo con su pertinencia, capacidad de síntesis y fidelidad con respecto al contenido descrito o etiquetado.
B. Integración de las transcripciones de las entrevistas en el programa MAXQDA. Se creó una carpeta por cada escenario. Como observaremos más adelante este programa asigna un número a cada párrafo de las entrevistas, lo cual permite identificar de forma sencilla la ubicación de una cita en particular.

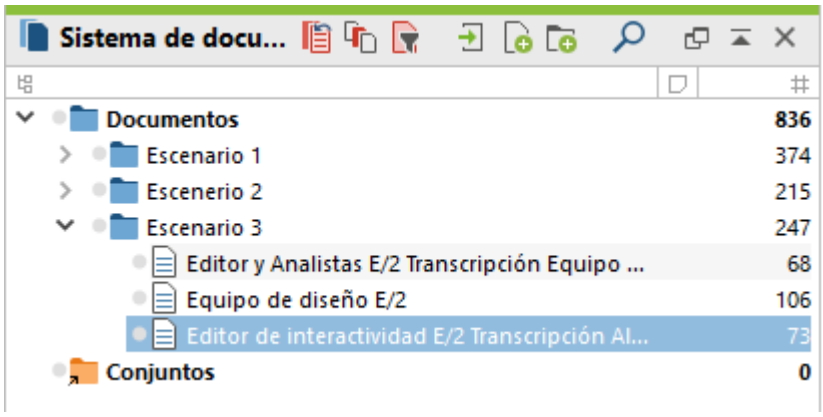


FIGURA 5: EJEMPLO DE GRUPO DE DOCUMENTOS EN MAXQDA

Nota: Ejemplo de cómo se agrupan los documentos por escenario en MAXQDA. El número de la derecha corresponde a la cantidad de fragmentos segmentados.

C. Codificación abierta a partir de la lectura exploratoria, que en principio es conducida por las preguntas de investigación: ¿Cuáles son los procesos que siguen los reporteros que se dedican al periodismo de datos, en relación con su comportamiento informativo, durante la elaboración de un producto noticioso de esa naturaleza? Así como: ¿qué patrones de comportamiento informativo siguen los profesionales que realizan periodismo de datos? y ¿cómo agregan valor a los datos los profesionales que desarrollan periodismo de datos?

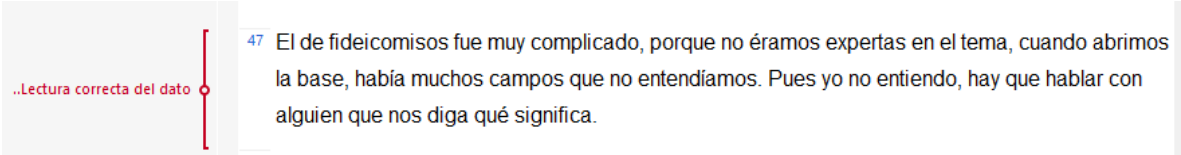


FIGURA 6: EJEMPLO DE CODIFICACIÓN EN MAXQDA

Nota: Ejemplo de codificación en MAXQDA, se trata de un proceso en el que el investigador realiza la codificación de forma manual. Conforme avanza en el análisis se va gestando una matriz de códigos que pueden utilizarse para asignarlos a fragmentos relacionados con lo que cada código expresa o contiene. En este caso nótese que el número asignado a este párrafo es el 47 de la entrevista a una reportera del escenario 1, de este modo si se desea citar este párrafo debe hacerse de la siguiente forma: Reportera E1: 47. En otras palabras, el sistema de citación de referencias para las citas textuales de entrevistas que se utiliza consiste en señalar el puesto o función, el escenario abreviado y el número del párrafo de la entrevista. Ejemplo:
(Reportera E1: 47)

D. Relectura de las transcripciones para verificar la existencia de contenidos que puedan ser descritos con códigos provenientes de lecturas previas de los mismos datos.

E. Identificación de códigos naturales (in vivo), es decir palabras o frases que son mencionadas por los propios participantes y que se reutilizan como códigos para referirse al fenómeno (Kuckartz y McWhertor, 2014), por ejemplo: “jugar con el dato”; “el dato no me da”.

F. Construcción progresiva de categorías provisionales que agrupen a los códigos de acuerdo con sus propiedades relacionales (codificación axial). Se utiliza el término categoría para referirse a una familia de códigos, o bien como lo sugiere Kuckartz y McWhertor (2014) como sinónimo de clase, grupo, familia o tipo. Es decir, de acuerdo con criterios específicos algunas características pueden ser examinadas y consideradas similares entre sí, por lo tanto, se engloban o conjuntan bajo una categoría asignada.

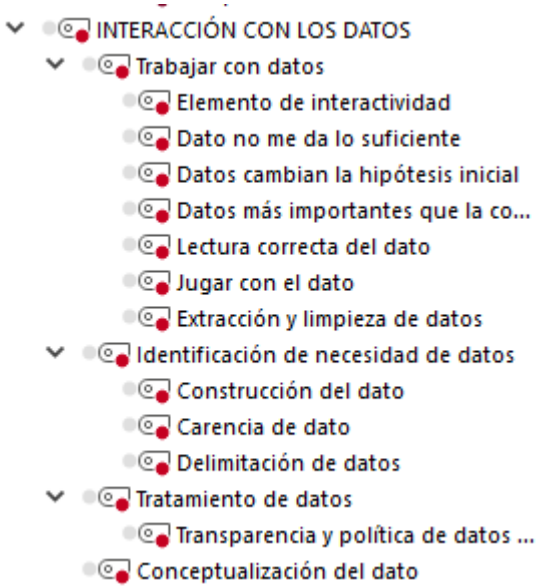


FIGURA 7: EJEMPLO DE AGRUPACIÓN DE CATEGORÍAS EN MAXQDA

Nota: Ejemplo de cómo se aprecia una categoría en MAXQDA. En este caso la categoría *Interacción con los datos* se divide en tres subcategorías: (1) Trabajar con datos; (2) Identificación de necesidad de datos y (3) Tratamiento de datos. De esta forma los códigos se van agrupando, por el investigador, de acuerdo con sus propiedades, relación y similitud.

G. Comparación constante de los códigos entre sí (Kuckartz y McWhertor, 2014) y búsqueda de alternativas para referirse a estos segmentos, los cuales, en algunos casos, se fusionaron en un solo término. Este proceso fue continuo y cubre toda la fase de codificación y análisis de los datos.
H. Conformación de una estructura preliminar de categorías y códigos que se valida con una nueva lectura de las transcripciones, en la que se asegura que los códigos y las categorías creadas en las últimas revisiones, o de forma intermedia, se apliquen a las transcripciones que se analizaron al principio. Esto implica regresar a las transcripciones para asegurar que los nuevos códigos se vean reflejados, si es necesario, en las transcripciones que se revisaron previamente. En otras palabras, se trata de un proceso de recodificación y recategorización.
I. Recodificación y reagrupación en categorías. Esto implica descartar, fusionar o reagrupar códigos para asegurar la consistencia de los datos. Este paso es dinámico e iterativo. Supuso un esfuerzo de análisis constante para el investigador frente a los datos. En esta etapa se considera también la participación de mis tutores cuyas aportaciones, observaciones y correcciones permitieron aclarar, afinar, ajustar y hacer más comprensible la estructura de códigos.

Etapa III. Codificación y segmentación del texto

Como se puede advertir, la construcción de códigos y categorías es un proceso paralelo a la codificación. Es decir, los códigos se aplican y construyen al mismo tiempo. En este caso el uso del programa MAXQDA fue útil para integrar, sistematizar y codificar los datos transcritos de las entrevistas realizadas. Como ya se comentó se concentraron las entrevistas en este programa y se realizaron lecturas sistemáticas de las mismas con el fin de identificar códigos que emergieran

de la propia lectura, y a partir de estos códigos construir categorías que agruparan subcategorías y códigos.

Es conveniente mencionar que en muchos casos un segmento concentra varios códigos empleados, de tal modo que hay códigos sobrepuestos en algunos segmentos de las entrevistas. Lo cual es comprensible dada la polisemia de los enunciados. Aquí un ejemplo tomado de MAXQDA en el que cada código empleado se diferencia por color:

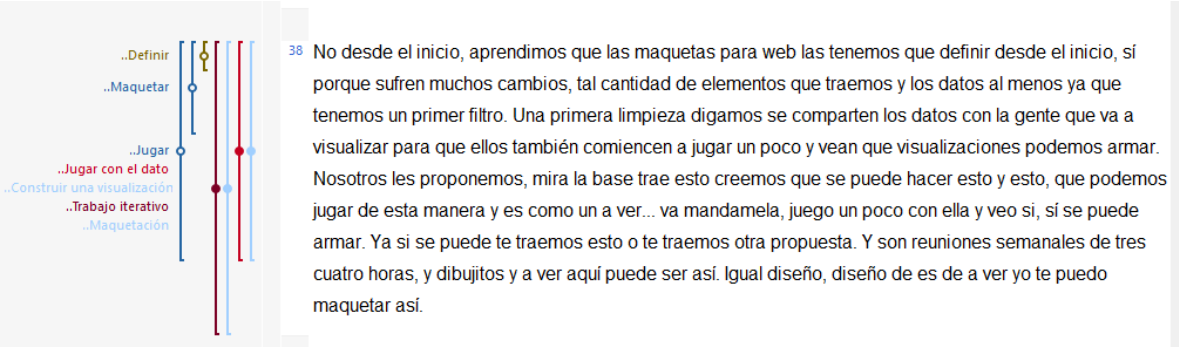


FIGURA 8: EJEMPLO DE SEGMENTOS CODIFICADOS

En total se crearon 185 códigos, agrupados a partir de su relación temática en 7 categorías centrales. Con estos códigos se etiquetaron en total 792 segmentos de las transcripciones y un informe del E1. En cuanto a las 7 categorías centrales son el resultado de menciones o alusiones frecuentes en las transcripciones o relaciones con otras categorías. Se cumplen los requisitos mencionados por Soneira (2006) en cuanto al hecho de que hay alusiones que apuntan a tal concepto; las relaciones establecidas entre subcategorías son lógicas y consistentes; los nombres empleados son abstractos de tal modo que den lugar a una teoría más general; conforme creció la estructura de categorías, subcategorías y códigos, la teoría también creció en profundidad y poder explicativo.

- PARTICIPANTE: integra referencias sobre la trayectoria personal de los participantes, sus motivaciones, alusiones sobre la forma en que se perciben, el papel de la intuición, sus prácticas en el consumo de fuentes de información, su función en la organización, sus tareas asociadas, así como sus valores y disposiciones.

- ESPACIO DE TRABAJO: conjunta los códigos que guardan relación con los requerimientos contextuales o institucionales, los valores del área en la que se desempeñan, la forma de organizarse, las tareas y obligaciones a los que están sujetos los participantes.
- PERIODISMO DE DATOS: agrupa los códigos que se asocian con las implicaciones de desarrollar este tipo de periodismo como los plazos, la percepción de otras áreas, los requerimientos internos. Integra también las menciones que hacen los participantes a otros medios que realizan este tipo de periodismo y que son vistos como referentes al momento de trabajar una nota. Conjunta, además, aquellos fragmentos que hablan de la conceptualización de lo que es o debería ser este tipo de periodismo.
- AUDIENCIA: códigos que refieren al destinatario o lector de los productos noticiosos que se elaboran, la concepción que tienen los participantes de los lectores y los ajustes que hacen a sus productos a partir de esta concepción.
- FORMATO DE PUBLICACIÓN: se refiere a los códigos asociados con los formatos de salida y las tareas que implican. En este caso se limita a los productos noticiosos diseñados para la versión impresa, electrónica y multiplataforma. Se codifican aquellas referencias a los procesos, limitaciones o tareas que supone cada plataforma.
- VALOR AGREGADO A LA INFORMACIÓN: integra los procesos de transformación de la información que realizan los participantes desde la óptica de la interacción con los datos, el desarrollo de la nota (su origen, su elaboración, los obstáculos que supone) los procesos para generar una visualización, así como las herramientas y tareas de programación asociadas.
- PROCESOS COGNITIVOS: agrupa los códigos que refieren a una acción cognitiva, en este caso la mayor parte de los códigos refieren a un verbo aplicable a una acción realizada por el participante. Los verbos se relacionan con conceptos, ideas, conocimientos y eventos que no son directamente observables.

Por último, se muestra la Figura 9 que refleja la vista general de códigos generados y empleados, del y para el análisis de la información recabada en los tres escenarios.

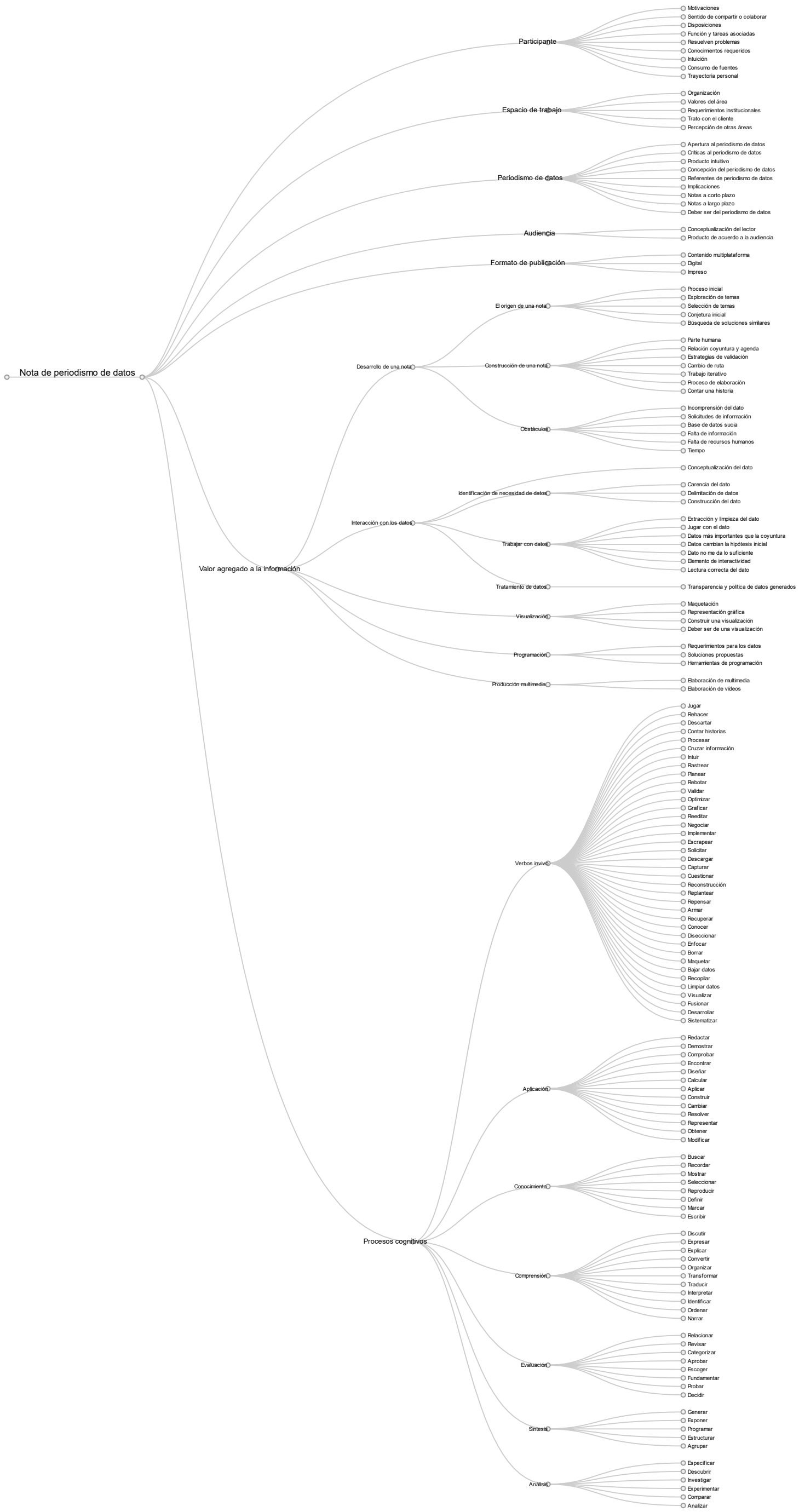


FIGURA 9: VISTA GENERAL DE CATEGORÍAS Y CÓDIGOS

Etapa IV Análisis

El planteamiento previsto para esta etapa es la utilización de un esquema basado en una matriz temática conocida también como *profile matrix* o *thematic matrix* (Ver, Kuckartz y McWhertor, 2014). Esta matriz temática, como lo advierten Kuckartz y McWhertor (2014), es el punto de partida de un proceso sistemático que permite elaborar resúmenes analíticos elaborados por el investigador y que se basan en extractos del material analizado. De esta forma el material de análisis se reduce a lo que es realmente importante para responder la pregunta de investigación de una forma documentada.

A continuación, se muestra una captura del trabajo realizado en MAXQDA en la que se puede observar en la parte izquierda la estructura de códigos establecida y en las columnas siguientes las entrevistas con segmentos codificados por cada código. Cada punto significa la presencia de segmentos codificados en las entrevistas. El segundo bloque corresponde a la vista de fragmentos codificados y el tercer bloque a las anotaciones del investigador en relación con este código y escenario.

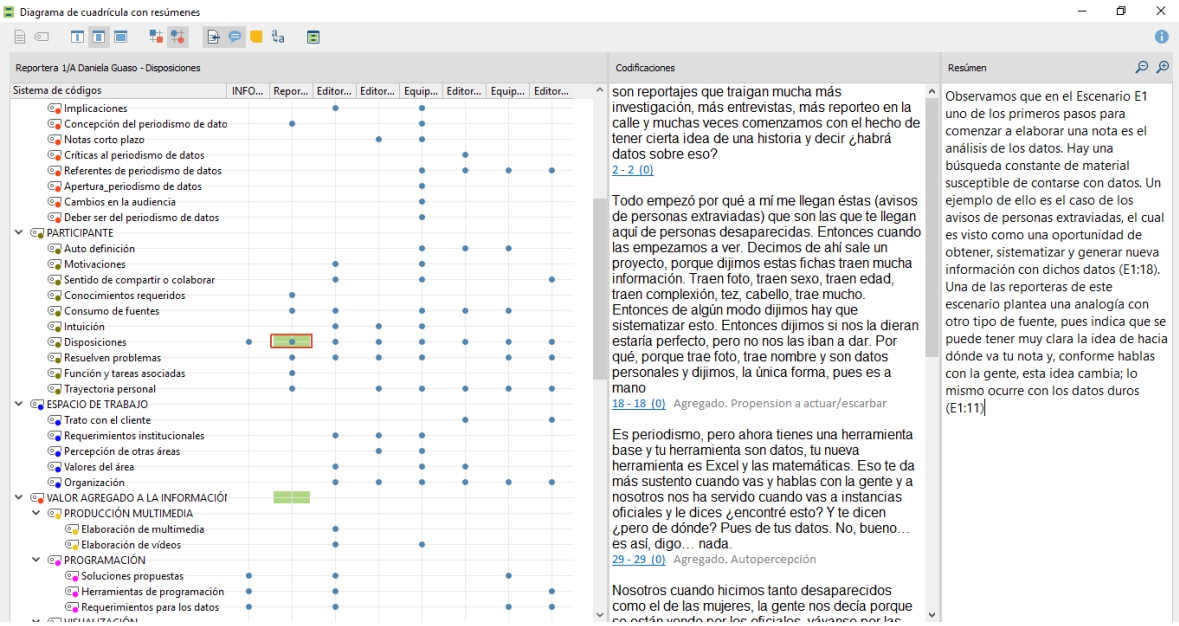


FIGURA 10: EJEMPLO DE TABLA DE RESUMEN EN MAXQDA

Originalmente esta forma de organizar la información consiste en incluir los tópicos principales en columnas y los participantes o escenarios en filas. Sin embargo, MAXQDA no permite esta vista por lo que se tienen que invertir los ejes pues en este caso los tópicos están en las filas y los participantes en columnas.

Hecha la aclaración se observa en el esquema original de Kuckartz y McWhertor (2014), (ver, Cuadro 6). que al realizar una lectura horizontal se analiza la perspectiva de un participante o de un escenario en particular en relación con cada uno de los tópicos (*case-oriented perspective*); o bien una lectura vertical que se centre en un tópico y se revise de una forma transversal de entre los escenarios o participantes (*topic-oriented perspective*). De ambas lecturas se obtienen resúmenes de casos o bien resúmenes temáticos, que pueden ser comparados o contrastados entre distintos escenarios, o bien, se pueden contrastar o relacionar las temáticas.

CUADRO 6: ESQUEMA BASADO EN EL PROPUESTO POR KUCKARTZ Y MCWHERTOR (2014)

	Código 1	Código 2	Código 3	
Escenario E1 (segmentos codificados)	Texto codificado del Escenario E1 sobre el Código 1	Texto codificado del Escenario E1 sobre el Código 2	Texto codificado del Escenario E1 sobre el Código 3	→ Resumen de caso E1
Escenario E2 (segmentos codificados)	Texto codificado del Escenario E2 sobre el Código 1	Texto codificado del Escenario E1 sobre el Código 2	Texto codificado del Escenario E1 sobre el Código 3	→ Resumen de caso E2
Escenario E3 (segmentos codificados)	Texto codificado del Escenario E3 sobre el Código 1	Texto codificado del Escenario E3 sobre el Código 2	Texto codificado del Escenario E3 sobre el Código 3	→ Resumen de caso E2
	↓ Resumen temático Código 1	↓ Resumen temático Código 2	↓ Resumen temático Código 3	

Para darle forma estructural a la descripción y análisis de los resultados se siguió, en un primer momento, la misma estructura de la categorización y codificación (ver, Figura 9). De tal modo que cada categoría daba lugar a un apartado en el que se describían los resultados de los tres escenarios sobre cada código. Este primer ejercicio, además de exhaustivo, resultó ser muy denso en términos descriptivos, pues se integraron todos los segmentos codificados por tema y esto provocaba la repetición de segmentos. Sin embargo, esta forma de trabajo permitió explorar a fondo cada uno de los elementos que dan forma a la estructura de categorías y códigos, incluso propicio la restructura y reordenamiento de elementos que lo conforman.

Las observaciones recibidas, a este primer planteamiento sirvieron para darle al estudio un segundo tratamiento en relación con la forma de narrar los resultados. De tal modo que se elaboró una línea narrativa para el reporte final de resultados que integra en las categorías y códigos trabajados una proporción de análisis y de relato, de un tono conceptual y descriptivo.

Este segundo tratamiento retoma el paradigma de codificación de Strauss y Corbin el cual representa un esquema que se puede utilizar para organizar las conexiones que van emergiendo (ver, Soneira, 2006). Los elementos del paradigma son los siguientes (la descripción corresponde al texto de Creswell (1998):

- Fenómeno: categoría central que da pie al estudio.
- Condiciones causales: categorías de condiciones que influencian el fenómeno
- Estrategias de acción intervención: acciones e interacciones que resultan del fenómeno central.
- Condiciones intervinientes: condiciones estrechas o amplias que influencian las estrategias.
- Consecuencias: resultados de las estrategias.

Aplicado a la estructura de este estudio queda de la siguiente forma:

CUADRO 7: ESTRUCTURA DEL ESTUDIO

Fenómeno	El comportamiento informativo durante la elaboración de un producto noticioso
Parte I Condiciones causales	El periodista de datos Espacio de trabajo
Parte II Estrategias de acción intervención	**Del dato a la nota:** Desarrollo de una nota: • Origen de una nota • Construcción de una nota **Trabajar con datos:** Interacción con los datos • Conceptualización de dato • Identificación de la necesidad de datos • Trabajar con datos • Tratamiento de datos
Parte III Condiciones intervinientes	**El tiempo y otros obstáculos** • Tiempo • Falta de recursos humanos • Falta de información • Solicitudes de información • Bases de datos sucias • Incomprensión del dato **Los elementos de un todo: visualizar, programar, editar:** • Visualización • Programación • Producción multimedia

	Una nota que entiendan todos: la idea del lector: • Conceptualización del lector • Producto de acuerdo a la audiencia **Formatos de salida:** • Impreso • Digital • Contenido multiplataforma
Parte IV Consecuencias	**El periodismo de datos visto desde su práctica** • Concepción del periodismo de datos • Referentes del periodismo de datos • Implicaciones • Notas a corto y largo plazo • El deber ser del periodismo de datos • Producto intuitivo • Apertura al periodismo de datos • Críticas al periodismo de datos

Cada apartado, a nivel de categoría, comienza con un balance general que incluye un análisis a la información recabada y le sigue un mapa con la estructura de categorías y códigos que da lugar a la temática del capítulo. La idea de ofrecer esta información al inicio de cada capítulo cumple con dos propósitos: (1) evidenciar la forma en que se codificó la información vertida y (2) funcionar como una guía para el lector en relación con los puntos de análisis.

Para la descripción de los códigos en cada capítulo se sigue una secuencia narrativa lógica que corresponde a los procesos de co-construcción propios del trabajo colaborativo, más que un orden en términos de frecuencia. Asimismo, en los casos en los que los códigos descritos figuran en dos o tres escenarios, con suficiente información, se hace un balance al final de cada apartado que lo analiza.

Por último, la Figura 11 muestra los procesos descritos en este capítulo, el cual está dividido en dos grandes bloques, (1) Acopio de información y (2) Análisis

de la información recopilada en campo. El primer bloque sintetiza los pasos que se siguieron para establecer contacto con las redacciones, diseñar y aplicar las entrevistas, la descripción de los participantes y la forma en que se vincula con la etapa de análisis (segundo bloque) donde se describe el proceso de construcción de la estructura de categorías y códigos que permitieron avanzar gradualmente en la revisión de los datos recabados y los procesos que siguieron para su análisis.

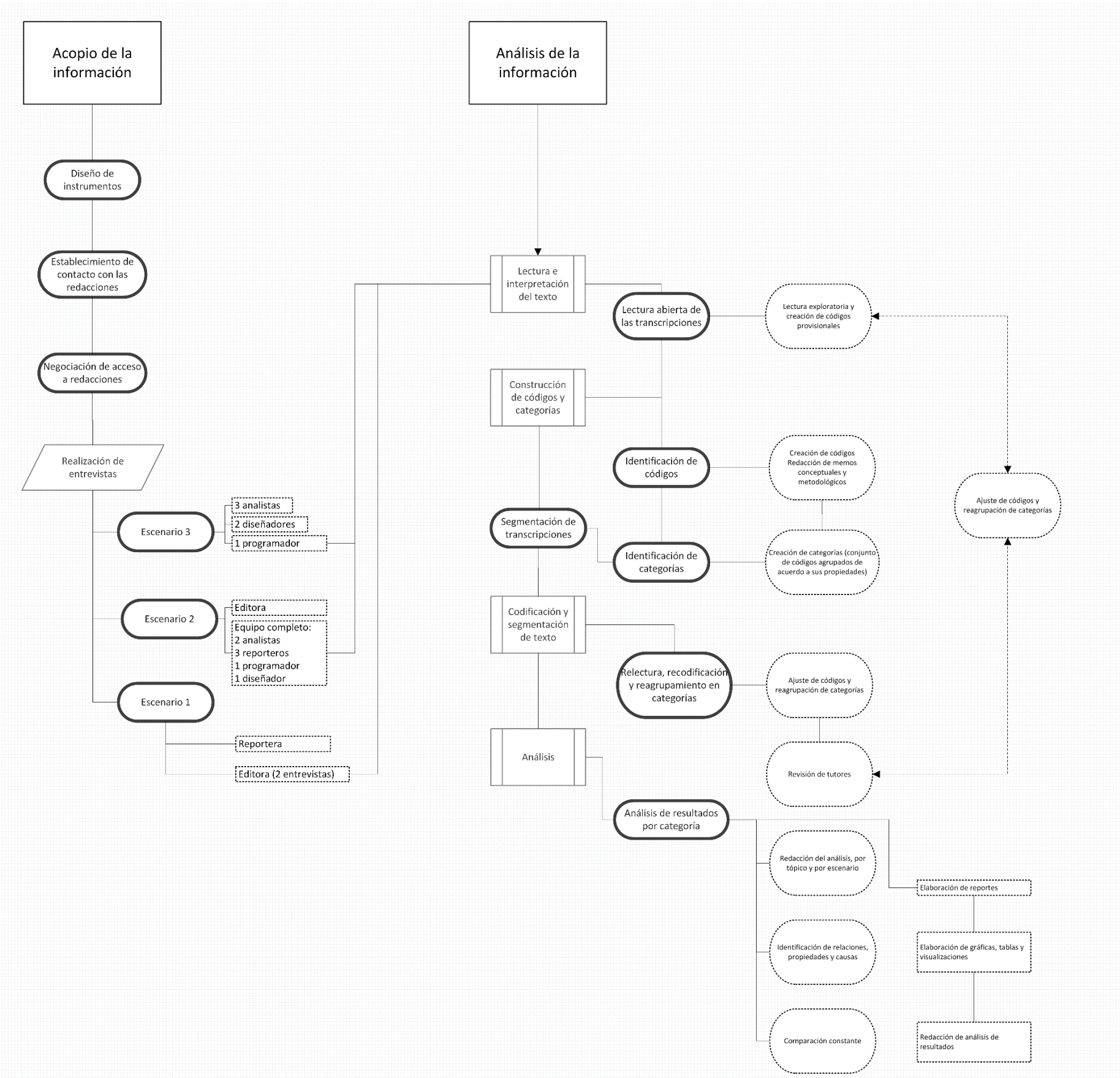


FIGURA 11: PROCESOS DE ACOPIO Y ANÁLISIS DE LA INFORMACIÓN RECOPILADA EN CAMPO

Visión general de categorías y códigos

El empleo de la metodología descrita permitió codificar 792 segmentos de las transcripciones y los documentos obtenidos de los 3 escenarios. Una vez que se han descrito las facilidades u obstáculos de cada escenario se entiende porque, a pesar de tratarse sólo de 2 participantes, el mayor número de segmentos se obtuvo del E1. Recordemos que en este caso la apertura de los participantes favoreció la recolección de información, pues además de las entrevistas hubo facilidades para aplicar el horizonte de información a la editora y reportera y nos facilitaron un informe de cómo se elaboró una nota en particular, este informe se incluyó como parte de la documentación analizada de este estudio. En este escenario (E1) se codificaron 352 segmentos; en el E2 se codificaron 206 segmentos (en este caso hubo una conversación sin registro de audio con la editora cuyas notas se retoman en el análisis de este escenario), y en el E3, se codificaron 234 (ver, Figura 12).

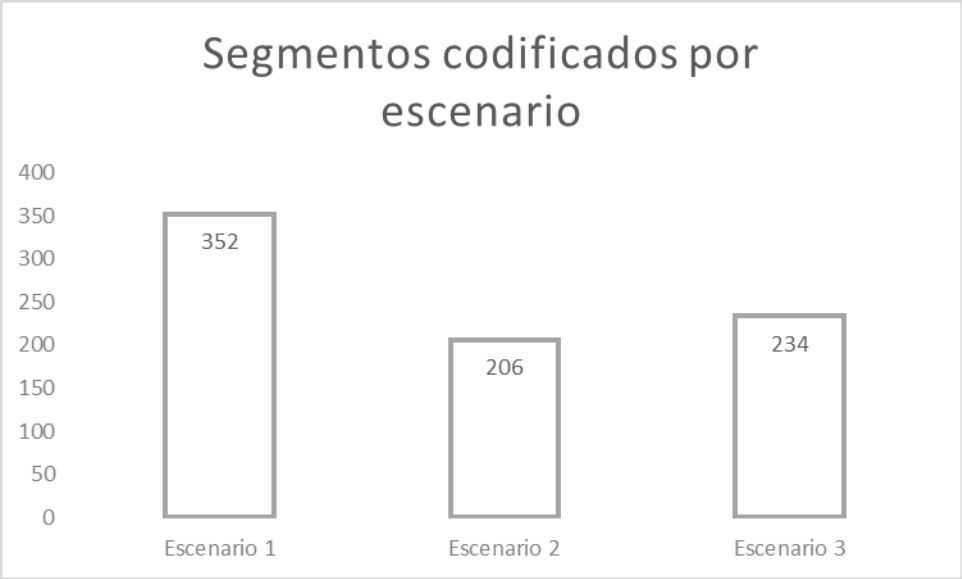


FIGURA 12: FRECUENCIA DE SEGMENTOS CODIFICADOS POR ESCENARIO

Los segmentos codificados se agrupan en las 7 categorías centrales (ya descritas), los cuales se pueden observar en la Figura 13. Las categorías con mayor ocurrencia son: *Procesos cognitivos* (309), seguido de las categorías: *Valor agregado a la*

información (265), *Participante* (108), *Espacio de trabajo* (44), *Periodismo de datos*, (32), *Audiencia* (23), *Formatos de publicación* (11). Recordemos que en el caso de la categoría Procesos cognitivos los códigos agrupados refieren a verbos de acciones realizadas por los participantes en el proceso de elaboración de una nota, en ese sentido se entiende la alta frecuencia de códigos concentrados en este rubro. En cuanto a la categoría *Valor agregado a la información* es una de las de mayor alcance en términos de subcategorías y códigos, de ahí su alto nivel de concentración numérica.

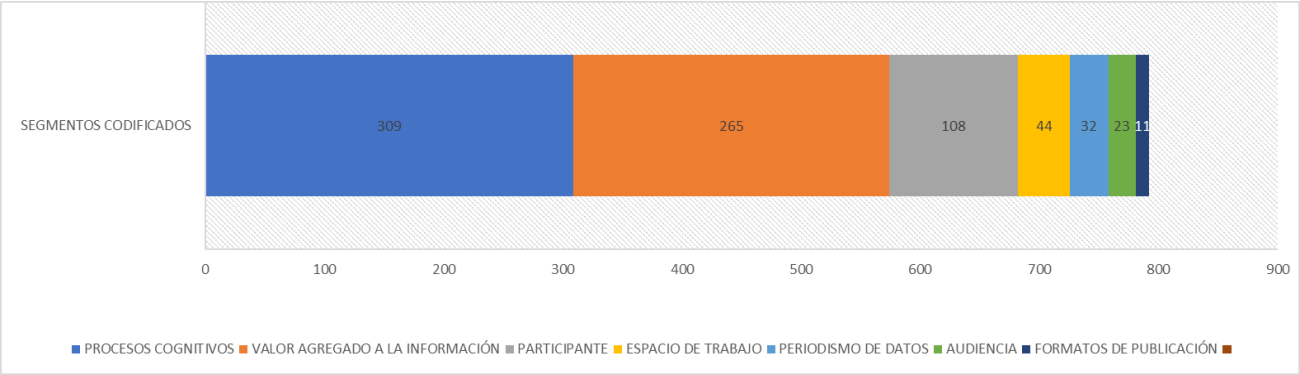


FIGURA 13: FRECUENCIA DE CÓDIGOS AGLOMERADOS POR CATEGORÍA

En la Figura 14 se observa que la categoría presente en todos los escenarios es la de *Procesos cognitivos* con una alta frecuencia en el caso de la entrevista al equipo del E2 y la reportera del E1. Se advierte, además, que el equipo del Escenario E2 y del E1 ofrecen el mayor número de segmentos asociados con los *Procesos de valor agregado a la información*. Por su parte, el E2 también integra el mayor número de segmentos asociados con las categorías *Participante* y *Periodismo de datos*.

El hecho de realizar una entrevista grupal con el Equipo del E2 hace justificable la concentración de temas en un solo bloque, pues el hecho de tener a todos los actores reunidos ofrece un mayor rango de perspectivas y opiniones sobre los procesos que se ponen en juego en la elaboración de una nota de periodismo de datos. Esto explica el despunte en este escenario frente a los escenarios E1 y E2 cuya distribución de códigos se dispersa en varias entrevistas.

A partir de estos datos podemos deducir que, de forma global, se cuenta con un mayor número de segmentos codificados sobre sobre los procesos cognitivos, los procesos de transformación en los que interviene el participante, así como, las motivaciones y disposiciones que pone en práctica durante la elaboración de un producto de periodismo de datos y su contexto.

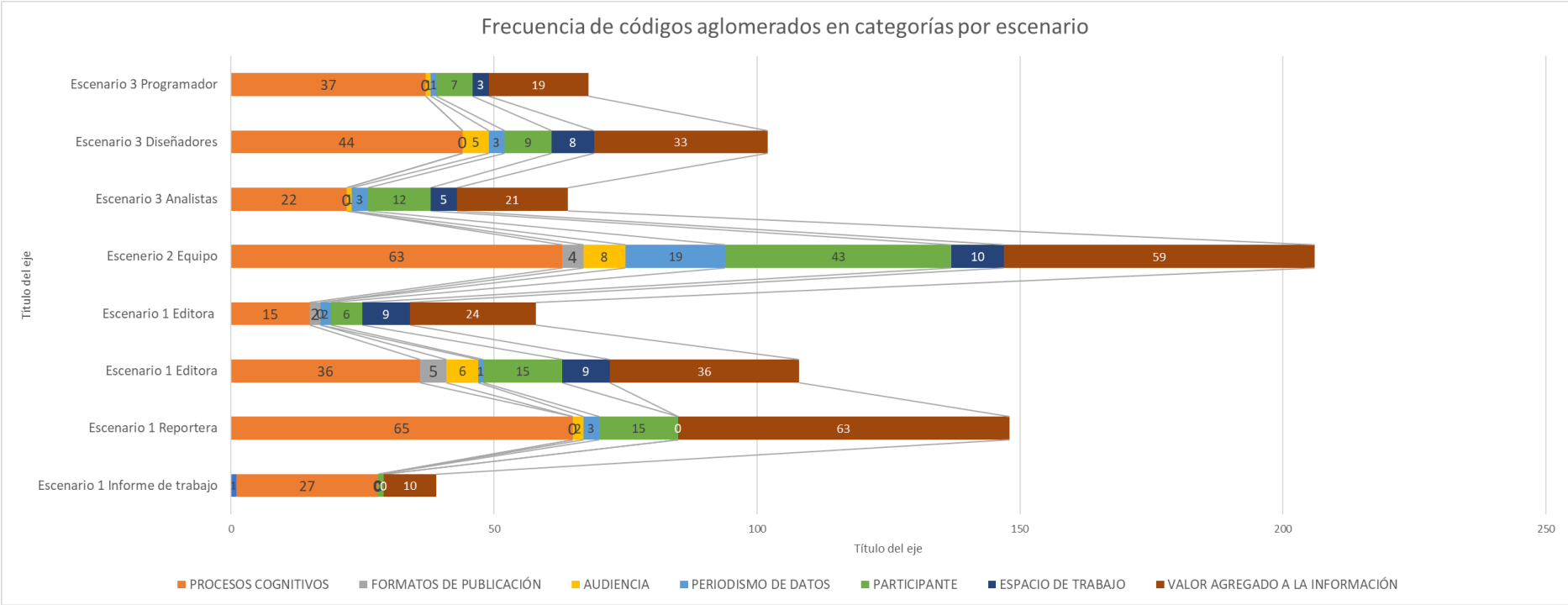


FIGURA 14: FRECUENCIA DE CÓDIGOS AGLOMERADOS EN CATEGORÍAS POR ESCENARIO

RESULTADOS

PARTE I CONDICIONES CAUSALES

La brújula interna del periodista de datos

En este capítulo se analizan los resultados obtenidos en relación con los participantes. En este caso se observan menciones a su formación profesional, sus áreas de interés, motivaciones y disposiciones personales para realizar el periodismo de datos. Se considera que estos datos dan muestra de cómo el sujeto se reconoce y visualiza en el entorno social.

Balance general

Las principales motivaciones de los participantes se ciñen a evidenciar información de difícil acceso, ubicar anomalías o posibles delitos y hacer comprensible una investigación para el público en general. Se apela a la función social del periodismo y a su capacidad de empoderar al ciudadano brindándole información para su toma de decisiones o para actuar. Los participantes operan bajo la lógica de la colaboración y la apertura de datos, lo cual funciona también como una motivación que influye en su actuar.

Hay casos en los que se habla de organizar talleres e intercambiar saberes específicos. La práctica del periodismo de datos requiere del conocimiento de dominios de saber, como el diseño, la programación y la visualización. Contar con

nociones básicas de estas áreas facilita la colaboración y el trabajo interdisciplinario orientado a un producto.

Hay una clara disposición a informarse para informar, a buscar ideas para el desarrollo de un tema, a consultar otros medios como *The New York Times*, *The Guardian*, *Washington Post, The Economist, Animal Político, Angeles Times, The Upshot* y otros diarios como *La Nación* de Argentina y Costa Rica; todos ellos son vistos como referentes de lo que puede hacerse en este campo o para verificar los temas que están dejando de atenderse. También se consultan para buscar oportunidades de actualización o para conocer experiencias de otras unidades de periodismo de datos. También se llega a estos medios a través de las redes sociales, usando etiquetas muy específicas para dar con ellos.

Los participantes marcan diferencias con el diarismo tradicional. Se habla de trabajar notas a largo plazo, que integran distintos saberes, no se limitan a publicar declaraciones, profundizan en los datos para ubicar hallazgos por medio de un cuestionamiento continuo de los mismos. Entre sus disposiciones más marcadas está la de crear productos fáciles de comprender, que sean amigables y nada complejos. Asumen el papel del lector para establecer empatía sobre lo que debería ser la nota en relación con la comprensión de ésta.

En una conversación con la Editora del E2, ella hablaba del sentido social que debería tener quien aspire a desempeñarse en esta área. Esto es claro en la disposición de colaborar, compartir experiencias y participar en iniciativas de datos abiertos. Ella utilizó el término “activistas sociales”, término que en este contexto es afortunado por el tipo de trabajo, motivaciones e incentivos que realizan.

Por otra parte, hay una aptitud por resolver dificultades para generar una nota, ya sea limitaciones de acceso a la información, falta de recursos humanos o técnicos, datos sucios o inconsistentes.

Estructura de códigos

La categoría *Participante* se caracteriza por reunir información sobre el plano personal en términos de intereses, motivaciones y forma de resolver necesidades

de información. Asimismo, nos ayuda a determinar cómo se conducen los participantes a partir de su contexto y universo de creencias. Integra los siguientes códigos relacionados con el ámbito personal: Disposiciones, Trayectoria, Consumo de fuentes, Motivaciones, Resolución de problemas, Sentido de compartir o colaborar, Intuición, Función y tareas asociadas, Conocimientos requeridos y Autodefinición en la Figura 15 elaborado con MAXQDA se presenta la distribución de los segmentos relacionados con esta categoría y sus códigos.

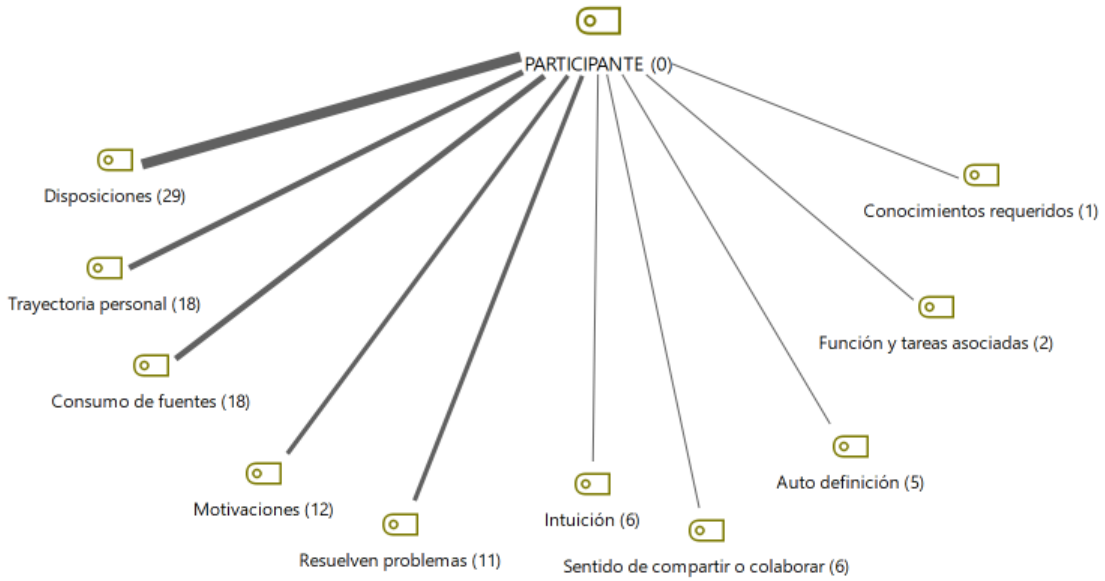


FIGURA 15: PARTICIPANTE, ESTRUCTURA Y FRECUENCIA DE CÓDIGOS

Como se puede observar en la Figura 15 el mayor número de segmentos codificados se refiere a Disposiciones (29). Le siguen los códigos de Trayectoria personal y Consumo de fuentes (18), y en un tercer bloque las Motivaciones (12) y Resolver problemas (11).

Autodefinición

Aunque no es una constante en los diferentes escenarios, la palabra que parece concentrar el concepto básico del periodista de datos es el de “datero”, la cual incluye su percepción como persona en constante formación y su dominio de un saber específico.

> “Esos trabajos que no tienen publicidad y que son altamente especializados, si no le dan idea, al menos, ni aspiran de una idea de lo que se puede hacer con los datos. Al menos así lo veo yo, como un datero en ese proceso.” (Reportero 1 E2:101)

Función y tareas asociadas

La Reportera del E1 explica de forma detallada las tareas asociadas a su puesto. Su papel es el de periodista, y parte de sus tareas es desarrollar reportajes especiales y publicarlos cada quince días. Se espera que este tipo de productos cuenten con mucho trabajo de investigación.

> “Mi función, pues yo soy periodista, no soy analista ni economista. Mi formación tal cual es en periodismo. Dentro de la sección me toca hacer los reportajes especiales, que son de al menos uno cada quince días. Son reportajes que traigan mucha más investigación, más entrevistas, más reporteo en la calle.” (Reportera E1:2)

Ella menciona una serie de obstáculos relacionados con el tiempo considerado para el procesamiento de datos y la integración de entrevistas o más información. Este aspecto se abordará más adelante.

Conocimientos requeridos

El trabajo que se realiza en el equipo del E1 permite puntualizar el tipo de conocimientos que deben conjugarse para generar un producto de periodismo de datos. En este caso se habla de diseño, programación y visualización.

> “Pero todo aquel en el que sacamos como un micrositio especial son muy complejos porque no sólo somos nosotras ahí, sí tenemos un diseñador, programadores, visualizadores, es trabajar ya con un equipo. Y eso es mucho más complejo, porque es definir maquetas. O sea, a lo que voy es a esa parte; periodismo de datos no es que sólo escribas. Además de eso debemos de pensar en cómo lo vamos a visualizar, definir una maqueta para web, definir cómo nos vamos a mover, definir si nos vamos con hojas, con gráficos… Si usamos *scroll,* no usamos *scroll*… Todo eso lo definimos con el equipo que se conforma para ese proyecto.” (Reportera E1:36)

También permite redundar en la necesidad de contar con equipos transdisciplinarios y en propiciar que estas áreas con saberes específicos puedan establecer una colaboración y comunicación constante.

Trayectoria personal

Se observa que los periodistas de datos cuentan con estudios universitarios que llegan hasta la maestría o la especialización. Existe un perfil transdisciplinar que incluye estudios en periodismo, ciencias de la comunicación, diseño gráfico, ciencias políticas, relaciones internacionales e ingenierías en sistemas y mecatrónica. No obstante, la disciplina que más destaca es la de economía. Muchos de ellos han colaborado en proyectos de periodismo de datos durante varios años y expresan que forman parte de las primeras generaciones de profesionistas formados en su campo. En el E1 tanto la editora como la reportera tienen formación en periodismo.

> "De la maestría de periodismo del CIDE. Estuve ahí, igual que [Y]. Nada más que [Y]. estudió creo que… es de la siguiente generación. Yo estuve una antes. A mí no me tocó el profesor Jiménez y a ella sí. No fuimos los conejillos de indias porque esos fueron los de la primera generación, yo fui de la segunda y [Y] de la tercera, creo." (Editora, E1:5)

> "Soy economista, aunque ya un par de años me he estado dedicando a hacer tareas de reportero. Estoy dejando el análisis de datos, pero ya lo combino mucho con lo de reportería." (Analista 3, E2:27)

El E2 también cuenta con dos analistas formado en economía (Analista 1 y 3 E2), un diseñador gráfico (Diseñador E2), un analista que es ingeniero en sistemas (Analista 2, E2), una persona que apoyan en la búsqueda de información en bases de datos (Becaria E2) y dos reporteros formados en periodismo (Reportero 1 E2 y Reportera 2, E2)

En el Escenario 3 los dos analistas y el becario fueron alumnos del editor general de este colectivo en el ITAM (Analista 1 y 2; Becario, E3); él fue quien los invitó a colaborar en proyectos vinculados con el periodismo de datos. La formación que tienen es de economía, ciencia política y relaciones internacionales. El

Programador (E3) tiene una formación en mecatrónica y programación; conoció al editor general por proyectos en común y decidieron formar el colectivo. Por lo que se refiere a los diseñadores (Diseñadores 1 y 2, E3) ellos tienen esa formación y llegaron a este escenario, pues se estaba buscando a personas con su perfil.

Disposiciones

La idea de una brújula interna que conduce el comportamiento o determinadas formas de actuar, de pensar o decidir por parte los participantes es un elemento implícito que opera como una guía para conducir acciones y comportamientos en un contexto dado. Informarse y tomar una posición inquisitiva frente a la información que se consume es una práctica recurrente en la búsqueda de temas por desarrollar. Esto supone que el participante permanezca en un estado de alerta constante para localizar notas (Analista 1 E3:20).

Esta forma de radar opera como una propensión a encontrar en la información posibilidades para el desarrollo de una nota. La información o los datos son vistos por los participantes como probables detonantes o insumos de un tema.

> "Todo empezó porque a mí me llegan éstas (avisos de personas extraviadas) que son las que te llegan aquí de personas desaparecidas. Entonces cuando las empezamos a ver. Decimos de ahí sale un proyecto, porque dijimos estas fichas traen mucha información. Traen foto, traen sexo, traen edad, traen complexión, tez, cabello, trae mucho" (Reportera E1:18)

Se asume que todos leen y que el consumo de nueva información va formando ideas en los participantes e incluso se formulan hipótesis que les va "jalando ruido". (Analista 2 E2: 92). Hay una propensión a ubicar hallazgos a evidenciar irregularidades y si la información que se requiere es difícil de ubicar, se busca la forma de "salir adelante" (Editora E1:25). El periodismo de datos es visto también como un campo provisto de herramientas que permiten enfrentar un mundo tecnológico en constante cambio.

> "Realmente creo que el periodismo de datos te otorga herramientas para enfrentarte a una realidad tecnológica cada vez más compleja, no sólo como

> profesional, sino como persona, y esto es algo distinto al diarismo." (Analista 1 E2: 175)

En este ejemplo el participante hace una clara distinción entre su trabajo como periodista de datos y lo que es el diarismo: hacer periodismo con base en declaraciones, boletines o notas a corto plazo. Como se verá más adelante esta distinción es recurrente y va conformando un grupo de nociones que dan forma a la idea de lo que es el periodismo de datos y como es visto por otro tipo de colegas. Por otra parte, se asume que realizar este tipo de periodismo es cumplir una función social al empoderar o ayudar a los lectores a que actúen informadamente (Analista 2 E2:180).

Otra disposición clara es la creación de productos que sean fáciles de comprender por el público, que sean amigables, que no requieran de estudios especializados para saber de qué se trata, es una inclinación constante en este campo. (Reportero 1 E2: 59 y Analista 3 E2:153). En el mismo tono, encontramos casos en los que el participante asume la postura del lector para comprender las expectativas que tienen sobre un producto en específico.

> "Entonces yo les dije, no, no puedo la gente no va a entender eso. No tienen por qué picarle 600 veces para enterarse de los fideicomisos que hay." (Editora E1:51)

En los tres escenarios hay ejemplos claros de esquemas incorporados que operan como filtros que cuestionan la información que se consume, de buscar ángulos o posibilidades de trabajar una nota con datos y ofrecer al lector información significativa y con valor social que propicie su empoderamiento.

Resolución de problemas

Hacer este tipo de periodismo plantea problemas de diversa índole, desde la búsqueda de financiamientos, falta de personal especializado o datos difíciles de obtener o precisar. Sin embargo, existe una vocación por resolver estas dificultades para generar la nota. Sobre este punto, en el E1 se observan dificultades al compartir los recursos humanos con otras áreas del mismo diario, situación que los

limita en cuanto al alcance de las notas que se plantean y los obliga a gestionar apoyos externos para subvencionar el pago a personal especializado que apoye en el desarrollo de la nota.

> "Ahora muchos de estos interactivos, la mayoría, los hemos estado sacando adelante con gente extraña, porque justo la gente del diario, por decir, el área de infografía se dedica a hacer infografías para todas las secciones y son 5 personas nada más. Entonces, todos los días sacar infografías para todas las acciones y aparte para nosotras, está cañón." (Editora E1:16)

Otro tipo de problemas planteados es que no hay acceso a la información, principalmente gubernamental. Esta limitación opera también como un incentivo en la voluntad del participante, pues habla de "empecinarse" y de que la nota "saliera adelante" (Editora E1:25). En este ejemplo, la búsqueda tenaz por desarrollar la nota los llevó también a beneficiarse de un financiamiento.

Resolver una dificultad es visto como un incentivo para quien desarrolla este tipo de periodismo. La resolución es una recompensa materializada en el logro de una nota. Esto se observa también en el E2, donde uno de los analistas habla de la dificultad de trabajar con archivos en PDF y la sensación que experimentan al momento de convertir los datos y hacerlos manejables.

> "Pero las cosas que aprendí sobre la marcha, la existencia de algunos programas que te pueden ayudar; pero muchas veces ese PDF era como fascinante convertirlo." (Analista 3 E2:112)

De manera similar, el E2 explica que los problemas relacionados con la sistematización de la información son comunes porque, al final, se trata de periodismo de investigación; es decir, se asume que es lo normal y que esto supone sortear dificultades para lograr lo que se propone y contar una historia con datos.

> "A final de cuentas es como el periodismo de investigación, o sea, alguien puede agarrar y meterse a solicitudes y tener bonches de papeles de proyectos, a cuerdos y, a final de cuentas, tienes que ver cómo procesar esos datos para contar una historia que se entienda. Ésta es la nota, esto es lo que quiero decir; lo mismo con los datos." (Analista 3 E2:158)

Por último, tenemos las manifestaciones del personal de diseño y programación, en el E3 quienes se suman a la idea de que su función es ofrecer soluciones y hacer

factible los planteamientos o requerimientos que les hagan, en tanto haya tiempo (Programador E3:30), (Diseñador 1 E3:86).

La falta de personal es un problema recurrente para el E1 y E2, esta situación orilla a los participantes a ubicar oportunidades de apoyo externas. Los problemas de carácter técnico y la falta de información son visibles en los tres escenarios y son vistos como retos por los participantes. La recompensa o incentivo que los motiva es la propia materialización de la nota en que trabajan. El empecinarse, e involucrarse de forma profunda en el desarrollo de la nota tiene la misma motivación.

Motivaciones

Evidenciar información de difícil acceso a la ciudadanía por medio de un producto profesional y claro. Este es entre líneas el principal elemento que, de acuerdo con los participantes, motiva a la gente que se dedica a desarrollar periodismo de datos. La búsqueda de información que evidencie una anomalía o un posible delito es una motivación para la Editora del E1 a tal grado que habla de empecinarse en el tema y buscar la forma de que el tema saliera adelante (Editora E1:25). Otras motivaciones ubicadas se ciñen a acaparar la atención de los lectores por medio de una mayor resonancia de su nota que sirva para difundir una problemática

> "Porque no es fácil entender el tema, entonces no es fácil entender el tema te enfrentas a que puede ser una investigación buenísima, pero si no lo entiende la señora que vende los tamales en la esquina ya te fregaste porque no se va a socializar y no se va a lograr difundir la problemática, ni le va a llegar a la gente y creo que al menos en una parte muy pequeñita creo que sí conseguimos que la gente por lo menos volteara a ver el tema que eso es algo muy difícil no." (Editora E1:45)

Darles claridad a temas "poco entendibles" es un objetivo que se persigue durante la elaboración de una nota, el proceso de transformación que va del acopio y análisis de los datos al desarrollo de la nota tiene esta motivación (Analista 3, E2:155) que incluso propicia la percepción de mayor acercamiento entre el periodista y el lector.

> “Uno tanto a descubrir un montón de historias que están ocultas en los datos, que pueden tener un impacto tremendo que no, pero también te da una medida de transparencia, y una medida de cercanía entre el lector y la persona […]También genera una confianza mayor en el lector sobre que se hizo y como se hizo, porque se hizo.” (Analista 1 E2:167)

Este tipo de trabajo que saca a la luz pública “revelaciones” es visto por uno de los participantes como una forma de enriquecerse de forma profesional. Es decir, hay un incentivo a nivel personal (Analista 1 E2: 174). El ejercicio de este tipo de periodismo es motivado también por lo que se podría denominar dimensión social del periodismo, en la cual la información es vista como una herramienta para empoderar al ciudadano y brindarle elementos para que se informe y actúe.

> “Es empoderar a la gente o ayudar a la gente y el periodismo tiene una función social también interesante esa combinación que tiene, por eso digo la parte exploratoria para mí es como muy interesante, la gente también es una herramienta para ellos, no es un dato digerido, es una ramita para que ellos mismos generen sus historias, se informen y de cierta manera actuar.” (Analista 2 E2: 180)

Una nota de periodismo de datos cumple su cometido si es comprensible a un mayor rango de lectores. Si la nota es motivo de un actuar especifico por parte de los lectores, se cumple en buena medida una de las motivaciones de las personas que colaboran en el desarrollo de periodismo de datos, es decir el empoderamiento del lector por medio de la información. Finalmente, es claro que en los dos escenarios E1 y E2 las motivaciones tienen un carácter social.

Sentido de compartir y colaborar

En la comunidad de periodistas de datos permea una cultura de colaboración, donde expresiones como “datos abiertos” o “compartir datos y experiencias” demuestran un sentido, una forma de trabajar en este campo. La organización de talleres, el intercambio de idea o experiencias son comunes y conforman una propensión a la apertura e intercambio de información e incluso formas de resolver problemas

comunes. “Simplemente son talleres gratuitos que damos aquí nada más para enseñar un poquito las tecnologías que utilizamos” (Programador E3:34).

Algunos de los mecanismos utilizados para mantener contacto con las listas de correos en las que se comparten trabajos para recibir retroalimentación o recomendaciones.

> “Y te hacen recomendaciones y te dicen esto pudo mejorarse así. Al mismo tiempo, ellos nos dan a conocer sus propios proyectos o trabajos que van sacando”. (Editora E1:56)

Otra práctica incipiente es el *crowdsourcing* (colaboración abierta distribuida) que consiste en integrar los esfuerzos de muchos voluntarios para capturar, analizar y transparentar información para hacerla accesible, con posibilidades de generar productos a partir de su sistematización colectiva (Analista 2 E2:121). Esta práctica es considerada exitosa, pues permite el intercambio o la sistematización de los datos a partir de la colaboración de varias entidades (Reportera 2 E2:101).

El sentido de colaborar es visible en los tres escenarios. La colaboración se sostiene por los valores sociales que se comparten y por los beneficios recíprocos que se obtienen en forma de una red informal de especialistas que enfrentan problemas similares, mecanismos de retroalimentación y apoyo en tareas que requieren del trabajo de muchos voluntarios especializados.

Intuición

La intuición es un elemento que figura en distintas etapas de este proceso. Al inicio, al momento de “estar viendo datos y tienes la impresión de que puede haber algo” (Analista 2 E2: 48). Al momento en el que “alguien llega con la intuición de un tema y se buscan los datos” (ídem) o cuando recién acaban de publicar datos sobre un tema y se intuye que se puede encontrar información de forma abundante (Editora E1:19). La intuición opera en la selección de un tema, en la revisión de los datos e incluso en el momento en el que se sabe que determinados datos están disponibles.

La intuición es vista por uno de los participantes como algo impreciso ajeno a una hipótesis. Puede ser vista como una intención de encontrar una idea que subyace en la información a analizar y que determina el rumbo de una nota.

> "Más que una hipótesis es una intuición porque no son tan puntuales, porque una hipótesis tratas de comprobarla, una cuestión en particular, o sea la compruebas o la descartas, en este caso lo que buscamos es tener una idea de más o menos que podría estar ocurriendo y ya dejamos que la información nos indique por dónde va la cosa, puede tomar un camino." (Analista 3 E2: 84)

La idea de comprobar o desechar "que hay algo ahí" (Equipo E2: 52) funciona también como una motivación, que en algunos casos también te puede llevar a un rumbo equivocado.

> "De hecho hay uno que estamos trabajando ahora, que teníamos una intuición antes de conseguir los datos, o sea queríamos demostrar la intuición que teníamos o tenemos y resulta que nos está dando todo lo contrario, o sea la intuición general está completamente equivocada, entonces allí nuestro análisis nos está dando justamente lo contario de lo que habíamos dicho." (Analista 3 E2: 87)

Tanto en el E1 como en el E2 hay ejemplos claros en el que la intuición está ligada con la motivación de encontrar algo irregular de entre un cúmulo de datos. Esa intuición responde a un interés de evidenciar información sensible o significativa para la ciudadanía. En algún momento la Editora del E1 reconoce la importancia de este elemento en el desarrollo de una nota "La intuición es muy importante porque siempre dicen piensa mal y acertarás" (Editora E1: 31) En su caso ella narra el caso de un fideicomiso que después de 16 años de no contar con recursos el gobierno, deciden "meterle miles de millones de pesos" (Editora E1: 31).

Consumo de fuentes de información

La consulta de fuentes de información es una tarea común en este campo. Los objetivos de informarse o consultar fuentes de información está ligado a los objetivos de buscar ideas para el desarrollo de un tema, para contextualizar o destrabar una idea sobre un tema. La revisión de fuentes de información puede ser visto como un detonante para la búsqueda de más información respecto a un tema, para

plantearse preguntas que funcionen como hipótesis o bien como un punto de partida para el desarrollo de una posible nota.

Se consultan fuentes también para revisar la forma en que otros medios resuelven determinados obstáculos. En el E3, por ejemplo, se habla de consultar diarios internacionales como *The Guardian*, para saber cómo resolvieron la visualización de datos sobre un tema complejo (Diseñador 1 E3:84, 89). En el E2 se habla de que la consulta de otros medios es un "punto de partida" permite saber que se está dejando de atender o bien para "traducir" un trabajo que ellos realizaron con la información que el participante tiene (Diseñador 2 E2: 95)

Por otro lado, la agenda opera como un listado de requerimientos de temas que son vistos como posibilidades potenciales para trabajar y se sopesa junto con otros temas durante el proceso de búsqueda. Se ubican temas potenciales de largo aliento, mientras se trabaja en notas para cubrir la cuota diaria tal como lo hace ver la editora del E1 "Son cosas que voy guardando y al mismo tiempo hay que trabajar en lo del día a día." (Editora E1:56)

Hay fuentes que de forma recurrente son consultadas en estos contextos, diarios como el *New York Times* y *The Guardian, Washington Post, The Economist, Animal político, Angeles Times, Herald Tribune* y otros como *La Nación* de Argentina y *La Nación* Costa Rica, se consultan con distintos propósitos por los participantes, principalmente son vistas como un referente para el desarrollo de notas de periodismo de datos (Editora E1:8; Analista 1 E2: 100; Diseñador 1 E3:101).

Por otra parte, hubo un grupo de participantes que, si bien consumen notas de los diarios, llegan a ellas por medio de redes sociales, incluso hablan de seguir determinadas etiquetas que les permite filtrar información.

> "Llegas a notas por Facebook, Twitter. Me gusta también leer *LondBrits, The Guardian*, y también por el trabajo que hacemos vemos muchos medios que tienen que ver con daily journals, nos metemos, bueno yo, y creo que también ustedes, a *The Upshot, The New York Times*, cuando alguien publica algún interactivo [,,,] También nos ayuda a nosotros para cuando tenemos que bocetar alguna nota o presentar datos de alguna forma por eso nos ayuda estar expuestos todo el tiempo." (Analistas 2 E3:26)

Esto lo ratifica también la Editora de este Escenario (E1) y añade también las listas de correo como una fuente de información para mantenerse actualizado y pendiente de los entrenamientos que ofrecen las agencias de noticias, o los grupos de trabajo en redes sociales.

> "Tenemos por ejemplo el Twitter nos ayuda mucho porque en este tablero del Twitter tanto [C] como yo estamos suscritas, tenemos varias como, etiquetas que seguimos. Entonces por ejemplo yo sigo la de #Periodismodedatos, la de #Dataviz, la de #Opendata, entonces tú puedes aquí darle seguimiento a gente que nos interesa conocer o nos interesa el trabajo que ellos están desarrollando […] Por ejemplo, estamos suscritos a varios correos, listas de correos, de periodistas que están en todo el mundo, por ejemplo, yo estoy suscrita a una de periodismo de datos en español que lo crearon los periodistas de IRI que es International Reporter Investigative y ellos cada año hacen una reunión en Estados Unidos para precisamente hacer entrenamientos desde cómo usar Excel, bases de datos, etcétera." (Editora E1:56)

Los espacios de trabajo

Esta capitulo reúne los elementos que guardan relación con los requerimientos, los valores del área en la que se desempeñan los participantes, la forma de organizarse y las tareas y obligaciones a las que están sujetos. Se habla sobre los elementos situacionales que enmarcan el trabajo de los participantes en términos del área en el que trabajan, objetivos del área, percepción de otras áreas, la visión y valores institucionales y el trato con el cliente.

Como se puede apreciar las condiciones para crear una nota de periodismo de datos son distintas en cada caso, desde la disposición física del espacio y lo que dice éste en términos de flujo y jerarquía, hasta los valores implícitos de cada organización. A continuación, se describe cada escenario y se analizan sus diferencias.

Balance general

La redacción es el punto de encuentro y de intercambio. Es el espacio de negociación y de formación. Al revisar de forma global los escenarios se advierte que, en efecto, no hay una forma única de trabajar una nota de periodismo de datos. Se habla de un proceso discontinuo, no lineal. Sin embargo, como se observa en la caracterización de los procesos, hay rasgos comunes en la forma de trabajo.

En cuanto a los contextos de trabajo, tanto el Escenario E1 como el E2 expresan que deben compartir recursos humanos con otras áreas, lo cual representa un obstáculo pues limita su capacidad de respuesta y se ven orillados a recurrir a apoyo informal entre las áreas o a buscar financiamientos para apoyarse de externos.

De la misma forma, la percepción que se tiene de las unidades de periodismo de datos entre las secciones y áreas de apoyo es que, estas unidades realizan un trabajo que se sale del esquema al que están acostumbrados, pues los requerimientos que hacen de diseño y programación son diametralmente distintos a los de otras secciones.

Por otra parte, al hablar del proceso de visualización se observa que en los tres contextos hay una fase de negociación entre diseñadores, programadores y quienes trabajan la nota acerca del posible producto final a desarrollar. Con este propósito, se realizan pruebas para ver qué se puede armar con los datos, ajustar o replantear. Otra característica en común es el objetivo de desarrollar notas que faciliten la comprensión de los datos de forma sencilla y visualmente atractiva. En los tres casos, los diseñadores son los que ofrecen mayores detalles de lo que se espera de una visualización.

En cuanto a los temas o la agenda que se plantea cada escenario para el desarrollo de notas, en los tres casos existe libertad para seleccionarlos; sin embargo, esta agenda debe atender, como prioridad, las solicitudes del área editorial, o en el caso del Escenario 3 los requerimientos de un cliente.

En cuanto a los flujos de trabajo, la principal diferencia que guarda el Escenario 3 con los otros contextos es que, mientras en los diarios se solicita apoyo a personal que colabora para distintas secciones, en el Escenario 3 hay áreas formadas *exprofeso* para apoyar los requerimientos de los periodistas o analistas de datos. Es así como confluyen de forma sistemática los esfuerzos de diseñadores, programadores y analistas en la generación de productos noticiosos, lo que resulta en flujos de comunicación más fluidos y locales.

Estructura de códigos

Esta categoría reúne los elementos que guardan relación con los requerimientos, los valores del área en la que se desempeñan los participantes, la forma de organizarse y las tareas y obligaciones a las que están sujetos.

También comprende los elementos situacionales que enmarcan el trabajo de los participantes en términos del área en el que trabajan, objetivos del área, percepción de otras áreas, la visión y valores institucionales y el trato con el cliente. La Figura 16 muestra la distribución de la categoría *Espacio de trabajo* que integra a los códigos: Organización, Requerimientos institucionales, Percepción de otras áreas, Valores del área y Trato con el cliente.

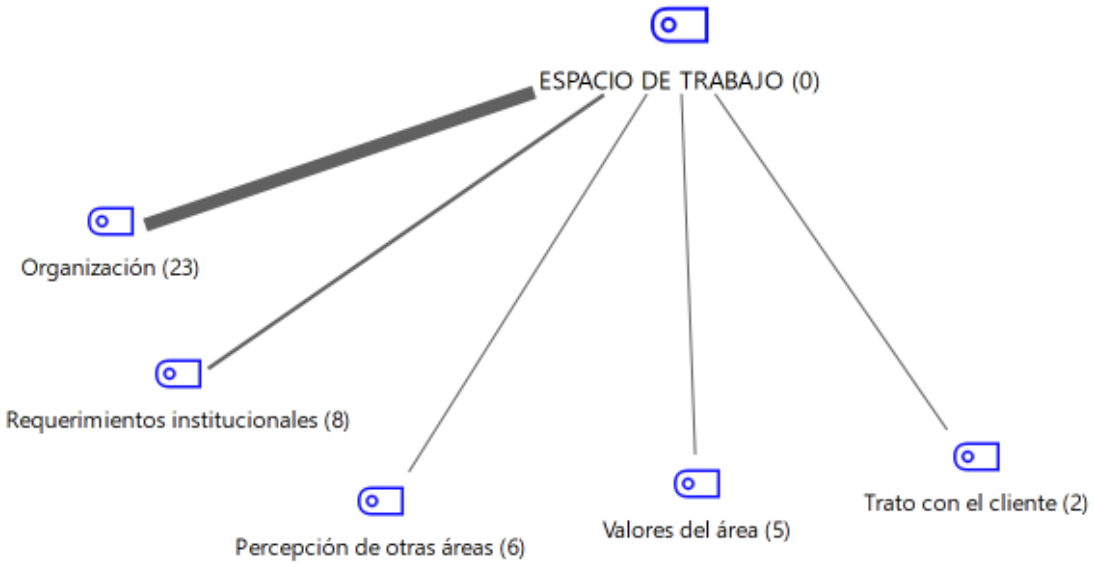


FIGURA 16: ESPACIO DE TRABAJO, ESTRUCTURA Y FRECUENCIA DE CÓDIGOS

En este caso, el código con mayor número de segmentos asociados es Organización (23), seguido de Requerimientos institucionales (8) y Percepción de otras áreas (6).

Escenario 1

Se trata de una de las pocas redacciones de periodismo de datos en un diario de circulación nacional. Con apenas dos años de creación esta área se distingue por contar historias a partir del uso de bases de datos, el análisis y visualización de los datos. La redacción de esta área comparte el espacio con el área comercial y de diseño. En la redacción hay un corcho con anotaciones sobre una de las columnas, hay pegadas notas de otros diarios, recordatorios, una entrevista a una de las reporteras de esta área e ilustraciones —Esas son cosas de Internet y que nos gustan— comenta la editora. A un costado hay cuatro computadoras de escritorio, un archivo en el que guardan los testigos (ejemplares del diario en impreso en el que publicaron una nota) el espacio de la editora y la reportera y a un costado dos becarios trabajando.

Organización

A decir de la editora de esta sección los flujos de trabajo en este escenario están condicionados por el tiempo y los requerimientos de otras áreas. La forma en que las otras áreas afectan los flujos está relacionada con el hecho de que comparten a personal especializado para tareas específicas como el diseño o el desarrollo de infografías. Por ello es necesario contar con determinada organización, pues no siempre se cuenta con el personal para llevarlo a cabo (Editora E1:53).

La principal limitación de este escenario es no contar con suficiente personal, pues esto restringe la capacidad de reacción frente a notas de coyuntura o para procesar datos ricos en contenido. En este caso se tiene claro la prevalencia del valor de la inmediatez, sin embargo, se reconocen limitaciones de personal:

> "Mira, quisiera tener más capacidad de reacción. Porque eso sí, por ejemplo, ahorita, está toda la temática de la violencia contra la mujer, acaban de sacar unos datos sobre las cárceles […] Si tuviera un equipo de, por lo menos, cinco personas, yo creo que sí tendríamos la posibilidad" (Editora E1:19)

El planteamiento de temas a desarrollar o la agenda del área obedecen a iniciativas personales que se apegan a una lógica institucional. En este caso la

editora narra el proceso desde que se plantea una agenda, se atienden las solicitudes del jefe del área editorial y se colabora con otras áreas del diario (Editora E1:15)

La lógica institucional en cuanto a la administración de recursos humanos limita en buena medida el flujo de trabajo, pues la editora debe tocar varias puertas para resolver un requerimiento de otra área relacionado con la nota que desarrolla. Esto implica hacer trabajos de manera paralela e ir de un área a otra (Editora E1: 53). Después de este proceso el material gráfico se concentra en el departamento de fotografía para que se suba a una plataforma desde la que se colocan los pies de foto y se les agrega una cabeza (Editora E1:59). Estas actividades son realizadas por el diseñador, el editor y el reportero.

Valores institucionales

El sistema de valores de esta área se concentra en hacer un esfuerzo por transparentar información de valor social para un conjunto de lectores, cuestionar, analizar y validar datos para brindar notas de periodismo de datos claras y comprensibles para un amplio rango de lectores (Editora E1: 21)

Requerimientos institucionales

Lo que espera el diario de esta sección es que al menos tenga presencia una vez a la semana en el diario en su versión impresa. En cuanto a la agenda que se plantea hay la posibilidad de concretar temas propios con interés periodístico, pero también deben atender solicitudes del área editorial que implica colaborar con otras secciones (Editora E1:15). Esto requiera hacer una pausa en otros proyectos para darle prioridad a la solicitud del diario.

> "Uno de los directivos lanzó la idea de 'yo creo que periodismo de datos podría entrarle a este tema'. Nosotros encantados. Claro que, en ese momento, hay que ponerle pausa a lo otro que estábamos haciendo y sacar (la encomienda del directivo)." (Editora E1:15)

Percepción de otras secciones

Es claro que el tipo de productos que elabora esta área representa una forma distinta de hacer periodismo, de tal modo que el área editorial puede solicitar trabajos con temáticas de otras secciones. Por otra parte, las relaciones que se pueden establecer con otras entidades, en este caso, de corte académicas brinda la oportunidad de pensar en otro tipo de productos que los distingan de su competencia por el tratamiento de la información y el valor añadido, en este caso colaborando con otra sección del mismo diario.

> "Trabajamos con un equipo de analistas y geofísicos expertos en análisis de datos. Porque ellos, platicando con nosotros, vimos que teníamos coincidencias en proyectos. […] Ellos aportaron la visualización de los datos y los análisis de los datos de todos los mundiales que ha participado México, cuántos goles ha metido México en todos los mundiales en los que ha participado, y era una visualización bastante agradable." (Editora E1:17)

Escenario 2

La sección de datos se ubica a un lado del área comercial del diario. En el momento de la entrevista tenía 6 meses de haberse formado. Es un espacio pequeño con mesas en la periferia y una mesa central para el trabajo colaborativo. Esta mesa tiene una división en el centro que inhabilita el trabajo colaborativo. La editora que coordina esta área señala que es una modificación que hicieron durante su ausencia —es el muro de Berlín— menciona en tono de broma, no entienden la forma en que trabajamos. El área de periodismo de datos está conformada por un analista de datos, reporteros, un diseñador y un ingeniero de datos y visualizaciones.

La forma en que la editora señala que conformó a su equipo de trabajo fue muy peculiar, pues los conoció en un hackaton una especie de taller para trabajar con datos y en ese momento trabajaron un tema sobre agua. Ella está convencida de que fue muy afortunada de coincidir con personas con un interés común por generar periodismo de datos. Señala que la forma de desarrollar el periodismo se ha modificado pues en el caso de su área no hay periodistas que trabajen como

“lobo solitario” o que señale que “esta nota es mía”, en ese sentido prevalece más el sentido de colaboración.

La dinámica de trabajo consiste en reunirse diariamente en la mañana para discutir los avances de las notas, la dinámica de trabajo consiste en asignar a un líder de proyecto de acuerdo con sus habilidades. En las etapas iniciales hacen un trabajo de “pre-reporteo” una especie de exploración de temas que sirve para descartar o confirmar la potencialidad de un tema para convertirlo en nota.

La editora señala que no hay periodistas “multitask”, pues se requiere el apoyo multidisciplinario para crear contenidos no sólo para medios impresos sino también para radio, televisión e Internet.

Organización

En un inicio se planteó un esquema horizontal para los flujos de trabajo, en el que todos participarían en la concepción de una idea, sin embargo, la forma de trabajo se ha decantado más por dividir el trabajo de acuerdo con la especialización de cada integrante.

> “Y eso finalmente no paso, cada quien por así decirlo se agarraba su parcela, los dateros analizaban la base de datos, los reporteros reporteando y éste estableciendo una narrativa y [B] por su lado estableciendo esa narrativa en el diseño de los datos.” (Equipo E2:43)

Al asumirse como un área que genera productos no tradicionales dentro de la estructura del diario se reconoce que hay requerimientos que salen de los esquemas de trabajo a los que están acostumbrados en otras áreas. En este caso se valen de colaboraciones informales en forma de “favores” para beneficiarse y lograr concretar una nota o un producto que complemente su trabajo.

> “hay gente con buena disposición, aparte el tipo de colaboradores o gente que llega sin saber, ellos colaboran, lo han hecho bien y de buena fe, y lo ves un poco en los videos que hemos sacado, la mayoría ha sido o casi todos han sido así.” (Equipo E2:79)

Valores institucionales

El sistema de valores de esta área se enfoca a innovar, generar cosas nuevas, lo que implica una nueva forma de pensar y de crear, así como una mayor inversión de tiempo. La editora habla de experimentar y equivocarse, jugar con hipótesis y hacer un tipo de periodismo basado en métodos científicos en los qué planteas un problema y lo válidas y resuelves de una forma lógica. Se utiliza la analogía del laboratorio lo cual supone crear cosas distintas.

> "El problema que nosotros tenemos es nuestro nombre dice laboratorio, entonces eso quiere decir que innovas propones cosas distintas a los moldes que ya están preestablecidos, exacto, con nosotros sí se puede dar esa dinámica, si nos podemos salir del lugar de confort." (Equipo E2:75)

La relevancia periodística, menciona la editora, es una de las variantes que consideran al momento de desarrollar un tema, algunos elementos que determinan este factor son el interés, la relevancia, el descubrimiento de algo nuevo y el sentido de brindar información oportuna. De acuerdo con uno de los reporteros, se asume que la coyuntura no necesariamente dicta la agenda de lo que publican. Ellos van generando su propia agenda. Este desmarque de la agenda periodística es una forma de señalar que la elección de sus temas se rige bajo otro sistema de valores, ajenos al periodismo del día a día, es decir al diarismo.

> "Nosotros vamos a contrapelo de la coyuntura periodística actual, la coyuntura nos influye pero no nos dicta la agenda, o sea por ejemplo tenemos un tema de la cruzada, es un tema que no está, tú no abres el periódico de hoy y vez un tema de la cruzada… Nosotros no tenemos nada de desapariciones no nos interesa, bueno al menos he notado que no nos interesa tanto como otros, nosotros marcamos nuestra propia agenda." (Equipo E2:52)

Requerimientos institucionales

La expectativa de este diario, a nivel institucional, supone que se publiquen notas de gran resonancia de forma recurrente. Sin embargo, se deja a un lado que este tipo de periodismo requiere de una mayor inversión de tiempo por lo que no "se puede medir con la misma regla" que el periodismo tradicional (Analista 2 E2:83)

Hay una exigencia tácita de proponer temas, con distintos niveles de dificultad y de requerimientos, así como de plantearse nuevas formas de narrar para hacer la diferencia frente a los demás medios (Analista 2 E2:83).

Percepción de otras secciones

En cuanto a la percepción que se tiene del área de datos al interior del diario es narrada por uno de los participantes como una visión de que se hacen cosas distintas "que se salen de los esquemas" lo cual implica trabajar de otro modo los procesos que en otras áreas se realizan bajo determinados esquemas (Diseñador E2: 81). La lógica de la producción de productos de periodismo de datos se asume como distinta al diarismo tradicional no se les puede "medir con la misma regla" en cuanto exigencias tácitas de publicar notas de corto plazo, rápidas y con gran resonancia en medios (Equipo E2: 83).

Escenario 3

El espacio de trabajo en este caso no es propiamente una redacción se trata de una empresa que desarrolla notas para otros medios. Sus instalaciones se ubican en un edificio blanco de la primera mitad del siglo pasado. El mobiliario es de madera, es improvisado y sin acabados, las paredes tienen pizarrones con trazos de distintos colores: hay fechas de entrega y diagramas de flujo, también se distinguen ejercicios de lluvia de ideas. Las mesas son compartidas y para hablar con su equipo mueven los monitores para mantener el contacto visual, sobre la mesa hay dibujos a lápiz y trazos a color, las formas son variadas como si se tratará de dibujos hechos para matar el tiempo durante una reunión o para centrarse en una idea. La disposición del espacio está orientado al trabajo colaborativo.

El analista menciona que diariamente tienen una reunión para ofrecer detalle del estatus de cada proyecto. Después de esa reunión cada equipo trabaja en su área. Las oficinas no tienen puerta, la circulación por el edificio es abierto —si alguno de los colaboradores requiere de apoyo de otra área puede acercarse de

forma inmediata— menciona el analista. Hay un área de diseño, otra de programación y una de análisis de datos.

Organización

En este caso la agenda de trabajo depende mucho de los requerimientos de los clientes. El flujo de trabajo considera, en primer término, la definición del tema en el área de datos, después la intervención del área de diseño y programación. Hay mucha comunicación en paralelo entre las tres áreas. Lo cual establece una dinámica iterativa.

> "Exacto. Obviamente diseño hace su parte, platicamos mucho con [A], etcétera. Pero todo el rebote de todos los proyectos, o sea, van y vienen, van y vienen, van y vienen comentarios, etc. Y pues nosotros nos encargamos de hacer eso." (Analista 1 E3:16)

Buena parte de la definición del producto se realiza entre las áreas de Datos y Diseño. Hay un proceso de negociación de ideas que van y vienen hasta que queda definido. "es como un pim pom de ideas, hacemos una lluvia de ideas hasta llegara a algo que se llama accesible, legible, adecuado, informativo o atractivo." (Diseñador 1 E3:18). En esta dinámica el área de diseño funciona como un nodo o puente para hacer legible la información al lector final, se concibe como la parte central que da coherencia a los datos (Diseñador 2 E3: 16).

Una vez que estas dos áreas aprueban pasa nuevamente por Datos y se integra al área de Programación para valorar la funcionalidad y entrar en un proceso de negociación y refinamiento de la idea entre las tres áreas para llegar a la visualización de los datos. La integración de saberes de los tres equipos de trabajo (Diseño, Datos y Visualización) y su combinación es de acuerdo con uno de los participantes la forma ideal de trabajar. (Programador E3:20). Se reconoce que hay un último filtro en este caso del editor general quien es el que tiene la última palabra junto con el equipo de datos (Analista 2 E3:11).

Valores institucionales

En este escenario parte de los valores son trabajar con rigor metodológico que incluye formularse preguntas, asentar una hipótesis y probarla (Editor y Analistas E3:46). Otro valor es la libertad de expresión.

> "Tenemos libertad total de escribir lo que queramos y nunca hemos tenido un problema de "no pueden publicar esto." (Editor y Analistas E3:18)

Trato con el cliente

A diferencia de otros escenarios, en este caso los requerimientos institucionales están muy relacionados con la figura del cliente. La figura del editor general opera también como representante pues una vez que se cierra un contrato con un cliente da la señal para que las tres áreas comienzan a trabajar y es el área de datos la que se encarga de que el cliente quede satisfecho con el trabajo (Analista 1 E3:12).

PARTE II ESTRATEGIAS DE ACCIÓN

Del dato a la nota

Este capítulo recoge parte de los procesos que se ponen en juego durante la elaboración de una nota de periodismo de datos. En este caso el proceso abarca desde el proceso inicial o de origen de la nota, la selección de un tema hasta la construcción de la nota en términos de elaboración, cambios de ruta y estrategias de validación. Si bien en el trabajo con los datos, los obstáculos, los elementos de programación y visualización también están imbricados en el desarrollo de la nota, éstos se tratarán en extenso en otro capítulo.

Balance general

Confeccionar una nota, en la práctica del periodismo de datos no se limita a ubicar un tema, difundir un boletín o transcribir una declaración, es un trabajo en conjunto cuyo producto es la suma de distintos saberes. A lo largo de este proceso la relación que los participantes establecen con la información y los datos es dinámica, pues buscan respuestas, cuestionan, ponen en balance lo hallado con lo que se quiere comprobar.

Los participantes se apropian de esquemas mentales estrechamente imbricados que se reflejan en acciones como buscar constantemente motivos para trabajar una nota, sea en forma de cuestionar desde otro ángulo la información

hallada, descargar un conjunto de datos de internet y revisar si “hay algo”; volver a las notas trabajadas o a solicitudes de información; o incluso poner en juego ese sentido, ese esquema incorporado, al tomar el metro y ver el aviso de personas extraviadas y saber que esos datos de edad, complexión, estatura, sexo… pueden sistematizarse para generar nueva información susceptible de convertirse en una nota.

La interacción con la información es constante, se consultan fuentes, se reciben resúmenes de notas, se revisa lo que publican otros medios (la competencia), se valora la agenda mediática, se revisa la forma en que esos otros medios resolvieron una nota, diseñaron una visualización compleja o trataron un tema.

Pocas cosas definen tanto a un medio como su agenda. En los casos narrados se arguye, que para proponer un tema de trabajo se hace a partir de una agenda personal, de líneas de investigación implícitas del área, de la coyuntura o de asignaciones institucionales. Trabajar un tema y darle forma a un producto noticioso representa también un proceso de negociación con todos los involucrados, pues el producto se rige bajo los términos de lo posible en términos técnicos, pero también temporales. Bajo esa lógica se reformula, se replantea, se rehace a partir de una amplia gama de posibilidades de lectura. Se evalúa constantemente el estado del desarrollo de una nota y se cuestiona qué se tiene y qué se puede contar con esto. Pueden tener claridad sobre “hacia dónde va tu nota” sin embargo esta idea cambia durante la interacción con los datos.

La evolución de la nota está condicionada, en buena medida, por lo que los datos respalden. En caso de que los datos “no den” se buscan otras alternativas o se desecha el tema propuesto.

Estructura de códigos de Desarrollo de la nota

El Desarrollo de la nota es el proceso con el mayor número de segmentos codificados (161). Forma parte de la categoría superior de Valor agregado a la

información. Al interior de su estructura, (Figura 17) las subcategorías asociadas, son *Origen de la nota* (48); *Construcción de la nota* (52) y *Obstáculos* (61). Esta última categoría la analizaremos en un capítulo más adelante.

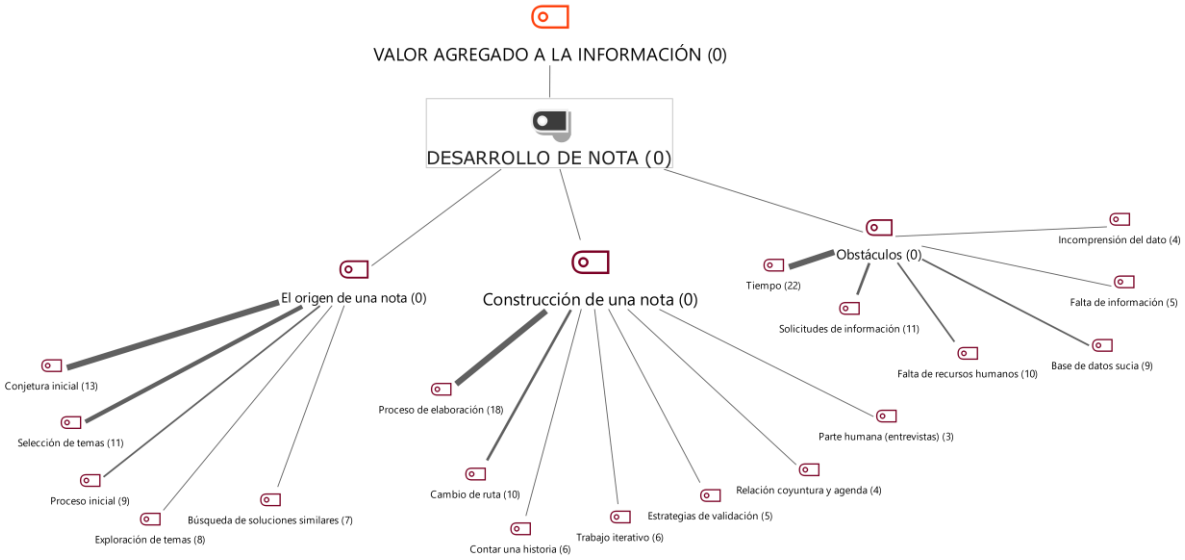


FIGURA 17: DESARROLLO DE LA NOTA, ESTRUCTURA Y FRECUENCIA DE CÓDIGOS

El origen de una nota

Proceso inicial

Todo comienza con una pregunta por responder, una hipótesis por comprobar, una revisión de los datos que se tienen para contar una historia. De acuerdo con lo que nos señalan los entrevistados (Reportera E1, Equipo E2), el origen de una nota no tiene una prescripción, no se trata de un proceso lineal; la idea de escribir sobre determinado tema puede surgir al descargar un conjunto de datos, al ingresar a una página de Internet (Equipo E2: 48), al revisar los datos que se obtuvieron de una solicitud de información al gobierno, al cuestionar los datos desde otro ángulo (Reportera E1: 2), o bien, al participar en espacios de intercambio de experiencias con programadores (Editora E1: 25). En otras situaciones porque el diario en el que se trabaja define que un tema es prioridad y el área de periodismo de datos lo puede sacar adelante (Editora E1: 15).

De acuerdo con lo que nos señalan los entrevistados, los periodistas de datos tienen una disposición latente de encontrar temas que puedan tratarse con datos. Acorde con ello la idea de escribir sobre determinado tema puede tener varias fuentes de inspiración, no siempre es el mismo camino, pero siempre hay un interés permanente por encontrar historias y generar notas.

> "Otro camino que tomamos es que estás viendo… por el mismo consumo de medios te topas con una base de datos interesante, no tiene que ser coyuntural, puede ser una base de datos, no sé, chistosos. Hemos hecho cosas de los Óscares, de futbol, de lo que sea. Y dices: 'ah, está padre esto'. No hay ahorita nada coyuntural que quisiéramos hacer; hacer gráficas, buscas preguntas que puedes contestar e igual, el mismo proceso pero con un tema más, no tan coyuntural pero que tiene que ver con datos." (Editor y Analistas E3: 21)

La relación que hay entre quienes hacen periodismo de datos y las fuentes de información es estrecha, indisociable. Se toman decisiones a partir de lo que se encuentra o de lo que se carece, de lo que emerge de una lectura inquisitiva, o bien de imaginar lo que se puede obtener si se sistematizan datos desestructurados. El participante es visto entonces como un agente que ha incorporado una serie de disposiciones reflejadas en un interés tácito de buscar temas, de analizar datos, de buscar enfoques o complementar información para dotar de valor un producto.

Observamos que en el E1 uno de los primeros pasos para comenzar a elaborar una nota es el análisis de los datos. Hay una búsqueda constante de material susceptible de contarse con datos. Un ejemplo de ello es el caso de los avisos de personas extraviadas, el cual es visto como una oportunidad de obtener, sistematizar y generar nueva información con dichos datos (Reportera E1:18). Una de las reporteras de este escenario plantea una analogía con otro tipo de fuente, pues indica que se puede tener muy clara la idea de hacia dónde va tu nota y, conforme hablas con la gente, esta idea cambia; lo mismo ocurre con los datos duros (Reportera E1:11).

Por lo que corresponde al Escenario E2 hay una similitud con el Escenario E1, pues se comienza con el análisis de los datos, guiados por la idea de que puede haber un tema susceptible a desarrollarse; sin embargo, se advierte que no hay una

manera específica de comenzar, ya que cada caso puede ser diferente (Analista 2 E2:48).

A diferencia de los escenarios anteriores, en el Escenario E3 se habla de formularse preguntas sobre lo que se quiere investigar o comprobar, a partir de ello se busca la información (Programador E3:26); es decir, antes de emprender la búsqueda de información o de datos, se formulan preguntas que definen los requerimientos de esta información. Hay un departamento en específico que trabaja con los datos y son ellos quienes definen en buena medida lo que se quiere contar y con qué datos se cuentan para hacerlo (Diseñador 2 E3:54).

A partir de las entrevistas podemos identificar que, en estos casos, los detonantes o fuentes de inspiración de los participantes ocurrieron:

- Al descargar un conjunto de datos o al ingresar a una página de Internet (Analista 2 E2: 48).
- Al revisar los datos que se obtuvieron de una solicitud de información al gobierno (Reportera E1: 2)
- Al cuestionar los datos desde otro ángulo (Reportera E1: 2)
- Al participar en espacios de intercambio de experiencias con otros miembros del equipo (Editora E1: 25),
- Cuando el personal del diario en el que se trabaja define que un tema es prioridad (Editora E1: 15).
- Al analizar temas socialmente relevantes. Por ejemplo, los avisos de personas extraviadas son vistos como insumos para obtener, sistematizar y generar nueva información con dichos datos (Reportera E1:18).
- Al analizar temas coyunturales. En los tres escenarios se observa que no hay una desvinculación de la agenda o de temas de coyuntura; pues los periodistas se mantienen informados y conocen los temas sobre los que está hablando.

Al comparar los escenarios podemos afirmar que el E3 tiene ventajas sobre los otros dos ya que hay un departamento específico que analiza los datos y define

las preguntas sobre lo que se quiere investigar y los requerimientos de datos, para a partir de ello buscar la información necesaria.

Exploración de temas

Los periodistas de datos realizan diferentes acciones para determinar si un tema tiene potencial para ser una NdePd. A este proceso se le ha denominado "exploración de temas" y requiere de procesos de análisis permanentes, entre los que se identificaron:

- Reflexión constante para identificar si se tiene "algo enfrente" (Reportera E1: 9), si es hay una tendencia, o si los datos son relevantes por su temporalidad o su alcance geográfico, o si "empiezas a ver algo... esto ya me está jalando ruido" (Analista 2 E2: 94). Hay un proceso inquisitivo que se pone en juego frente a los datos, del cual puede emerger un tema y entonces lo que sigue es delimitar.
- Lectura y exploración de bases de datos y noticias para identificar ideas y generar preguntas. Esta aproximación implica que el periodista se informe como parte de su rutina y que mantenga un constante estado de alerta sobre los temas que se está hablando o que "salen en los medios todos los días" (Reportera E1: 41)
- Intención de emular una nota publicada en otra fuente con datos locales. En esta acción el punto de partida es la propia competencia de otros diarios nacionales e internacionales (Analista 2 E2: 94). Al respecto se mencionan como fuentes de inspiración *The New York Times*, *The Guardian*, *El Mundo* o bien las redes sociales como son Facebook y Twitter.

> "Ah, por ejemplo, este que sacó *El Mundo* [señala el artículo pegado en la columna] lo vimos y dijimos: 'ah, mira, está padre cómo poder graficar datos de los diputados e incluir los tuits y está padre que la puedes incrustar dentro del sitio y queremos hacer algo parecido'. (Editora E1:56)

- Exploración de bases de datos especializadas o informes. Por ejemplo, las del INEGI, o como la de la Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública (ENVIPE), o la Encuesta Nacional de Seguridad Urbana (Reportera E1: 4).

La aproximación a temas por trabajar en forma de nota supone en los tres escenarios la constante exposición a noticias. En los tres casos está presente esa disposición de explorar noticias o datos para “empezar a ver algo”. En el Escenario E1 se habla de esta exploración, pero también de atender una agenda propia que obedece al medio en el que se trabaja o a intereses personales.

Se observa que la dinámica de trabajo influye en la aproximación a los temas de la NdePD. Al respecto el reportero del E1 señala que va identificando temas, mientras realiza otras actividades cotidianas. En los tres escenarios se considera que la lectura y exposición a productos noticiosos es una condición favorable para explorar temas. Otra coincidencia entre los tres escenarios es que refieren el trabajo que ha elaborado la competencia en el ámbito nacional e internacional, lo que incluye la identificación de temáticas que no se están cubriendo.

Selección de temas

La agenda periodística es vista como el conjunto de hechos o situaciones relevantes para una sociedad en concreto, los cuales se espera sean tratados en los medios de información en un espacio y tiempo determinados. Estas expectativas son cubiertas a través de la “selección del tema”, proceso que se refiere a la forma en que las redacciones evalúan y determinan las notas que se elaboraran.

La dinámica en las redacciones al hablar o proponer temáticas responde tanto a la existencia de una agenda propia, como a la asignación de temas por parte de los directivos del diario. La selección depende en gran medida del contexto y de las estructuras de organizaciones horizontales y verticales; pero también en concordancia con los temas debido a la coyuntura o los intereses personales.

En el análisis se identificaron las siguientes formas de selección:

- Por asignación. Las prioridades se establecen de manera jerárquica como se observa en el caso de la Editora E1:15. En donde se señala que “uno de los directivos lanzo la idea de yo creo que el periodismo de datos podría entrarle al tema. Nosotros encantados”. Lo que se puede observar es la apertura de este grupo de periodistas para asumir retos.
- Por la coyuntura. Se reconoce la influencia de la coyuntura en la selección de temas, por la Editora E1 narra el caso en el que la publicación de un informe de un tema coyuntural por parte de una institución es el hecho que debe aprovecharse para generar agenda.

  “En cambio, te digo, bueno, ayer fue la publicación de este informe de la CNDH sobre las cárceles, entonces en ese instante es cuando uno tiene que agarrar el informe y decirle al reportero “siéntate a analizar los datos ahorita”. Entonces, ya, desde ayer los analizó, hoy ya me trae un avance, y el día de mañana, quizás, podamos sacar algo.” (Editora E1:21)

- El dato sobre la coyuntura. Otra forma de selección se refiere a la alternativa de selección en donde los datos son más importantes que la coyuntura, dado que se identifica que son interesantes y que tienen valor para ser narrados, al respecto el señalan

  “No hay ahorita nada coyuntural que quisiéramos hacer, haces gráficas, buscas preguntas que puedes contestar e igual, el mismo proceso, pero con un tema más, no tan coyuntural pero que tiene que ver con datos.” (Analista 2 E3:21)

- Análisis grupal. La selección del tema se puede derivar de una situación en donde un grupo de profesionales comparte información, identifican las coincidencias y se conjuntan las habilidades para analizar datos, crear visualizaciones y conformar la noticia de datos.

  “Platicando con nosotros, vimos que teníamos coincidencias en proyectos. Y entonces ellos nos dijeron: ‘nosotros traemos un proyecto que tiene que ver con el tema de los mundiales’... ellos aportaron digamos que las historias, los reportajes, los textos, y nosotros, a través de estas personas. Ellos aportaron la visualización de los datos y los análisis de los datos de todos los mundiales en que ha participado México, cuántos goles ha metido México en todos los mundiales en los que ha participado, y era una visualización bastante agradable y, al mismo tiempo, útil, porque incluso podías moverle, picarle. Y eso fue pues… a raíz de una plática en la

que estábamos viendo, bueno, ¿cómo nos insertamos en este momento con tanta información que hay del tema de los mundiales." (Reportera E1:17)

- Agenda personal. Tanto en el Escenario E1 como en el E2 se habla de una libertad para marcar una agenda noticiosa propia. Esta puede corresponder a intereses individuales, a líneas de investigación que se han tratado en el pasado y que se les puede dar continuidad o bien, a un tema coyuntural como puede ser un momento político. Al respecto la Editora del E1 señala que "Yo trato de empezar siempre como una agenda que pudiera sonar como muy personal" (Editora E1: 15).
- Solicitudes de los clientes, este tipo de selección de temas sólo está presente en el Escenario 3.

Lo que podemos anotar, es que en la selección de temas inciden tanto influencias del contexto organizacional, el entorno social o los clientes de una organización; así como la propia iniciativa e interés del periodista. También se observa una tendencia a tomar decisiones de manera grupal a partir del análisis de publicaciones o informes o por medio del intercambio de ideas con otros colegas que encuentran una temática coyuntural, como un mundial, para colaborar. Es decir, cada escenario tiene sus propias formas de seleccionar las notas del periodismo de datos sin que haya una forma única o estable.

Conjetura inicial

La conjetura inicial se relaciona estrechamente con la intuición, pues funciona como una brújula que lleva a los participantes a analizar con minuciosidad un conjunto de datos. Las conjeturas se presentan como pensamientos, preguntas, hipótesis o intuiciones que motivan a los periodistas a buscar datos. Las acciones para elaborar conjeturas son variadas, A continuación, se exponen las formas identificadas en los periodistas para hacer conjeturas:

- Pensar mal. Una conjetura puede surgir a partir de una intuición expresada en la forma de "piensa mal y acertarás" (Editora E1:31), como producto de

este pensamiento se hace una deducción lógica que lleva a la elaboración de la conjetura.

- Preguntas iniciales. La exposición y consumo de noticias, como se vio antes, está vinculado a varios procesos, en este caso se usa para la formulación de preguntas que conducen a la elaboración de una nota.

> "Nos gusta consumir noticias… Entonces, cuando hay algo interesante, en este caso en particular el asesinato de la alcaldesa y todo lo que se ha dicho del Mando Único… Entonces, es decir, qué tanto ha funcionado el Mando Único. O sea, el gobierno estatal ha publicado sus datos, pero no sabemos bien cómo lo calcularon, y empiezo con esas preguntas iniciales: ¿sí ha funcionado el Mando Único? … Pues vamos a comparar la violencia dos meses antes de que entrara el Mando Único, o tres meses, todavía no lo decidimos. Y el mismo lapso después de que entró el Mando Único. Vamos a ver si hay un salto en la violencia y o si disminuyó, y ya…, y eso va a ser algo sencillo." (Analista 1 E2:20)

- Hipótesis e intuición. El analista 3 del E2 (85) marca una distancia entre lo que puede verse como una hipótesis y la intuición. De acuerdo con este participante, la hipótesis es un planteamiento puntual que se comprueba o se descarta. En el caso de la intuición se trata de una idea no acabada que conduce el trabajo a partir de la información con que se cuenta; es decir, es la propia información la que va marcando un camino. El Programador (E3:12) concuerda con esta forma de concebir una hipótesis.

Esto conlleva a la necesidad de buscar datos e información, para llegar a una conclusión, la cual se expresa con afirmaciones. Estas afirmaciones se formulan como hipótesis que se arraigan al pensamiento del personal y conducen la búsqueda de datos e información.

Búsqueda de soluciones similares

La búsqueda de soluciones o notas similares a las que se están trabajando o se piensa trabajar es una práctica recurrente. La consulta de la versión electrónica de diarios como *Washington Post*, *The Guardian* o *New York Times* y su sección *The*

Upshot funciona como una forma de nutrirse, de saber cómo se han resuelto con sencillez temas complejos, también es un punto de inicio para bocetar o una forma de ejemplificar y comunicar lo que se pretende desarrollar en una nota.

> "De cómo visualizar o graficar presupuestos que ya se hicieron en otra parte del mundo o hasta ejemplos muy concretos o muy fáciles que se han publicado en el *New York Times* o en el *Washington Post* y que ves y, 'mira, así graficaron los presupuestos'; 'así graficaron para ver recursos de dinero, dinero público', y ya en función de ello decimos 'podemos hacer algo similar algo parecido'" (Editora E1:33)

En el caso del escenario E1 sirve como una forma de ubicar ejemplos que puedan ayudar a transmitir a otros colaboradores una idea de lo que se quiere lograr como resultado con una nota que se pretende desarrollar; así como conocer la forma en que se construyó una nota en particular. La constante exposición a medios que realizan periodismo de datos en los Escenarios E1 y E2 es un ejercicio continuo que permite conocer soluciones a problemas comunes en términos de visualización, interacción y nuevos formatos de Storytelling.

Construcción de una nota

Una vez establecido el tema y la conjetura, el siguiente proceso es la identificación de los datos e información que se necesitan, su búsqueda, tratamiento y limpieza, para su posterior análisis y conclusión con la edición de la versión impresa y digital. Procesos que llevan tiempo y que por lo regular no son evidentes para el usuario final.

Proceso de elaboración

Como se ha señalado previamente el proceso de elaboración de un producto noticioso está asociado con la dinámica de trabajo de cada escenario. Sin embargo, la interacción con los datos es uno de los primeros elementos con los que se trabaja en los contextos revisados. La delimitación, la búsqueda y la interacción con los datos en forma de exploración son de los primeros pasos que se dan. El punto de partida de la nota es el área de datos, ya que es donde se define lo que se pretende hacer y se establece dónde ubicar los datos (Diseñador 2 E3: 54).

Todo esto repercute en las tareas que se asumen, las tareas que se delegan y el tiempo de elaboración. Por ejemplo, en el E3 se realiza un trabajo colaborativo en los proyectos en curso, esto supone una interacción entre todos y una constante comunicación. En el E3 el punto de partida es el área de datos, pues es donde se define lo que se pretende desarrollar y se ubican los datos con los que es posible realizar el producto.

> "Las dos áreas que les puede tomar más tiempo en sacar como la nota periodística, una es la de datos, porque obviamente hay veces que los datos no existen, hay que buscarlos, hay que sacar solicitudes de información, hay que limpiarlos, procesarlos, hay que encontrar una hipótesis, algo que se quiere comprobar y ver si los datos respaldan esa hipótesis o no, y si es relevante, periodísticamente hablando." (Programador E3:12)

A esta etapa le sigue la limpieza de datos la cual está relacionada con dos procesos, revisar y ordenar los datos para darles significado (Reportera EI: 21). En caso de que las personas tengan dudas sobre los datos y la información obtenida, lo que se valora es buscar datos alternos, se estipula el nivel de detalle que se requiere y las fuentes que se utilizarán para elaborar el producto noticioso, lo que lleva a la construcción de una hipótesis (Analista 1 E3: 20).

La nota es entonces un ensamble en la que colaboran muchos actores para su realización, Hay una parte humana que es tan importante como los datos que se analizan e integran en la elaboración del producto noticioso. La parte humana, en este caso se entiende como el conjunto de entrevistas, información de campo o de "reporteo" que complementa las cifras (Reportera E1: 25).

Los Escenarios E1 y E2 guardan similitud en relación con que en ambos contextos se comparten recursos humanos con otras áreas. En el Escenario E1 el proceso de elaboración está marcado por las limitaciones de personal, pues el equipo es muy reducido y comparten un equipo de diseñadores con otras áreas. Por su parte, en el Escenario E2, al trabajar con personal que apoya en otras áreas del diario, sus solicitudes salen comúnmente del esquema de trabajo al que están acostumbrados, ya que se trata de productos muy distintos a los que se generan en otras secciones.

En el caso del Escenario E1 se habla del tipo de salida que se le daría al producto en términos de formato, pues además de los procesos de búsqueda, análisis, comprensión y limpieza de los datos, solicitudes de información, planeación y elaboración, deben ocuparse de darle seguimiento al formato de salida en versión impresa y digital.

Por lo que corresponde al E3, la elaboración de notas de periodismo de datos en este contexto se apega a un flujo condicionado por el tiempo en el que colaboran distintas áreas con tareas particulares. La principal diferencia que guarda este escenario con los otros es que, mientras en los diarios los procesos se concentran en una sola área y se solicita apoyo a personal de apoyo común, en el E3 son varias áreas, las que confluyen de forma sistemática para generar el producto.

Trabajo iterativo

Durante el desarrollo de un producto noticioso hay un nivel considerable de interacción entre distintas áreas, pues se integran aportaciones o saberes especializados de diseño, programación o análisis que se "rebotan" entre sí para darle forma a una nota o bien, para hacerla accesible, legible, adecuada, informativa o atractiva. Rebotar en este contexto puede entenderse como una forma de compartir una propuesta y esperar una respuesta en relación con la misma de forma repetida.

> "Entonces es como un *ping-pong* de ideas, hacemos una lluvia de ideas hasta llegar a algo que se llama accesible, legible, adecuado, informativo o atractivo." (Diseñador 1 E3:18)
>
> "Tac, tac, tac, y al final, y luego programación, y sí como que ya son cosas de estar como rebotando constantemente, siempre es así, nunca es así como lineal, como 1, 2, 3." (Diseñador 1 E3: 59)

Los datos se comparten, se interactúa y se juega con ellos con el fin de armar una propuesta que pueda integrarse con otra de distintas áreas. Son comunes las referencias a la iteración en forma de onomatopeyas; por ejemplo: "ping-pong" o "tac, tac, tac", "hasta que quede bien" (Diseñador 2 E3:25). El proceso de rebote tiene como principal propósito definir y tomar decisiones en relación con el producto

que se elabora. El proceso lleva tiempo, por ejemplo, en el caso de la Reportera (E1:38) menciona varias reuniones semanales, probando y proponiendo.

Es así como la nota refleja una suma de adiciones, de valores diferenciados que se van sumando en una unidad que va tomando forma; se juega, se ajusta hasta que tenga una forma definitiva.

Las diferencias entre un escenario y otro radica en el hecho de que en el escenario 3 los flujos de comunicación con otras áreas son más fluidos, se realizan de manera local, a diferencia del E1 que debe recurrir a instancias externas al diario, dificultando así la comunicación.

Cambio de ruta

El trabajo progresivo en una nota supone tomas de decisión en el camino. Las conjeturas funcionan como guías que orientan los procesos descritos anteriormente; sin embargo, durante el trabajo con los datos puede ocurrir que estas conjeturas se desechen o se modifiquen sustancialmente con nuevos ángulos o hallazgos que tengan más fuerza.

Como consecuencia de lo anterior, las conjeturas iniciales se refuerzan, sufren modificaciones, se desechan o se alteran sustancialmente, dependiendo de los resultados del análisis de datos. A continuación, se presentan los casos de cambio de ruta identificados.

- Replantear la situación para fortalecer la conjetura, Este cambio de ruta es positivo porque se refiere cuando una hipótesis o conjetura se fortalece a partir del análisis de los datos. Por ejemplo, cuando se identifica un ángulo que le da más fuerza a una nota a partir de un tratamiento distinto de los datos disponibles. Este es el caso que narra la Reportera del E1.

> "Eso sí nos ha pasado, o replantear la situación. Eso también, porque muchas veces… o sea… vamos a sacar algo sobre enfermedades, hay que replantear la forma en que estás analizando las cosas... Nos salía los tumores malignos, que es cáncer, y nos salía como muy desperdigados… Entonces le dije a [D]; 'si lo dejo así parecería que casi no hay cáncer en México; si los junto todos sí se sube a una gran posición'. Pero eso pasa después como de varios análisis. Qué dije… Sí hay tumores malignos, pero que como que los dividen por tipo: en el riñón, en la boca, en el no sé qué… Están muy diseminados y no se nota tanto su fuerza. Entonces

necesitaría juntarlos a todos y que sea todo cáncer y después seguir. Entonces, ese tipo de cosas sí te pasan, que empieces el análisis y digas no lo puedo dejar así porque la gente va a pensar que no hay cáncer en México y, no, es que sí hay." (Reportera E1:13)

- El dato no me da. Las primeras interacciones con los datos permiten estimar si se cuenta con suficiente información para respaldar una nota. En caso de que la persona que trabaja una nota identifique que no hay datos que respalden su idea, hipótesis o intuición, entonces busca más información o cambia, la replantea. En esos casos se habla de que el "dato no me da tanto" (Reportera E1:13).

- Los datos no se pueden mapear. En el siguiente extracto ejemplifica la tensión del periodista al no poder encontrar y organizar los datos. A este problema lo denomina mapear.

"Entonces, llega un momento en el que dices "no hay forma". Los datos no te dan como para que puedas desarrollar una investigación que realmente, o sea. Si quieres dar un primer hitazo (sic) sería la investigación que te dice "no hay forma de cómo mapear estos datos", y la otra es que no tienes una forma tan realista como la que tú quieres mostrar." (Editora E1: 21)

- Hablar con la gente. Este cambio de ruta se deriva de la información "humana" que compila el periodista y que le obliga a modificar el foco de su investigación. Al respecto la Reportera E1:11, señala que "que conforme hablas con la gente, conforme reporteas en la calle" la hipótesis inicial va cambiando.

- "Vuelta de tuerca". Este proceso se observa como un cambio de 360 grados en la conjetura inicial, en donde los datos no "cuadran", situación que se presenta en la fase final de la construcción de la nota. Lo cual supone buscar otra opción.

"De hecho, desde que comenzó, cuento ésta como vuelta de tuerca en cuanto al tema de cómo visualizar los datos… Entonces, incluso yo ya tenía los ejemplos de cómo iban a visualizar los de gobierno, No cuadraban los datos y los números con las gráficas que ellos querían hacer. Entonces, pues digo eso, o sea, no hacen falta

dos dedos de frente para saber que, si no puedes mezclar peras con manzanas, pues no puedes, y tienes que buscar otra opción. (Editora E1:49)

- ‘‘Nos está dando todo lo contrario”. En esta situación el periodista se enfrenta a que su análisis no contribuye a confirmar la conjetura planteada inicialmente. Un ejemplo de ello es lo descrito por el Analista 2 E2 (88), quien señala que después del análisis de datos se advierte que “nos está dando todo lo contrario, o sea la intuición general estaba completamente equivocada”.

- No era tan interesante. Este cambio de ruta se presenta cuando después de analizar los datos, se identifica que esta no era tan interesante (Analista 3 E2: 87).

- Nos dijeron que estaba mal. En el E3 se puede observar que el analista tuvo que cambiar de ruta después de que se entrevistó con un grupo de académicos, quienes señalaron que “estaba mal’. “No había considerado una cosa, no había considerado un factor propio que hay que considerar y fue rehacer la base para considerarlo” (Analista 1 E3: 33).

En función de lo anterior, se puede afirmar que la elaboración de una nota tiene como una de sus características la posibilidad de cambiar de ruta. Esto se atribuye a una variedad de situaciones que inciden en el rumbo del producto en donde hay un vínculo claro entre el trabajo con los datos y la comprobación de las conjeturas e hipótesis.

La obtención del dato es, entonces, el regulador e incentivo que permite estimar su valor como respaldo o argumento en el tratamiento de un tema de una NdePD. Todos los casos anteriores implican la constante reflexión y análisis de los involucrados para saber si deben detenerse y cambiar la estrategia, el análisis o el enfoque para estar en la ruta correcta.

El periodista en estos casos debe estar abierto a reafirmar o descartar las conjeturas iniciales de su investigación. Lo mismo ocurre con el diseño de una propuesta final, hay ocasiones que, al exponerse a soluciones o referentes no necesariamente similares, se logra “destrabar” (Diseñador 2 E3: 91). Se observa que las situaciones negativas son las que prevalecen sobre las positivas, por lo que el periodista debe buscar con alternativas y formas de solucionar los problemas que se le presentan.

En cuanto a experiencias negativas de cambio de ruta el escenario E1 habla de los primeros acercamientos a la información, pues se narra un caso en que los datos desde el inicio no brindan información suficiente para respaldar una hipótesis, es decir, los datos “no dan”. Por su parte en el escenario E2 se narran experiencias en las que después de trabajar prolongadamente con los datos se advierte que lo que se encuentra “no es tan interesante”. Por último, en el caso del E3 se habla de una ocasión en que se invirtió mucho tiempo para trabajar con los datos y tras validarse con académicos se tuvo que rehacer.

Contar una historia

Una nota incluye también un ingrediente que es la narración o inclusión de una historia la cual propicia una fácil comprensión de la información y ayuda a que sea percibida como significativa. El ensamble de la nota no sólo incluye la interpretación del dato, sino también se conforma de un discurso narrativo que busca atraer la atención del lector contándole una historia. Hay una constante actualización o una constante búsqueda, según uno de los participantes, de conocer nuevas formas de contar historias incluyendo datos (Reportera E1:44).

Un dato es significativo para los lectores cuando permite generar historias cercanas a su cotidianidad o contexto "No sé, por ejemplo, hablando del agua [...] que la gente pueda bajar esas historias donde ellos viven, la historia para mí que vivo en Iztapalapa es diferente para quien vive en la Condesa (Analista 2 E2: 159).

Se observa también que el trabajo de campo se relaciona más con la visión del diarismo y sus prácticas como la inclusión de entrevistas a especialistas, pobladores o testigos de un suceso. Para uno de los participantes el periodismo de datos se vale del diarismo para una mejor comprensión de la realidad (Reportero 1 E2: 178)

En los tres escenarios son recurrentes las referencias a la idea de “contar una historia” (Diseñador 1 E3: 44) de ofrecer claridad y una fácil comprensión de temas que, a primera vista, pueden ser complejos pero que con el tratamiento adecuado pueden “bajarse esas ideas” y ser percibidas como simples.

Parte humana (entrevistas)

La inclusión de voces permite contextualizar el dato. Incluir la parte humana es incluir las voces de expertos o involucrados. Este mecanismo permite aligerar una nota pues "si dejamos nuestras notas sin eso, se pierden son muy pesadas y la gente no las lee" (Reportera E1: 27)

La integración de entrevistas conforma una de las piezas adicionales al ensamble o dispositivo de comunicación que, en este caso, da forma a la nota de periodismo de datos. Hay referencias a este elemento en los escenarios E1 y E2 (Reportero 1 E2: 31)

Relación coyuntura y agenda

La coyuntura es el elemento que, de forma habitual, marca la agenda en los diarios o medios informativos. Esa es la idea que subyace del análisis de los fragmentos de este código. Sin embargo, hay casos como el de uno de los reporteros del equipo E2 en el que se desmarca de esta coyuntura para marcar su propia agenda (Reportero 1 E2: 54).

En otros casos puede ocurrir que la publicación de un informe por parte de una institución sea vista como una coyuntura que deba aprovecharse para generar agenda y un tema tenga una mayor resonancia en los diarios. Como en el caso narrado por la Editora del E1 (21) en el que al enterarse de que la Comisión Nacional

de Derechos Humanos acababa de publicar un informe sobre las cárceles, le solicita a una persona que le apoye y se siente a analizar los datos en ese momento,

La práctica en este caso consiste en conocer la agenda periodística para retomar, desmarcarse o enriquecer una nota.

En los Escenarios E1 y E2 se narran casos muy distintos; por una parte, se reconoce que hay una influencia de la coyuntura periodística, pero ésta no les “dicta la agenda” al Equipo E2. Si bien ellos observan lo que hacen otros medios, asumen que no hay un interés sobre los temas que otros tratan. En cambio, el E1 menciona un informe de la CNDH recién publicado sobre las cárceles; ya que este hecho marca una agenda para varios medios, se pide a un reportero que analice los datos para publicar algo al respecto.

Estrategias de validación

La validación de la información que se integra a una nota, puede verse también como un añadido, pues reviste de credibilidad a un cúmulo de datos. La correcta interpretación de hallazgos e interpretaciones es un requerimiento de consultoría que permite saber "que estábamos leyendo de forma correcta y no al aventón" (Reportera E1:36)

A veces el proceso de verificación lleva a replantear o rehacer una base de datos (Analista 1E3:33) En el caso del E3 la estrategia de validación de la nota completa consiste en compartir la información con académicos para que sea revisada antes de que se publique.

> “Toda la línea del tiempo por municipio, y después de esa base entrevistamos a académicos que nos dijeron “estaba mal”. No había considerado una cosa, no había considerado un factor propio que hay que considerar y fue rehacer la base para considerarlo.” (Analista E3: 33)

Trabajar con datos

Este apartado integra la forma en que son vistos los datos como insumo para la elaboración de una nota, los procesos en los que se identifica la necesidad de trabajar con éstos, delimitarlos, obtnerlos y construirlos. Se explora también los requisitos para sus uso desde los procesos de extracción y limpieza, su correcta interpretación, el significado de “Jugar con los datos” y la forma en que los datos intervienen con la idea de la nota que se trabaja.

Balance general

Si hubiese una analogía para el trabajo con datos, se diría que es un proceso artesanal, pues al igual que un artesano que antes de trabajar el barro debe prepararlo, el periodista debe asegurar que los datos cumplan con determinadas condiciones como lo son su limpieza, es decir que sean consistentes para permitir su sistematización y comparación. También hay un proceso de moldeado que busca dar forma a un productos, este proceso es visto como un juego en el que se hace y rehace una idea de lo que se puede formular con los datos. Hay un involucramiento activo en el que se habla de “meterse con los datos”, formularse preguntas y dar forma a tendencias.

Siguiendo con la analogía del artesano, podemos pensar en una situación en la que sus materiales o sus insumos no se prestan o no están listos para lo que pretende realizar. En el caso de los periodistas se observa que los datos “no dan lo suficiente” es decir no le sirven para respaldar una idea o conjetura. Lo cual lleva a replantear la idea de lo que se pretende desarrollar o en el peor de los escenarios la idea se deshecha.

En otro orden de ideas, se habla de la carencia del datos por diversas circunstancias como los son la inoperancia gubernamental para atender solicitudes puntuales, la inconsistencia de los datos proporcionado o bien la inexistencia de los datos requeridos. Estos son algunos de los obstáculos que los participantes afrontan para sacar adelante su investigación la construcción manual del dato, o su ubicación y sistematización provenientes de distintas fuentes son algunas de las alternativas que el participante asume para el cumplimiento de su tarea.

Estructura de códigos de Interacción con los datos

La categoría *Interacción con los datos*, que también forma parte de la categoría superior *Valor agregado a la información*, concentra 56 segmentos codificados, los cuales se encuentran distribuidos en tres grandes subcategorías: Trabajar con datos (29); Identificación de la necesidad de datos (16); Tratamiento de datos (2) y el código Conceptualización del dato (9), (ver, Figura 18).

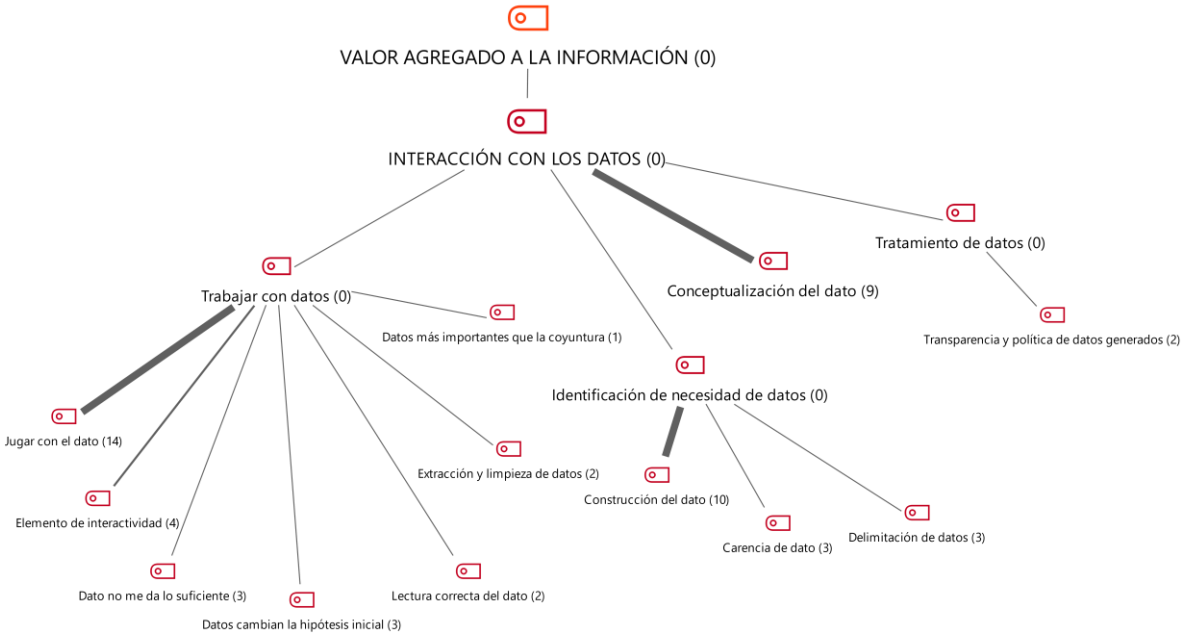


FIGURA 18: INTERACCIÓN CON LOS DATOS, FRECUENCIA

Conceptualización del dato

El dato es visto como un insumo, un argumento, una fuente y un arma de doble filo. Su valor como argumento radica en la procedencia de su fuente. Es un insumo maleable cuyo manejo lleva a comprobar una hipótesis, a desecharla o a encontrar una de mayor fuerza. Tal como lo define la Reportera del E1.

> "Los datos son… un arma de dos filos… si tú traes una hipótesis muy clavada en la cabeza, tú puedes hacer casi cualquier cálculo para que salga eso, pero entonces hay… De decir mi hipótesis se cayó, lo que quería demostrar se cayó o si continuo… sí continua pero además encontré otro ángulo que es mucho más fuerte." (Reportera E1:9)

La confiabilidad de los datos para ser utilizados como argumento reside en su origen. En el caso narrado por la Reportera del E1 (30) se habla de una marcada reticencia a utilizar datos de organizaciones civiles para trabajar una nota relacionada con desapariciones, pues se corre el riego de ser descalificados por organizaciones gubernamentales, por no tratarse de fuentes oficiales. En cambio, si se utilizan este tipo de fuentes, es decir fuentes gubernamentales, puede resultar

que los datos sean erróneos, pero esto ocurrirá porque los datos del gobierno están mal.

El dato es un insumo no dado, es un recurso que se trabaja con herramientas como Excel y las matemáticas. Su validez, reside en su origen y tratamiento.

> "Es periodismo, pero ahora tienes una herramienta base y tu herramienta son datos, tu nueva herramienta es Excel y las matemáticas. Eso te da más sustento cuando vas y hablas con la gente y a nosotros nos ha servido cuando vas a instancias oficiales y le dices ¿encontré esto? Y te dicen ¿pero de dónde? Pues de tus datos. No, bueno… es así, digo… nada." (Reportera E1:29)

Los datos son vistos también como elementos que aportan objetividad y transparencia. Es también un recurso que se trabaja, trata y cruza para ubicar "historias que están ocultas" (Analista 1 E2: 167) y contarlas. Es justo el trabajo que se realiza con los datos lo que permite brindar mayor confianza con el lector, pues se brinda mayor transparencia y esto puede ser visto como un valor adicional "Te da ese nivel adicional de transparencia que es muy poderoso." (*idem*)

Como periodista de este campo, el uso de datos en las notas es visto también como un elemento que empodera, pues el dato brinda objetividad. "Te empodera como periodista, porque tiene elementos objetivos de una certeza que no te da una declaración de ahí." (Reportero 1 E2:171)

Hay un vínculo claro entre el trabajo con los datos y la hipótesis. En el caso del Escenario E1 se menciona que una hipótesis se cae o se comprueba; o bien se encuentra un ángulo con mayor fuerza. Esta probabilidad de cambiar una idea que se tiene en mente es latente cuando se interactúa con datos. En el caso del Escenario E2 se habla del dato como un elemento que genera confianza con el lector, brinda grados de objetividad y transparencia. En ambos escenarios el dato es visto como un elemento para respaldar un argumento.

Identificación de la necesidad de datos

Delimitación de datos

La delimitación de los datos es un proceso que lleva a identificar lo que se necesita y se establecen las estrategias para encontrar estos datos. Como en el caso de la información, la delimitación es uno de los primeros pasos a seguir para emprender una búsqueda.

> "Muchas veces comenzamos con el hecho de tener cierta idea de una historia y decir '¿habrá datos sobre eso…?'; empezamos a buscar esta información o a veces comenzamos a revisar solicitudes que ya contestaron y a ver qué datos son los que ya entregaron y ver de qué otra forma se le puede sacar otro ángulo." (Reportera E1:2)

Formular una definición de lo que se requiere, buscar estos datos, enfrentar la posibilidad de que sean datos inexistentes o negados por instancias gubernamentales y otros organismos no garantiza acceder a los datos requeridos, tampoco asegura que estén limpios o que sean congruentes y con valor periodístico. De esta forma la obtención de datos puede ser frustrante, pues ocurre que las dependencias y organismos no los tienen, están reservados, no están organizados o requieren de un tratamiento sistemático para su interpretación (Programador, E2:12). Por ello, al buscar datos, el periodista se enfrenta la posibilidad de que una nota no se concrete, independientemente de lo depurado de las estrategias aplicadas.

Construcción del dato

La construcción del dato es el proceso que sigue el periodista de datos cuando agota las alternativas para hacerse de datos relevantes para desarrollar una nota y entonces opta por compilar y capturar manualmente los datos por medio de la consulta a otras fuentes como las hemerográficas, Internet o los reportes que se publican en medios de comunicación masiva (Reportera E1:7).

> "Hace dos años hicimos uno sobre policías detenidos por crimen organizado entonces se tuvo que hacer toda una reconstrucción hemerográfica de cuatro años, pero eso no lo hay, en ningún lado te lo van a dar. Entonces, fue como trabajo propio." (Reportera E1:7)

Concebir un dato para tomarlo como insumo implica, en ocasiones, hacer un levantamiento de información de distintas fuentes, consultar una por una y concentrarla de forma sistemática. Se integra una “basesita de datos” para emplear el término de la editora de E1 (21).

> “Yo me acuerdo de algunas notas que hemos hecho algunas han surgido del literalmente rastreo manual de la web, levantar información porque no existen data set, durante x tiempo, yo recuerdo que levantaron información sobre otra federación, no es como decir vas al dataset y lo ves, vete a las fuentes de información pública una a una esa es una manera.” (Analista 2 E2: 53)

Otro caso narrado refiere a los datos proporcionados, pero no accesibles. Es decir, los datos son entregados en este caso por las Secretarías de Estado, pero en formatos cerrados o con restricciones técnicas para trabajar con ellos y la única alternativa es capturar la información de forma manual.

> “Te pongo un ejemplo: hemos estado haciendo todo este proyecto de lo de los viajes transparentes porque es hacer solicitud de información de cada uno de los funcionarios de las secretarías de Estado y, aunque nos los dan en formatos que son totalmente cerrados porque te dan puros PDF, puros documentos que no hay forma de que los puedas tú ver y pegar, ahí vamos haciendo la basecita de datos.” (Editora E1:21)

En el caso del Escenario E1 la Reportera (16) narra el caso del trabajo periodístico sobre desapariciones, la cual se basa en las descripciones asentadas en las fichas de personas desaparecidas, esta información había sido negada por tratarse de datos confidenciales. Sin embargo, el propósito de este trabajo era visibilizar datos como los rasgos, el nombre y la edad de mujeres desaparecidas. Ante la imposibilidad de sistematizar esta información que se encontraba disponible pero no accesible, se determinó estructurar una base de datos y capturar estos datos de forma manual.

El acceso a la información es un problema común en los Escenarios E1 y E2. Se mencionan los problemas con las solicitudes de información, los formatos y la incomprensión de los datos. La obtención del dato, en ambos escenarios, es un proceso paralelo con la construcción de éste. Cuando el dato o los datos son inexistentes la alternativa que se toma es construirla de distintas fuentes,

capturarlos de forma sistemática y trabajarlos manualmente. En otras palabras, el dato que se busca para desarrollar una nota toma forma a partir de la obtención de datos de distintas fuentes y la capacidad para estructurarlos y sistematizarlos.

Carencia del dato

Los datos son un detonante de una investigación, la mayor parte del tiempo se invierte en “buscarlos […] sacar solicitudes de información, limpiarlos y procesarlos” (Programador E3: 12). Se puede tener una hipótesis en mente y emprender la búsqueda de los datos que la sustenten, sin embargo, el reconocer que no hay datos o que los datos “no te den” como expresaba la Editora el E1 (21) ponen en perspectiva al participante sobre la viabilidad de una nota.

Trabajar con datos

Extracción y limpieza del dato

Como se ha señalado en reiteradas ocasiones el dato es un insumo no dado, requiere a veces de un trabajo de búsqueda o de captura, de sistematización y de limpieza. En este caso escrapear es un proceso de extracción de un grueso de datos alojados en un portal y la limpieza es un complejo proceso de sistematización, consistencia, corrección, eliminación e integración de datos con miras a hacerlos comparables o manejables por medio de programas especializados. Este tipo de tareas consumen tiempo en la elaboración de un producto noticioso al que se añade valor al integrar este tipo de elementos.

> “Puede sonar que es mucho, pero a veces te tardas dos o tres días en encontrar los datos que realmente necesitas. Otros dos días en limpiar en analizar, y ahí ya se te fue toda una semana.” (Reportera E1:25)

Una base de datos muy sucia puede requerir la intervención de un programador, del uso de software especializado y de técnicas sofisticadas como lo menciona la Reportera (1E:16).

> “Entre las dos lo hicimos y estado de México tenía una página similar a la de CAPEA, pero trabajé también con unos programadores y uno de los programadores me dijo, déjame ver si la puedo escrapear. Que es como ver el back de la página para jalar todo, de Domain se pudo jalar todo, me dijo nada más hay que limpiar y eso… pero ya es una chamba mucho más fácil. Porque incluso, me dijo, pude jalar las fotos. Le dije ah pues está bien.” (Reportera 1E: 16)

Acorde con lo anterior la limpieza consume tiempo y requiere de conocimientos especializados, todo ello con el fin de agregar valor en la nota de periodismo de datos, pues los datos limpios y consistentes permiten avanzar en el desarrollo de la nota, pues son un requisito para la visualización y la programación.

Jugar con el dato

Al hablar de “jugar con los datos” los participantes refieren la interacción con éstos, el cuestionamiento, la búsqueda de patrones o tendencias para moldear o desechar una idea. Se empieza con las preguntas más simples, se analiza desde lo más mínimo, se diseccionan los datos y se cuestiona lo obtenido.

> “A veces no, a veces si tienes como una idea. O sea, sí tienes como una idea planteada, pero es mucho empezar a jugar con ellos.” (Reportera E1:9)

Jugar con el dato guarda una estrecha relación con la Intuición, pues la interacción con los datos o el juego con ellos es motivada por esta percepción de una idea formada o por formarse. De esta forma jugar con el dato permite al periodista dejar que “la información nos indique por dónde va la cosa.” (Analista 3 E2:85). Algunas de las estrategias identificadas en este juego con los datos son:

- Meterse con los datos. Esta estrategia consiste en el involucramiento activo y profundo. Para ello los periodistas no deben tomar los datos ya procesados, sino que deben "meterse con los datos”, para hacer su propio análisis.

  > El punto es quieres hacer algo realmente con datos métete. Es meterte ni siquiera a las Queries que tiene INEGI en la que tú nada más pides cosas y te saca tablitas, no. Metete hasta atrás metete a los microdatos de INEGI empieza a analizar que quiero buscar por estados, por años voy a sacar más, voy a sacar menos quiero buscar tendencias, quiero enfocarme sólo en el último semestre de 2016. Se requiere de muchas más preguntas de lo que la gente cree que se hacen.” (Reportera E1:31)

- Formularse preguntas y delinear tendencias. En este proceso, tal como refiere la Reportera del E1 implica un cuestionamiento constante de los datos para identificar si hay algo, al respecto señala "se requiere de muchas más preguntas de lo que la gente cree que se hacen", para poder llegar a una tendencia.

- ¿Por qué no buscar algo más? Aquí el jugar con los datos implica buscar nuevas formas de analizarlos y visualizarlos con la finalidad de tener nuevos ángulos que permitan visualizar tendencias de forma clara y sencilla, atractiva y enriquecida.

  > "Sabemos que podemos mostrar lo datos con claridad, sencillez, con unas barras, pero ¿por qué no buscar algo más?, por qué no buscar una forma más atractiva, más rica a me está faltando este dato pero además este otro y porque no actuar de esta forma." (Diseñador 2 E3:47)

- Rebotar los datos. Esta estrategia consiste en pasar los datos de una parte del equipo a otro. Por ejemplo, de la redacción al equipo de visualización. Esta estrategia permite mirarlos desde varias perspectivas, lo que enriquece el análisis.

  > "Una primera limpieza digamos se comparten los datos con la gente que va a visualizar para que ellos también comiencen a jugar un poco y vean que visualizaciones podemos armar. Nosotros les proponemos, mira la base trae esto creemos que se puede hacer esto y esto, que podemos jugar de esta manera y es como un a ver... va mándamela, juego un poco con ella y veo si, sí se puede armar." (Reportera E1:38)

Jugar con el dato nos indica que no hay una manera inicial específica de trabajar con los datos, cada caso es diferente. Sin embargo, se reconoce el papel de la intuición, la reflexión y el auto-cuestionamiento permanente de lo que se tiene enfrente, para conducir el análisis.

El trabajo de jugar o rebotar datos es una tarea común en los tres escenarios, sea con el propósito de explorar una idea o con la finalidad de mostrarla por medio de una visualización de los datos obtenidos.

En el E1 se observa una forma inicial del trabajo con los datos, misma que se transmite a los practicantes, se habla de formular preguntas simples, ubicar tendencias, datos por año, delimitar el alcance geográfico del análisis. Se hace énfasis en trabajar de forma directa con los datos y profundizar en el análisis de estos. No limitarse al tratamiento que le haya dado la instancia emisora en forma de comunicado, “hacer tu propio análisis” cuestionar para llegar a la idea.

Por parte de uno de los participantes del E2 se advierte que no hay una manera inicial específica de trabajar con los datos (Analista 2 E2: 48), cada caso es diferente. Sin embargo, se reconoce el papel de la intuición en el juego con los datos, pues es la propia información la que va conduciendo el análisis.

En cuanto al equipo que se dedica a visualizar los datos en el E3 se habla de “rebotar” los datos con la finalidad de visualizarlos de forma enriquecida sin perder de vista la claridad y la sencillez.

Datos cambian la hipótesis inicial

La idea de una nota o un tema funciona como una hipótesis que conduce en buena medida la búsqueda de datos, en ese sentido la interacción que se tenga con éstos se encuentra condicionada por una conjetura que se busca demostrar. Durante la interacción con los datos, puede ocurrir que esta hipótesis se deseche o se desplace por un nuevo ángulo o hallazgo de mayor fuerza ya hemos citado la percepción de la reportera del E1 que advierte que los datos son un arma de dos filos, pues se corre el riesgo de empecinarse con una hipótesis o ángulo determinado (Reportera E1:9).

En el caso del E3 se asume que se es transparente en aquellos casos en que los datos son contrarios a lo que se pensaba como resultado. Por lo que se menciona es algo recurrente en este contexto.

> “Pero sí. Si encontramos que los datos van en contra de lo que nosotros pensamos que iba a salir, somos muy transparentes en ese sentido. Nos pasa mucho.” (Analista 1 E3:51)

La interpretación de los datos y su continua contrastación con una hipótesis es un proceso iterativo de cuestionamiento y prospectiva, en el que se pone en juego los datos que se tienen y la idea de lo que se quiere producir a partir de éstos.

Lectura correcta del dato

Los datos deben ser comprensibles como parte del análisis, esto implica realizar una serie de estrategias, entre las que se encuentran:

- Consultar glosarios. Esta situación se presentó en el escenario E2 debido a que había términos que no se comprendían pues no había un glosario de cómo era empleado cada término.
- Consultar a expertos. En estos casos el periodista reconoce que no "entiende" dado que son temas que escapan de su dominio por lo que deben recurrir al apoyo de expertos en el tema para asegurar la correcta lectura de los datos.

  "El de fideicomisos fue muy complicado, porque no éramos expertas en el tema, cuando abrimos la base, había muchos campos que no entendíamos. Pues yo no entiendo, hay que hablar con alguien que nos diga qué significa." (Reportera E1:47)

- Validar los datos. La validación de datos implica la consulta con el organismo emisor de los mismos, a fin de que ayude a darles significado a los datos y a las variables.

  "Teníamos campo de dinero en negativos y decíamos esto qué… si lo agarro así nos cambia mucho el análisis, qué quiere decir esos negativos… entonces habla a Hacienda. Que te hagan caso en hacienda y ya que te hagan caso, qué quieres… no pues sobre esta base en específico porque tienes negativos y ellos decían ¿cuáles campos negativos? Pues tienes campos negativos, qué significan … y decir, no entiendo." (Editora E1: 25)

En los tres escenarios se hace evidente que la comprensión del dato es muy importante ya que de la lectura y correcta interpretación es lo que dará validez y confiabilidad a una nota.

El dato no me da lo suficiente

Las primeras interacciones con los datos permiten estimar si se cuenta con suficiente información para respaldar una nota. Si los datos existentes no respaldan suficientemente una idea, hipótesis o intuición, entonces se busca más información o se replantea. En esos casos se habla de que el "dato no da lo suficiente". En otras palabras, se tiene un dato existente, pero el usuario determina que el dato no es funcional para el propósito que se persigue.

> "Por lo general el análisis es de las primeras cosas que hacemos. Entonces… lo que sí ha pasado es que desde el inicio dices las cifras la verdad no me dan tanto… o necesito buscar más o necesito mejor cambiarlo. Eso sí nos ha pasado, o replantear la situación." (Reportera E1:13)
>
> "O bien te dices la verdad es que los datos no me dan, no me dan algo dé lo suficiente. Digamos, ya le buscaste ya le rascaste y no hay nada significativo. Y ha ocurrido a la inversa de que tienes una hipótesis, una idea, y lo quieres validar con datos… pero no hay datos." (Reportera E1:14)

La idea de la nota o tema que se desea desarrollar es un regulador e incentivo al momento de valorar los datos que se obtienen. Es también la medida que permite estimar su valor como respaldo o argumento en el desarrollo de esta nota o en el tratamiento de un tema. En este contexto el dato no da lo suficiente cuando se identifica que este insumo no respalda o no puede operar como un elemento que brinde validez a la idea que se pretende exponer.

El dato es más importante que la coyuntura

La agenda periodística es vista como el conjunto de hechos o situaciones relevantes para una sociedad en concreto, puede ser vista como aquellos temas de interés o coyuntura que se espera sean tratados en los medios de información en un espacio de tiempo determinado. En el diarismo tradicional la coyuntura marca en buena medida los temas de la agenda. El siguiente extracto narra otra alternativa para determinar un tema y trabajarlo con datos.

> "Otro camino que tomamos es que estás viendo, por el mismo consumo de medios te topas con una base de datos interesante, no tiene que ser coyuntural, puede ser una base de datos, no sé, chistoso. Hemos hecho cosas de los Oscares, de futbol, de lo que sea. Y dices "ah, está padre esto. No hay ahorita nada coyuntural que quisiéramos hacer, haces gráficas, buscas preguntas que puedes contestar e igual,

> el mismo proceso, pero con un tema más, no tan coyuntural pero que tiene que ver con datos." (Analista 2 E3:21)

En este tipo de casos, no hay un tema concreto en mente durante la búsqueda de datos, pero sí un tipo de disposición o interés incorporado por contar un tema con datos. En este caso el tema no se articula a partir de la coyuntura, sino de la posibilidad de armar una nota con datos.

Elemento de interactividad

El desarrollo de elementos interactivos en la nota es visto como un valor adicional para la nota que permite propiciar una mayor cercanía entre el lector y la información.

> "Acercando la nota a algo más íntimo porque puedes interactuar con la información, te da una medida de transparencia sobre todo que estamos sentados como mejores prácticas, que fue lo que hicimos." (Analista 1 E2:167)

La interactividad es un requisito implícito de los productos que se desarrollan. Es un proceso que supone también el hecho de traducir una idea a un entorno tecnológico con posibilidades de salida en distintos dispositivos (Programador E3:14)

En ambos escenarios se observa que la interactividad es un elemento adicional que le añade valor a la nota en la presentación de los datos que la sustentan. En el caso del Escenario E2 el participante resalta la cercanía del lector con la información por medio de este elemento y en el E3 el Programador destaca la dificultad de pensar con otro ángulo lo que debe ser la nota. Se observa que una diferencia esencial entre ambos escenarios es que el E3 es el único entorno que cuenta con un departamento y un editor dedicados a trabajar en la interactividad de las notas que desarrollan.

Transparencia y política de datos

El E2 es el único en el que uno de los participantes habló abiertamente de una de las políticas que se tienen en su contexto sobre el manejo de datos, la cual destaca la transparencia y la apertura a compartir datos y códigos una vez publicada la información.

> "Bueno, tenemos la política de publicar siempre nuestros datos, ser muy transparentes. Y si alguien nos contacta para decirnos "oye, pásame esto", nosotros no tenemos bronca de pasarlo. Hay mucha gente que es muy celosa con sus datos, acá no tanto. Más bien, si vamos a publicar algo. Pues sí somos un poco celosos, pero ya una vez que se publica, si hay alguien que lo quiera en diapositivas, etcétera, pues sí les damos todos nuestros datos, les damos códigos. A mi más de una vez me han llegado a preguntar "oye, ¿cómo hiciste esto?, ¿qué significa esto?", usualmente amigos." (Analista 1 E3:52)

PARTE III CONDICIONES INTERVINIENTES

En este apartado se trata las condiciones que influyen en las estrategias que emprende el participante, en este caso el tiempo, la falta de recursos humanos, la falta de información, los obstáculos para solicitar información, las bases de datos sucias y la incomprensión de los datos por parte de quién trabaja con ellos.

El tiempo y otros obstáculos

El tiempo es visto como una dimensión finita que juega en contra casi siempre para quienes desarrollan este tipo de periodismo, no sólo se trata de quien lo percibe sino también del tiempo de los demás involucrados, del tiempo de inversión en recopilar, organizar, limpiar, visualizar o programar. Por otro lado, la falta de datos para solventar una idea, la frustración, el proceso para hacer una solicitud de información, una plataforma sin mantenimiento, bases de datos sucias y la falta de recursos humanos son algunos de los principales obstáculos que se recogen en este capítulo

Balance general

Hay una clara asociación entre el tiempo, el producto deseado y los recursos con los que se cuentan, sea en forma de nueva información requerida o la falta de personal especializado para el cumplimiento de tareas adicionales como capturar, limpiar, visualizar o generar un producto audiovisual. El tiempo resta “capacidad de

reacción”, posibilidades de ser más creativos e innovadores, Es un referente para estimar el alcance de lo posible y de aquello que requeriría más tiempo.

Otros obstáculos en los procesos de elaboración de una nota se asocian con la deficiente atención de los sistemas de transparencia que pueden retrasar entre 20 o 30 días el desarrollo de la nota, los formatos de entrega pueden abonar a este retraso, más si se tratan de PDF’s de imagen o datos dispersos en distintas instancias. Los datos confidenciales son otro elemento que se suma a este tipo de obstáculos que el participante debe afrontar, pues en ocasiones este recurso se utiliza para negar información importante para la nota que trabajan.

En otros casos el dato que se requiere para el desarrollo de una nota no existe y, por lo tanto, se estima en tiempo si es posible capturarlo de forma manual, construirlo a partir de consultar distintas fuentes o construir una base de datos propia para este propósito.

Hay obstáculos que dependen del acceso o existencia de un dato, pero también hay obstáculos personales como lo son los propios esquemas de percepción o de conocimiento que se ponen en juego para la correcta interpretación de un dato, sea por el nivel de especialización sobre una disciplina especifica o bien por la forma en que están dispuestos los datos en un conjunto o set de datos.

Estructura de categorías y códigos

Esta subcategoría forma parte de la categoría Desarrollo de una nota, la cual ya hemos descrito en el capítulo denominado Del dato a la nota. En este caso se observa en la Figura 19 que el código Tiempo es el que presenta el mayor número de segmentos codificados (22), seguido de Solicitudes de información (11) Falta de recursos humanos (10), Base de datos sucia (9), Falta de información (5) y, finalmente, incomprensión del dato (4).

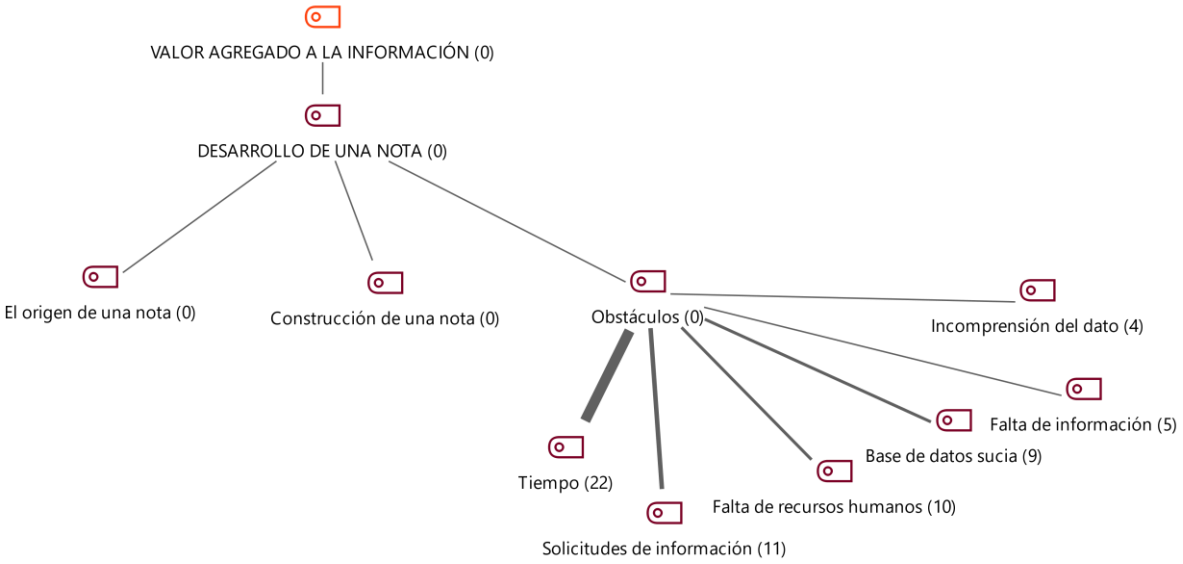


FIGURA 19: OBSTÁCULOS, ESTRUCTURA Y FRECUENCIA DE CÓDIGOS

Tiempo

La relación con el tiempo para el periodista está marcada por el valor de la inmediatez y de lo oportuno, de lo que hace o deja de hacer un medio, en términos de lo que cubre o no. La inversión de tiempo en una nota es el tiempo que se invierte en añadirle un valor agregado que los distinga de otros medios en el competido espacio mediático de las noticias. El tiempo destaca como el obstáculo más importante en la elaboración de una NdePd, ya que puede afectar varios procesos y características de la nota provocando la frustración en los periodistas. En el siguiente ejemplo se muestra los procesos que se realizan en un tiempo determinado y cómo a pesar de parecer un tiempo razonable, no lo es.

> "Pues más o menos, nosotros debemos de entregar de menos uno como éste, cada quince días. De menos. Este fue un poco más sencillo porque toda la parte del reporteo ya se había hecho, porque yo estuve en Cartagena y esa es una historia de Cartagena. Yo ya traía redactada toda la parte humana, ya nada más tenía que buscar todas estas cifras, pero incluso a veces los quince días, puede sonar que es mucho, pero a veces te tardas dos o tres días en encontrar los datos que realmente necesitas. Otros dos días en limpiar, en analizar, y ahí ya se te fue toda una semana y te queda sólo una semana y ni siquiera completa, te quedan netamente como tres días, entre redactar, sacar más entrevistas, buscar la parte humana y terminar todo esto." (Reportera E1: 25)

Dado que la elaboración de una nota requiere la participación de profesionales de diferentes áreas y profesiones, es más difícil estimar el tiempo. Al respecto la Editora E1:23, señala:

> "El tema que a nosotras nos ocupa mucho de tiempo son los procesos, que no son nuestra responsabilidad directa, me refiero al área de infografía, al área de diagramación, etc. El tener que convivir con estas otras áreas, y que yo no soy la jefa, me crea un timing en contra porque, aunque no lo creas, sí es estar luchando." (Editora E1:23)

La carrera contra el tiempo es una constante en todas las áreas y departamentos de las redacciones, sobre todo en las que hay procesos que convergen, como lo es el diseño gráfico para la elaboración de las infografías (Editora E1:23). En los siguientes ejemplos se puede advertir que el *tiempo* es un constante obstáculo durante la elaboración de la nota; es una restricción temporal que traza el alcance de lo posible en términos de nivel de profundidad y características.

> "Hay veces en las que son restricciones de tiempos, entonces dices: 'oye, eso está perfecto. Lo podríamos hacer, pero nos tardaríamos dos meses; entonces necesitamos encontrar otra solución". (Programador, E2:18)

La falta de tiempo causa frustración en los periodistas, por ejemplo, el Equipo de diseño E3:44, señalan "créeme que se puede hacer sólo necesitamos un buen de tiempo". Afecta su creatividad, ya que se trata de tareas nuevas "cosas distintas a los moldes que ya están preestablecidos" que implican "mayor carga de trabajo, capacitación, inversión de tiempo personal" (Diseñador E2: 77). En el siguiente caso se observa como el factor tiempo afecta el diseño interactivo de una nota.

> "Porque muchas veces hemos publicado textos que salen en la versión impresa pero no tenemos como la posibilidad para salir en web con un interactivo, que nos gustaría, pero no tenemos como la forma de cómo lograr hacer eso tan, tan a corto plazo … no tenemos el equipo necesario para hacerlo con tanta velocidad." (Editora 1:7)

El tiempo es visto en los tres escenarios, es un factor poderoso que puede limitar la creación de una nota, el formato de esta y la elaboración del material complementario. La organización del tiempo incluye el plano personal y el de los involucrados en otras áreas de apoyo. Por ello la elaboración de una nota implica

una organización adecuada, pues cada formato requiere de un tratamiento distinto por parte de personal de apoyo, sea para un vídeo o un interactivo.

Falta de recursos humanos

La falta de recursos y el tiempo son inherentes al contexto, los cuales son mencionados reiteradamente como factores que afectan el trabajo y desarrollo de la nota. La relación con otras áreas de apoyo en los escenarios E1 y E2 es vista como un obstáculo, pues se trata de áreas que apoyan simultáneamente a otras secciones del diario ya sea que elaboren infografías, diagramen, programen o editen video, por lo que son secciones que no cuentan con suficiente personal.

Se reconoce, en uno de los casos, que esta limitación de recursos humanos dificulta tener "más capacidad de reacción"; crea conflictos de autoridad y hace que el tiempo juegue en contra (Editora E1:19). Otro ejemplo es el caso del escenario E2 se reconoce que hay propuestas que han quedado fuera por esta misma limitación.

> "Pero siempre se han hecho propuestas, así como, si es verdad que se han quedado un poco fuera en ese sentido porque no alcanzaban los recursos humanos." (Editora E2:70)

Cuando hay más de una versión (digital, impresa o multiplataforma) es posible que se requieran otros productos como un vídeo o una infografía (Analista 2 E2: 70), lo que supone el conseguir la colaboración de más personal, inclusive de otras áreas sobre quienes no se tiene autoridad para establecer prioridades.

Esto mismo ocurre con los programadores cuyo trabajo está supeditado a las posibilidades y prioridades establecidas en el trabajo; por esta situación los periodistas de datos deben acoplarse "con los tiempos" del programador pues éstos no pueden "producir solamente trabajos para periodismo de datos" (Editora E1:11). Una solución planteada en el E1, es buscar financiamiento para contratar programadores y diseñadores que trabajen por proyecto.

> "Precisamente una de las formas que yo he tratado de ir sorteando para evitar precisamente estar caminando al mismo ritmo que ellos, es por eso que nos hemos

> dedicado en el área a buscar financiamientos externos para poder conseguir recursos para programadores o diseñadores que puedan trabajar a la par otros proyectos con nosotros." (Editora E1:11)

La falta de personal especializado en los escenarios E1 y E2 genera conflictos de autoridad, problemas con el establecimiento de prioridades, y una baja capacidad de reacción para trabajar notas coyunturales e inclusive para descartar notas.

Falta de información

A decir de una de las participantes del Escenario E1, la falta de información es uno de los primeros obstáculos a los que se enfrentan al momento de trabajar en una nota. Uno de problemas a los que se enfrentan son las restricciones para el acceso a los datos y a la información de dependencias gubernamentales, pues éstas la reservan al público, la declaran inexistente o la presentan con errores o inconsistencias.

> "En cambio, pedimos información para hacer todo el mapeo sobre los baches. Solicitamos información a cada una de las delegaciones y solicitamos información a cada dependencia relacionada con el tema: información reservada, información inexistente, información que no es consistente con lo que tú estás viendo en el momento." (Editora E1:21)

Otra limitación es la confidencialidad de la información, pues no se puede compartir información que integre nombres, direcciones o número de cuentas. Aquí el caso de la reportera que tuvo que replantear la estrategia para trabajar el tema de desaparecidas:

> "Sí, el que hicimos de las mujeres desaparecidas no sé si lo has visto. De ese teníamos como que la idea, pero los datos no son públicos." (Reportera E1: 16)

Estos obstáculos se presentan en los tres escenarios, y de igual forma se presentan alternativas, como las de elaborar solicitudes; "hay que buscarlos, hay que sacar las solicitudes de información" (Programador E3: 12) o bien se debe trabajar en rastrear manualmente en la web y levantar la información pues no existe un dataset (Analista 2 E2: 52)

Solicitudes de información

El proceso de acopio de datos para el desarrollo de una nota puede prolongarse debido a que las solicitudes que se hacen por medio de Instituto Nacional de Transparencia, Acceso a la Información y Protección de Datos personales pueden tardar entre 20 o 30 días. Esto implica que el periodista cuente con habilidades para lidiar con la burocracia, la calidad del dato proporcionado, la desorganización de la información; entre otros problemas.

Una de las fuentes principales de datos e información para el periodista de datos son las dependencias gubernamentales. Para obtener la información ellos tienen que llenar solicitudes de información en formatos en línea o de manera presencial.

En ocasiones ellos tienen que llenar varias solicitudes debido a que los datos están dispersos o no son equiparables. Al respecto el Analista del E2 (140), señala que, al solicitar información a la SEDESOL, fue necesario llenar solicitudes para cada uno de los 9000 registros que se necesitaban. Otro tipo de obstáculos en ese sentido son:

- Información reservada e inexistente. Las solicitudes no son atendidas por las dependencias gubernamentales, pues se arguye que se trata de "información reservada, información inexistente, información que no es consistente con lo que tú estás viendo en el momento" (Editora E1:21).
- Registros confidenciales. En este caso los datos y la información no están accesibles o se encuentran reservados por contener información confidencial. Lo cual es una situación que, de acuerdo con uno de los participantes, es exclusiva de México.

> "Aquí en México hay una categoría que no existe en ningún otro país de registros confidenciales, este archivo que tiene versión pública y que la gente no puede utilizar, llegan a tal extremo que en una galería no te prestan el archivo, pero en la otra sí. Entonces creo que en general los gobiernos se aprovechan cualquier rendija para respetar la ley, pero [a su vez] violarla" (Reportera 2 E2:141)

- La rendición de cuentas es una quimera. En este caso el problema se deriva del desinterés o la conveniencia de los organismos para no proporcionar datos o información.

> "Es un reflejo de nuestro federalismo fallido, en el que se le otorgan privilegios a los gobiernos estatales y la rendición de cuentas este es una quimera, no se les exige la ley, bueno sí se les exige la ley pero en términos tácticos es un albur, te pueden dar toda la información y de buena calidad, como no te pueden dar nada o una copia escaneada a mano" (Reportero 1 E2:145).

- Datos personales. La información también está restringida por contener datos personales, al respecto la Reportera del E1 (18), señala "entonces dijimos si nos la dieran estaría perfecto, pero no nos las iban a dar. Por qué, porque trae foto, trae nombre y son datos personales".

Se puede apreciar que las solicitudes de información representan una serie de retos para los periodistas quienes se encuentran con trabas burocráticas y legales que obstaculizan la obtención de datos e información. Todo ello provoca que se desarrollen estrategias emergentes o que inclusive se cancele la elaboración de la nota.

Se puede apreciar que las solicitudes de información gubernamental no resuelven del todo, las necesidades de información o de datos. Algunos problemas son la dispersión de la información, la falta de consistencia, el mantenimiento de las bases de datos, los datos personales, el desinterés y la falta de complimiento con las leyes de acceso; entre otros problemas. Todo ello provoca que el desarrollo de una nota requiera una mayor inversión de tiempo para obtener lo que se necesita. En el caso del periodismo de datos se observa que el tratamiento de datos personales confidenciales es muy complicado, sobre todo porque en ocasiones son parte de lo que se quiere evidenciar; en este sentido los datos no pueden obtenerse de los organismos pues se trata de información reservada.

Bases de datos sucias

La disposición de conjuntos de datos no siempre significa la oportunidad de trabajar en su análisis de forma inmediata. Hay un proceso de por medio para comprender su estructura y darle uniformidad. El proceso de limpieza es una forma de solventar:

- La mala calidad de los datos.

  "Uno es muy mala calidad en los datos a veces, son muy pobres, hay cientos de miles de dataset por lo regular muy pobres, ya todos muy agregados, están todos como resumiditos, están formateadas desactualizados o no histórico." (Diseñador E2:105)

- La imposibilidad de tratar a los datos de forma evolutiva o histórica dada su estructura.

  "Pero esos reportes trimestrales están de alguna forma estructurados que no te permiten hacer un análisis de la evolución del dinero que se tiene en esos fideicomisos, no sé si me explique." (Editora E1:19)

- La imposibilidad de tratamiento y análisis a causa del formato. Este es uno de los obstáculos más recurrentes en esta etapa. La entrega de datos numéricos en archivos PDF los cuales dificultan la extracción y visibilidad. En el caso de los participantes del Escenario E1 y E2, es claro que se trata de una situación vista como contradictoria, pues sí se cuenta con la información, pero, al ser entregada en PDF, esta debe ser tratada para que pueda ser manipulada y obtener lo que se requiere (Editora E1:21; Analista 3 E2:110).

En los siguientes extractos de entrevistas se exponen dos formas en que se puede resolver esta inaccesibilidad. La primera por medio de un OCR y trabajo humano y la segunda por medio de la existencia de algunos programas que pueden ayudar al respecto.

> "52, 000 hojas en 80 y pico volúmenes, que no eran PDF eran imágenes dentro de un PDF y después de procesarlas en OCR la mitad contenía un número legible, entonces literalmente hay que procesarlo con humanos." (Analista 2 E2: 113)

> "Casi imposible, la verdad si, a mí una vez me mandaron una base de datos en PDF que eran como 700 hojas imagínate, en aquel tiempo yo no sabía absolutamente

> como resolver, pero las cosas que aprendí sobre la marcha, la existencia de algunos programas que te pueden ayudar." (Analista 3 E2: 112)

Después de la conversión del archivo, el siguiente paso es la limpieza de los datos lo cual no siempre tiene buenos resultados, como lo mencionan los entrevistados del E1. Por otro lado, la limpieza de los datos, como se puede apreciar, es vista como una condición previa para el proceso de programación.

> "Por ejemplo el de fideicomiso que sacamos estaba muy sucia la base, muy muy sucia. Esa base pasó como por tres personas además de nosotros que nos ayudaron en cuestión de limpieza y todavía al final era una base muy sucia, muy confusa." (Reportera E1: 21)

Los escenarios E1 y E2 coinciden en que la entrega de información en formato PDF puede resultar un obstáculo, pues no permite procesar la información de forma adecuada, de tal modo que se tiene que trabajar manualmente, esto es capturarla con el apoyo de personas. Asimismo, la Editora del E1 habla de la estructura de las bases que no permiten trabajar los datos de una forma específica acorde con el producto que se pretende desarrollar. Esta percepción negativa es observable también en el equipo E2 que habla de la pobreza, desactualización y mala calidad de los datos.

Incomprensión del dato

Hay dos tipos de incomprensión de los datos que obstaculizan su correcta interpretación.

- La primera en términos intrínsecos de la base de datos, pues por una parte la base puede no tener una estructura bien definida (en el caso narrado por la Editora E1, se habla de la estructura en relación con la variable tiempo) y por otra no se cuenta con un diccionario para interpretar (Diseñador E2).

> "Teníamos campo de dinero en negativos y decíamos esto qué… si lo agarro así nos cambia mucho el análisis, qué quiere decir esos negativos… entonces habla a Hacienda. Que te hagan caso en Hacienda y ya que te hagan caso, qué quieres… no pues sobre esta base en específico porque tienes negativos y ellos decían ¿cuáles campos negativos? Pues tienes campos negativos, qué significan." (Editora E1:25)

> “Que no tienen diccionario para interpretar los datos. Puedes tener una base que puede ser muy buena, pero no encuentras de que se trata y pasa mucho tiempo buscando el diccionario, tienes que asumir la variable que está mal, en términos prácticos una base que no sabes exactamente cuál es la variable y no la puedes interpretar.” (Diseñador E2:104)

- El segundo tipo de incomprensión es en relación con la temática, pues en el caso narrado por la Reportera del E1 el tema que trataba era “cuasi económico” y en este caso la participante asumía no conocer dicha temática y por ello recurría a expertos.

> “Entonces teníamos, además de todo, hablar con muchos expertos, para entender bien que era lo que queríamos. Porque es un tema cuasi económico y yo no manejo el tema de fideicomiso, los temas sociales son mucho más fáciles que los temas económicos.” (Reportera E1: 36)

Como se puede observar la correcta interpretación de los datos asegura la validez de la nota. De esta forma el periodista de datos debe tener una idea clara de cómo están registrados los datos en la base y a qué se refieren sus variables.

Los elementos de un todo: visualizar, programar y editar

Las notas del periodismo de datos pueden ser vistas como un ensamble en el que intervienen las perspectivas de distintos agentes que dotan de significado a un producto de forma lógica. La participación de diseñadores, programadores y visualizadores van conformando un producto que puede ser publicado en formato impreso o visual o ambos. Estas características sugieren un constante replanteamiento de lo que se quiere desarrollar y una interacción constante con otras áreas. En siguiente el ejemplo se muestra la intensa interacción que se da entre los miembros de los equipos. En este caso entre periodistas, programadores y diseñadores.

> "Para que estas infografías después tengan su salida a la versión digital en el diseño web: el infografista le tiene que pasar los archivos al área de diseño web. Esta área me pide a mí los textos y las fotos, las cabezas, los sumarios y cómo vamos a ordenar los mapas. Todo eso yo se los debo de pasar entonces al mismo tiempo. Tengo que ir con la gente de diseño web y decirle: 'ah, bueno, necesita esto, esto, esto y esto'. Esto implica otro proceso que a veces a mí ya no me da la vida. Porque sólo somos [C] y… entonces, o reviso la edición impresa o me dedico a hacer los gráficos interactivos, y entonces a veces no me da tiempo y dejamos de lado la parte digital." (Editora E1:16)

En los apartados siguientes revisaremos los procesos que intervienen en este ensamble que se materializa en una nota y las interacciones con la información.

Balance general

Construir un producto, ensamblarlo y conformarlo como una unidad lógica, conlleva trabajar en distintos procesos de forma no lineal o asincrónica. Implica también una predisposición a la frustración por parte de los participantes, pues se juega con los datos al prueba y error. No sólo para la definición del tema, sino también en el cuestionamiento sobre la mejor forma de representarlos gráficamente, correrlos en una interfaz o darles claridad para un entorno impreso.

En este punto recordemos que composición de una nota implica además la integración de los hallazgos en datos con la “parte humana” (Reportera E1: 25. Es decir, la interpretación de los datos a partir de entrevistas a personas que ayudan “aterrizar“ los datos a la realidad social o bien integra la perspectiva de los involucrados u afectados con miras a exponer el porqué es significativa la nota para determinadas personas.

La idea de lo que debe ser el producto depende de la perspectiva de quien interviene en su elaboración. Se habla de funcionalidad, de claridad, de que sea fácil de recordar, legible, enriquecidos y sobre todo comprensibles.

El tiempo está presente de forma transversal, no sólo se trata de un obstáculo para una persona, sino para un equipo. Se trabaja y planea el alcance de lo que se produce a partir de este factor.

Estructura de categorías y códigos

Estas categorías forman parte de la categoría Valor agregado a la información la cual agrupa a las subcategorías ya descritas: Interacción con los datos y Desarrollo de la nota. En este apartado analizaremos las subcategorías que corresponden a

Visualización (26), Programación (16) y Producción multimedia (6). Como se puede apreciar el proceso con el mayor número de segmentos codificados se ubica en la categoría Visualizar y corresponde al código construir una visualización (ver, Figura 20).

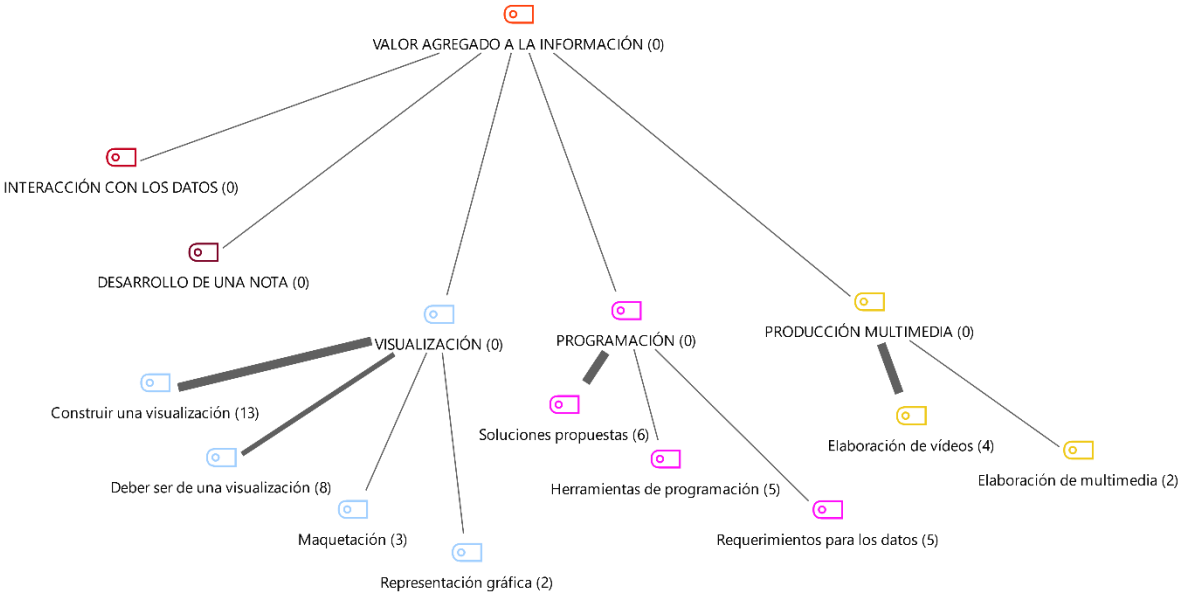


FIGURA 20: VISUALIZACIÓN, PROGRAMACIÓN Y PRODUCCIÓN MULTIMEDIA

Visualización

Al hablar de visualización de los datos, se alude a los requisitos previos para que el personal de visualización comience a “jugar” con éstos: limpieza, consistencia y congruencia estructural de los datos son indispensables para empezar a trabajar con ellos. Se elaboran maquetas o borradores de visualizaciones, en ocasiones se utilizan datos ficticios con el propósito de ilustrar una idea. Al igual que en el trabajo textual, durante este proceso se replantea lo que se quiere desarrollar en función de los datos y de lo que se pretende transmitir. Las visualizaciones están condicionadas a su vez por el público al que están dirigidas; es la audiencia la que determina su complejidad.

A la pregunta de cómo debe ser una visualización, se describe que ésta debe ser directa, elaborada con información validada, clara, que permita interactuar con los datos, memorable, legible, enriquecida, de fácil comprensión y visualmente atractiva.

Maquetación

La maquetación consiste en el esbozo de la representación visual de los datos, este esbozo puede iniciar desde el momento en que se están analizando los datos; como ocurre en el E1, en el que se considera necesario compartirlos con los programadores para que "comiencen a jugar un poco y vean que visualizaciones podemos armar" (Reportera E1: 30). Como es de esperarse estas maquetaciones "sufren muchos cambios" (Reportera E1:38), a partir de la incorporación de los datos, llegando inclusive a descartarse conforme avanza el análisis de los datos.

En los escenarios E2 y E3, el proceso de maquetación involucra la negociación con el editor, los diseñadores y los periodistas; con quienes se analiza y definen las características de la visualización del producto final (Diseñador 2 E3:21). La maquetación supone dos estrategias:

- Realizar simulaciones en las que se verifica la compatibilidad de tecnologías empleadas y se identifica lo que "se puede armar".

> "Hasta visualmente, de repente hay que ocupar este tipo de graficación y estos son los datos, y dices ah pues podría verse así no, pero ya cuando le metes los datos reales, así como css no luce el dato, no era así. Entonces es como tener los datos más reales que puedas para tú poder desarrollar una buena investigación y que digas al final tu producto se va a parecer muy similar a esto, no es así como datos ficticios que se ven bien, pero que no funcionan este tipo de graficación que se está haciendo." (Diseñador 1 E3:44)

- Solicitar apoyo a otras áreas, pues la complejidad de trabajar con una gran cantidad de datos y atender tareas específicas bajo la premura del tiempo lo demanda. Un ejemplo es lo que ocurre en el E1, pues se requería el apoyo de un programador porque se necesitaba un algoritmo para poder visualizar los datos (Editora E1:51).

Dado que la maquetación es una etapa previa a la realización de la graficación de la totalidad de los datos, se observa que puede tener cambios considerables en relación con el producto final. También es necesario contar con una política editorial que permita delimitar o centrar el análisis "en dos o tres aspectos fuertes" (Diseñador 1 E3:50) relacionados con el impacto en la audiencia final, y la facilidad de uso.

Construcción de una visualización

Como ya se ha señalado los requisitos previos de una visualización son los datos limpios y consistentes. Se establece, además, un proceso de negociación con las personas que van a visualizar centrado en la forma en que se armará la visualización.

> "Nosotros les proponemos, mira la base trae esto creemos que se puede hacer esto y esto, que podemos jugar de esta manera y es como un a ver... va mándamela, juego un poco con ella y veo si se puede armar." (Reportera E1:38)

Al igual que un producto textual, la visualización se replantea a partir de la incorporación de los datos. Se puede tener una idea de cómo resolver la parte visual, pero en ocasiones se tiene que descartar o replantear.

> "Yo ya tenía los ejemplos de cómo iban a visualizar los de *gobierno fácil* y pues sí estaban padrísimas sus grafiquitas, pero a la hora de quererlas visualizar ya con los datos de fideicomisos, no se podía, porque no es lo mismo. No cuadraban los datos los números con las gráficas que ellos querían hacer […] y tienes que buscar otra opción." (Editora E1:49)

La Editora del E1 narra la complejidad de armar una visualización a partir de una gran cantidad de datos. Por momentos, como ya se señaló, la Editora toma la postura del usuario final para argüir la necesidad de que la visualización sea sencilla y de fácil comprensión. Asimismo, narra la forma en que tuvo que recurrir a una instancia externa al diario para resolver técnicamente una necesidad de programación, que consistía en visualizar la evolución del dinero de determinados fideicomisos a lo largo de los años (Editora E1:51).

En el caso del Escenario 2, la diferencia operativa en el flujo de trabajo es muy clara pues cada área tiene un dominio específico. El área encargada de Datos define las preguntas que quiere resolver y trabaja en conjunto con el área de Diseño y Programación para darle forma. Como se puede apreciar, el *proceso inicial* se basa en la prueba y el error "Pues es que es muy prueba y error, de que nos juntamos diseño, con datos y empezamos a decir no es que podríamos hacer tal forma de graficar, entonces empezamos a hacer la pruebas" (Diseñador 2 E3:43).

El tiempo es un factor importante en la elaboración de una visualización, "créeme que se puede hacer solo necesitamos un buen tiempo como para poder desarrollar una visualización con tantas características." (Diseñador 1 E3:69)

Tanto en el Escenario E1 como en el E2, el proceso de visualización involucra la negociación entre los diseñadores y quienes trabajan la nota, en ese sentido se habla de realizar pruebas, ver que se puede armar con los datos, ajustar y replantear.

En el E1 se habla de un proceso que comienza con la maquetación para la web, pues se prevé que habrá muchos cambios y por ello importante establecer un primer filtro. Por la estructura del diario y la demanda de este tipo de productos, esta área solicita apoyo a instancias externas cuando el trabajo que se requiere es complejo y requiere de una gran inversión de tiempo y personal.

En cuanto al E2 la visualización se trabaja en conjunto entre las áreas de Datos, Diseño y Programación. Las propuestas se discuten, se prueban se corrigen y ajustan con datos ficticios muy apegados a lo que serán los reales.

Representación gráfica

Trasladar los datos a un lenguaje gráfico requiere de un trabajo de prueba y error para quien lo desarrolla, en un inicio se puede esbozar con datos ficticios. Sin embargo, como menciona el diseñador 1 E3 (44), al momento de ingresar los datos reales el dato que se quiere representar de forma gráfica no funciona. Todo esto requiere de un replanteamiento de la parte visual.

Para el diseñador E2 (165) la representación gráfica de los datos tiene variantes a partir de la plataforma en la que se podrá apreciar sea en una página de

internet o de forma impresa. Por una parte, hay interacción y por otra hay un trabajo previo para que sea suficientemente entendible en un medio fijo como el impreso.

> "Es la posibilidad que te da, hay como una línea entre extremos, el exploratorio, donde el usuario puede explorar esa y el otro es el medio impreso donde ya todo está digerido de una manera entendible" (Diseñador E2:165).

Al igual que la elaboración de la parte textual de la nota, el trabajo visual con los datos requiere de un trabajo iterativo que sugiere un constante replanteamiento de lo que se quiere desarrollar y lo que se tiene como producto en un momento determinado.

Deber ser de una visualización

La visualización consiste en traducir a un lenguaje visual los datos que se han analizado. Este proceso también puede llevar a realizar ensayos de prueba y error en la plataforma en la que se publicará. De acuerdo con los participantes una visualización debe cubrir los siguientes requisitos:

- Que sea fácil de recordar. Una representación debe garantizar que el lector recuerde la "pieza que leyó en algún momento" por su color o sus formas. Esto puede llevar a una competencia sobre cómo contar la historia en donde los diseñadores gráficos prefieren "que su historia se recuerde... por las formas, aunque después se le haya olvidado de que se trata" (Diseñador E2:165)
- De fácil comprensión. Se habla de que los datos estén "aterrizados", "para que se pueda entender muy, muy bien" (Diseñador 2 E2:35) para que sean comprensibles para el usuario final.
- Pensados para una audiencia en específico. La representación gráfica puede estar condicionado por el público al que va dirigido, pensando en que este puede ser especializado, pero al mismo tiempo puede leerlo alguien que no tenga conocimientos del tema.

> “Muchas veces sí, porque el rango de usuarios es más amplio, no tan especializado y hay veces que formas de hacer una visualización para alguien que está más especializado, este te da permiso de hacer un poco más complejas las cosas, porque ellos sabes perfectamente de que estás hablando, no es lo mismo una noticia para público general que para alguien que sabe que está viendo.” (Diseñador 1 E3:31)

En los tres escenarios la representación gráfica debe permitir la comprensión de los datos de forma sencilla y visualmente atractiva. Los diseñadores es el perfil que ofrece mayor detalle de lo que se espera de una visualización. En ese sentido, se habla de que sean diseños claros, legibles, de fácil comprensión y pensados para una audiencia en específico.

Programación

La programación es un proceso que consiste, en este, caso de dotar de la infraestructura lógica el despliegue de la información en forma de texto o visualización para entornos digitales. En una descripción lineal del proceso de desarrollo de una nota de periodismo de datos le correspondería la última etapa de confección de una nota como lo señala el Programador E3:12. Sin embargo, está presente en el análisis de datos, la visualización y la publicación en entornos digitales.

> “A raíz de esa etapa pues ya se dice “ah, mira, la nota es eso, esto es lo que se quiere hacer, lo que se quiere presentar”, entonces se pasa mucho con el lado de diseño de cómo podemos presentar la información. Y la última etapa pues es la de programación” (Programador E2:12)

Requerimientos para los datos

Como se ha señalado en otros apartados, el primer paso para trabajar con datos es asegurar que la base de datos se encuentre limpia. Esto les permite a los programadores realizar pruebas. Es decir, la sistematicidad de los datos y su homologación de atributos permite que sean comparables (Editora E1:25).

Se les da continuidad a las expectativas del producto noticioso. Hay un trabajo previo por parte de analistas y diseñadores del que se valen los programadores para elaborar o poner a prueba propuestas. El área de programación puede ser vista como el área de lo técnicamente posible, en términos de despliegue y funcionalidad del producto en un entorno digital.

> "Y la última etapa pues es la de programación en donde me dicen "oye, queremos hacer todo esto, ¿se puede?" Bueno, nunca preguntan si se puede, nomás "queremos hacer esto". Sí, básicamente nos toca ser muy creativos en nuestro lado, del lado de código, del lado de tecnología, del lado de seleccionar qué frameworks, qué bases de datos, formato, inclusive en qué formato requerimos nosotros la información que la parte de análisis sacó." (Programador E3: 12)

A partir de la solicitud que les hacen los analistas, diseñadores o reporteros el programador propone de forma creativa soluciones en relación con el código, la selección de *frameworks*, bases de datos, formatos, y más para asegurar que el producto final sea accesible para el usuario final.

Soluciones y herramientas para los datos

Las soluciones que se realizan desde el área de programación están muy ligadas con las herramientas que se utilizan, por esta razón se tratan ambos códigos de forma conjunta. El caso narrado por la editora del E1 es ilustrativo en ese sentido pues, ella narra que un trabajo que estaban elaborando de forma conjunta con otra instancia externa no habían logrado obtener los resultados deseados. Hasta que compartió los datos con una tercera instancia la cual le resolvió su requerimiento a partir del uso de una librería denominada D3JS, la cual requiere de un saber especializado y que representaba algo completamente nuevo para los participantes de este escenario.

> "Y en el caso de los fideicomisos cuando ya me presentó la propuesta la gente de Morlan eran gráficas que estaban hechas en una librería que se llama D3 que no cualquiera sabe manejar entonces para nosotras era así como de que padre pues de entrada esto no se parece a nada de lo que hemos hecho, ni a nada de lo que nos han propuesto o lo que hemos visto." (Editora E1:27)

El programador del E3 ofrece más información sobre esta herramienta:

> “El D3JS lo inventaron tres cuates de Stanford, si no me equivoco, y uno de los más importantes se llama Mike Bostock. Es el que tiene una mayor cantidad de D3 y lo que puedes hacer con eso. A raíz de eso pues siempre es jugar y experimentar, ver qué se puede y qué no se puede hacer.” (Programador E3:28)

El punto en común en ambos escenarios es el uso de la librería D3. Sin embargo, el Escenario E1 llega a esta herramienta por medio de la propuesta de la instancia externa que le resuelve su necesidad, no así el caso del E3 que tiene un departamento encargado sólo a la parte de programación.

Producción multimedia

La producción multimedia implica, la generación de gráficos interactivos y vídeos explicativos. Se trata de traducir o darle un nuevo tratamiento al contenido y análisis hecho a los datos, con el fin de añadir elementos que favorezcan la comprensión de la nota. Como es de suponerse, se trata de procesos que requieren de una mayor inversión de tiempo por parte de los participantes y también implica determinado tratamiento de los datos para trabajarlos y darles salida.

Elaboración de multimedia

La salida de gráficos multimedia implica de suyo una inversión mayor de tiempo para el desarrollo de una nota, pues es visto como una forma de darle una salida atractiva a un producto noticioso sobe un tema complejo. Sin embargo, “Implica cierto trabajo que no es de un día para otro, aparte lo que conlleva estar armando lo que te digo: los textos, ver cuáles van a ser los tipográficos, la visualización.” (Editora E1:16)

Para el E1 este añadido al ensamble que da forma a la nota significa también un elemento que puede retrasar la salida de una nota, pues requiere del trabajo adicional de un diseñador web que trabaja la parte digital o multimedia, pero este diseñador trabaja con distintas secciones del E1 de forma simultánea lo que resta

“capacidad de reacción” para publicar una nota sobre algún tema de coyuntura (Editora E1:57).

Elaboración de vídeos

La realización de elementos adicionales que añaden valor a la nota mostrando explicaciones sobre temas complejos, se ve condicionada, nuevamente, por el tiempo, pues “hay veces de que por tiempos no hemos podido generar ciertos contenidos” (Analista 2 E2: 69).

Se tiene claro en los escenarios E1 y E2 que la falta de personal especializado que trabaje de forma exclusiva con esta sección limita el alcance de una nota, una vez que “hay veces que la colaboración da como resultado videos de más corto alcance de lo que nosotros queríamos, no, como que, si teníamos unas ideas como más amplias, pero luego los tiempos como que dictan otra cosa” (idem).

Buena parte del trabajo que se realiza en estos escenarios depende de redes informales de apoyo, que dan pie a casos como el que narra la editora E1 (53) “en el área de vídeo tienen a alguien que sabe manejar ilustrador, entonces ellos hacen eso.” Se percibe que el trabajo de crear contenido multimedia se ve limitado por el tiempo y, por ello, los productos deseados por los editores se quedan a medias.

Una nota que entiendan todos: la idea del lector

Balance general

La concepción que se tiene del lector entre los participantes define en buena medida el nivel de complejidad de una nota, su discurso y su visualización. En distintas participaciones se reitera la importancia de trabajar notas que sean comprendidas por todo el público para darles más posibilidades de que se socialicen y de que tengan una mayor resonancia. Por otro lado, se perciben cambios en los hábitos de consumo de información de los lectores, reconociendo que consultan distintas fuentes de información y que ya no son leales a un medio.

Estructura de categorías y códigos

La categoría audiencia se refiere al destinatario o lector de los productos noticiosos que se elaboran. Incluye también la concepción que tienen los participantes de los lectores y los ajustes que hacen a sus productos a partir de dicha concepción. Los

códigos que dan forma a esta categoría y su frecuencia son: Conceptualización del lector (15) y Producto de acuerdo con audiencia (6) (ver, Figura 21).

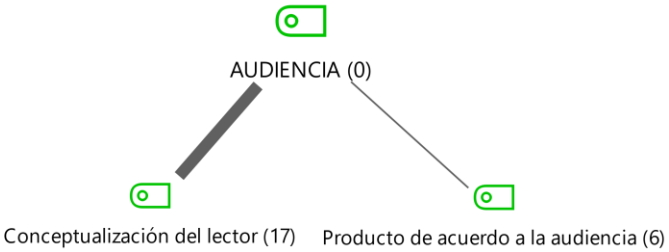


FIGURA 21: AUDIENCIA, FRECUENCIA DE CÓDIGOS

Conceptualización del lector

La idea del lector está vinculada con la concepción de un producto de periodismo de datos, pues durante el desarrollo de la nota la figura del lector conduce buena parte de su contenido y su visualización. Se habla de tratar un tema de forma que sea comprensible para todo público, de evidenciar irregularidades sin la necesidad de que el lector cuente con presupuestos especializados.

> "Algo que pueda entender tu público, si tú no puedes lograr ese paso de hacer lo que estaba poco entendible, oculto o que necesite estudiar economía, maestría o estudiar Ingeniería, si al final tú no puedes transformar esa información a algo que se entienda no estamos en ese proceso." (Analista 3 E2:155)

Se trata también de brindarle elementos al lector, de conducirlo para que se forme una opinión y tenga mayores elementos para comprender una nota. Cada parte de la nota se trabaja bajo este supuesto, el propósito es atraer la atención y "aterrizar" temas complejos para que sean fáciles de comprender "tiene que ser una combinación de factores que haga que al lector le guste, no sé es que no hay una regla mágica." (Equipo E2:146). A continuación, se caracterizan las formas en que el lector es visto:

- El lector y la comprensión de la nota: un requisito de las notas que se producen es que sean comprensibles para un amplio rango de lectores esto

facilita la socialización de un tema, su resonancia social y su reproducción en otros medios.

> "Porque no es fácil entender el tema, entonces no es fácil entender el tema te enfrentas a que puede ser una investigación buenísima, pero si no lo entiende la señora que vende los tamales en la esquina ya te fregaste porque no se va a socializar y no se va a lograr difundir la problemática, ni le va a llegar a la gente y creo que al menos en una parte muy pequeñita creo que sí conseguimos que la gente por lo menos volteara a ver el tema que eso es algo muy difícil no." (Editora E1:45)

- El lector como referente: el trabajo frente a las propuestas de visualización o del diseño del sitio, el lector es una figura constante no sólo para la comprensión de un tema, sino también en cuanto a centrar el diseño de una visualización en relación con la facilidad de uso. El público o la audiencia para la cual se piensa un producto puede determinar el nivel de complejidad del producto, según uno de los participantes del Escenario 2. Si el usuario tiene mayores presupuestos para la comprensión de una nota, eso te permite "hacer un poco más complejas las cosas" (Diseñador 1 E3:31). De la misma forma la concepción que se tiene del lector y el tema a tratar incide en la forma de desarrollar una visualización. (Diseñador 1 E3:29).

> "Entonces yo les dije, no, no puedo, la gente no va a entender eso. No tienen por qué picarle 600 veces para enterarse de los fideicomisos que hay, y en el buscador nunca pudieron resolver el asunto. […] Aquí, por ejemplo, cómo explicarle a la gente cómo utilizar el sitio, pues porque eso luego la gente no lo entiende, no. Y pues eso básicamente." (Editora E1:51; 53)

- El lector como consumidor: se parte de la idea de que el lector consulta distintas fuentes de información, y que no es leal a un solo medio.

> "Es que ya no hay lealtad a los medios en este sentido, la gente ya no entra al medio para informarse, llega al medio y se va a otro medio, y así, supón que la gente regresa a leer, pero ya no es el medio." (Analista 2 E2:154)

- El lector como destinatario crítico: se atribuye al periodismo la función social de informar y empoderar al lector.

> “Al fin y al cabo lo que intentar es empoderar a la gente o ayudar a la gente y el periodismo tiene una función social también interesante esa combinación que tiene, por eso digo la parte exploratoria para mí es como muy interesante, la gente también es una herramienta para ellos, no es un dato digerido, es una ramita para que ellos mismos generen sus historias, se informen y de cierta manera actuar.” (Analista 2 E2:180)

Producto de acuerdo con la audiencia

Se escribe, se desarrolla o se diseña pensando en ajustar u orientar el contenido en un tipo de lector. El discurso de lo que se estructura debe ser apropiado para el usuario final. La adecuación del producto se realiza con base en la idea de un lector.

> “Y a aterrizar como este tipo de lenguaje, para que se pueda entender muy muy bien. A veces como se utilizan muchos términos con datos, hay como ese argot, de palabrillas y todo eso, hay que aterrizarlo como al usuario final que dice esta palabra no, para mí no dice nada, entonces a que te parece si utilizamos un filtro en lugar de otro termino, es trabajar con este tipo de información y aterrizarla.” (Diseñador 2 E3:35)

El producto se “aterriza” desde distintos ángulos, pero con el mismo propósito: hacer comprensible algo complejo. Una estrategia que se describe, para que una nota no sea percibida como pesada por el lector es la inclusión de entrevistas con expertos o historias que resalten la parte humana en una nota para que la gente la lea (Reportera E1:27).

Los formatos de salida

Balance general

El formato de salida de una nota representa una dificultad porque cada entorno de publicación sea digital o impreso, requiere de un tratamiento distinto e implica que la unidad de periodismo de datos y el personal de apoyo inviertan una gran cantidad de tiempo, que, como ya se mencionó, no siempre se tiene, lo cual hace imposible la publicación de la nota en uno o ambos entornos.

Estructura de categorías y códigos

La categoría de entornos de publicación**:** contiene los códigos asociados con los soportes de salida y las tareas que implican. En este caso se limita a los productos noticiosos diseñados para las versiones en: Impreso (5), Multiplataforma (3) y Digital (2) (ver, Figura 22). Esto incluye las referencias relacionadas con los procesos, limitaciones o tareas que supone cada plataforma.

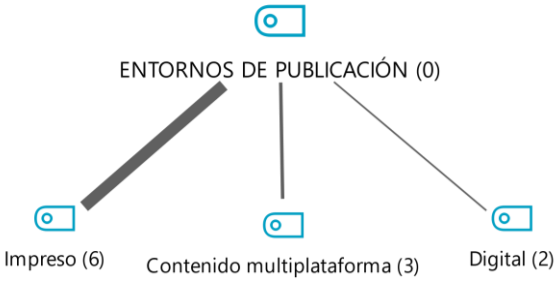


FIGURA 22: ENTORNOS DE PUBLICACIÓN

Impreso

La perdurabilidad de una nota es más probable en entornos digitales, sin embargo, la exposición o visibilidad es mayor en la versión impresa. De acuerdo con la Editora del E1 (Editora E1:11) al publicar una nota en soporte impreso se corre el riesgo de que se pierda, mientras que si se publica en distintos formatos o en formato interactivo hay más posibilidades de que un trabajo perdure.

> "En cuanto al tema de reconocimiento tanto interno como externo, lectoras y lectores sigue siendo la versión impresa. El referente es lo impreso. Nos hemos dado cuenta y prácticamente eso es muy obvio. Cuando nosotros aparecemos en la portada del impreso la nota se catapulta 100 veces más en cuanto a lectura que cuando se queda en interiores; o cuando ni siquiera aparece en el portal, digo en el impreso." (Editora E1:15)

Subyace la idea de que hay una exigencia institucional de darle mayor prioridad a la publicación de notas en la versión impresa, lo que es contrario a la percepción de la Editora, quien señala que ésta no debería ser un referente (Editora, E1:16).

Por su parte en el Escenario E2 se habla de que el formato impreso supone las limitaciones intrínsecas de la versión impresa en términos de espacio (Analista 2 E2:63).

Digital

Tanto en el Escenario E1 como en el E2 se tiene la expectativa interna de que se publiquen notas en formato impreso y digital, sin embargo, el desarrollo de contenido digital requiere de distintos esfuerzos por parte de infografistas, diseñadores web, desarrolladores de interactividad, entre otros.

> "Y que nosotros no trabajamos en una sola cosa, en un solo patín como el entorno digital, sino también la versión impresa." (Editora E1:2)

En todo caso el proceso de periodismo de datos incluye, en varios casos la versión digital de la nota.

Contenido multiplataforma

La forma en que se publicará una nota determina en buena medida el tiempo de inversión que se le dará y el tratamiento del contenido. El E1 presenta la dificultad de publicar en formato impreso y digital las notas que produce, esto implica un mayor nivel de planeación. Se afirma que ambos formatos tienen la misma importancia, sin embargo, en ocasiones no es posible publicar a la par en impreso y digital por falta de personal.

> "No quiero decir que sea obligatorio, pero, o sea, nuestro trabajo es publicar, o sea, que se visibilice en el medio impreso el trabajo que hacemos para la web. Tiene la misma importancia, o incluso mayor, porque muchas veces hemos publicado textos que salen en la versión impresa pero no tenemos como la posibilidad para salir en web con un interactivo, que nos gustaría, pero no tenemos como la forma de cómo lograr hacer eso tan, tan a corto plazo (Editora E1:7).

Este mismo problema se presenta en términos de limitaciones de tiempo y personal en el E2 "hay veces de que por tiempos no hemos podido generar ciertos contenidos porque es como multiplataforma […] hay que hacer el impreso, hay que hacer el digital y ya el video quizás ya no alcanza el tiempo." (Analista 2 E2:69)

PARTE IV CONSECUENCIAS (RESULTADO DE LA ESTRATEGIA)

El periodismo de datos visto desde su práctica

En este apartado se analizan los códigos que se asocian con las implicaciones de desarrollar este tipo de periodismo como los plazos, la percepción de otras áreas, los requerimientos internos. Integra también las menciones que hacen los participantes a otros medios que realizan este tipo de periodismo y que son vistos como referentes al momento de trabajar una nota. Conjunta, además, aquellos fragmentos que hablan de la conceptualización de lo que es o debería ser este tipo de periodismo

Balance general

Este tipo de periodismo es percibido por los participantes como una vertiente que integra tanto la perspectiva matemática como el uso de herramientas especializadas para el manejo y la visualización de datos. No se trata de un periodismo que excluya al diarismo (periodismo tradicional), sino de una confluencia entre el periodista de datos, los datos y la información orientada a comprender la realidad. Algunos de los diarios que son considerados como principales referentes del periodismo de datos son *The New York Times, The Guardian, Washington Post*, entre otros.

De acuerdo con los participantes, el periodismo de datos debe trabajar en el desarrollo de productos intuitivos que, de forma sencilla, evidencien nuevos datos a partir del cruce de información; que cuenten una historia de forma atractiva, sencilla

clara y contribuyan a elevar el discurso para cumplir con la finalidad de empoderar a los lectores.

Estructura de categorías y códigos

Esta categoría conjunta los fragmentos que tratan del origen de las unidades de datos, el concepto y el deber ser del periodismo de datos; así como los códigos de las acciones implicadas en el desarrollo de este tipo de periodismo, como los plazos, la percepción de otras áreas y los requerimientos internos. Integra también las menciones que hacen los participantes sobre otros medios que realizan periodismo de datos y que son vistos como referentes en el momento de trabajar una nota a corto y largo plazo, la apertura a este tipo de periodismo y sus críticas. Los códigos de que se compone esta categoría son: Concepción del periodismo de datos, Deber ser del periodismo de datos, Producto intuitivo, Notas a corto y a largo plazo, Implicaciones; Referentes del periodismo de datos; Apertura al periodismo de datos y Críticas al periodismo de datos (ver, Figura 23).

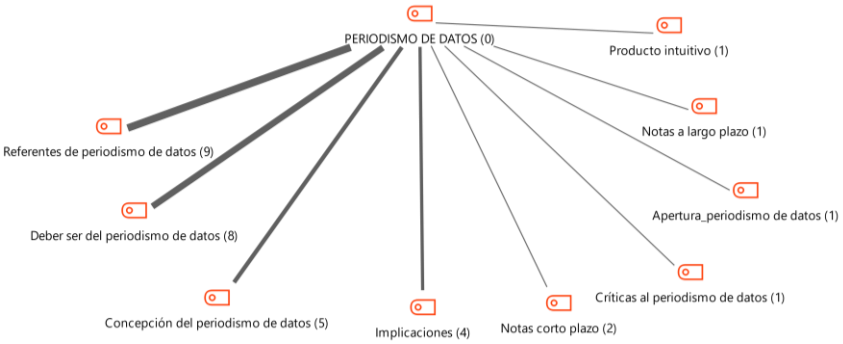


FIGURA 23: PERIODISMO DE DATOS, ESTRUCTURA Y FRECUENCIA DE CÓDIGOS

En la Figura 23 se aprecia que el código que se presenta con mayor frecuencia se refiere a las fuentes de información que son vistas como Referentes del periodismo de datos (9), así como la posición de los participantes sobre el Deber ser del periodismo de datos (8) y la Concepción de lo que es este tipo de periodismo (5).

Concepción del periodismo de datos

El periodismo de datos es visto como un área de oportunidad y de colaboración con especialistas para crear notas periodísticas basadas en datos con visualizaciones interactivas y rigor metodológico (Programador E3: 4). Este tipo de periodismo implica el uso de computadoras, herramientas y programas especializados para el tratamiento de datos.

> "Es periodismo, pero ahora tienes una herramienta base y tu herramienta son datos, tu nueva herramienta es Excel y las matemáticas. Eso te da más sustento cuando vas y hablas con la gente" (Reportera E1:29)

La nota o el producto derivado del periodismo de datos incluye texto, gráficas, la posibilidad de visualizar datos, la definición de una maqueta web y el diseño de la interacción con la información.

> "O sea, a lo que voy es a esa parte, periodismo de dato no es que sólo escribas. Además de eso debemos de pensar en cómo lo vamos a visualizar definir una maqueta para web, definir cómo nos vamos a mover. Definir si nos vamos con hojas, con gráficos. Si usamos scroll no usamos scroll. Todo es lo definimos con el equipo que se conforma para ese proyecto." (Reportera E1:36)

No se deja de hacer periodismo o de contar historias que integren el lado humano, sino de contar historias integrando números, ofreciendo así un enfoque diferente (Reportera E1: 27). En este sentido el diarismo y el periodismo de datos son vistas como opciones que se complementan para la comprensión de la realidad (Equipo E2: 178). La diferencia entre ambas es que el periodismo de datos proporciona al reportero herramientas para enfrentar la realidad tecnológica (Analista 1 E2: 175).

Referentes del periodismo de datos

La mayor parte de los actores, independientemente del rol que juegan, consultan otros medios de información. Los periódicos o medios que representan la competencia a nivel nacional y el seguimiento a diarios internacionales como el *New York Times* o el diario *The Guardian*, son vistos como medios que se consultan con frecuencia para conocer la forma en que resuelven problemas de visualización, de tratamiento de datos o relacionados con la forma de contar una historia. Se consultan también con el propósito de sondear los temas que cubren y el tratamiento que le dan a sus productos. Aquí dos ejemplos:

> *"The Guardian*, si hacen cosas muy interesantes y sobre todo ver como resuelven cosas que a lo mejor parecen como tan complejas, y hacen cosas con tal sencillez, que dices ok lo resolvieron muy bien." (Diseñador 1 E3:84)
>
> "Pues yo sí *The New York Times*, *The Guardian*, cosas, así como muy padres, tienen como, sí es como un referente para ver que está haciéndose." (Diseñador 1 E3: 101)

En el caso del E2, además de mencionar algunos de los diarios citados, hablan de consultar a la competencia y a diarios latinoamericanos de Argentina, Costa Rica, Colombia.

> "La Nación de Argentina, *La Nación* de Costa Rica, *El tiempo* de Colombia, latinoamericanos esto está como más... *Ojo público*, *La silla vacía* y obviamente los otros medios *The Guardian*, *The Economist*, *Angeles Times, Herald Tribune, Texas Tribune*". (Analista 1 E2: 100)

Implicaciones

De acuerdo con lo anterior la elaboración de productos de periodismo de datos implica una gran inversión de tiempo no sólo de quien la elabora, sino también de quienes colaboran con otro tipo de insumos como las visualizaciones, la página web, el procesamiento de datos, entre otros.

> "Implica cierto trabajo que no es de un día para otro y aparte lo que conlleva estar armando lo que te digo: los textos, ver cuáles van a ser los tipográficos, la

> visualización porque además muchos de estos gráficos que salen en digital no todos los aporta la gente de infografía. Porque no todo se puede resolver en un mapa interactivo." (Editora E1:16)

Las implicaciones que conlleva el periodismo de datos obedecen a un sentido de brindar mayor profundidad a un producto, lo cual requiere de una constante negociación con otras áreas de un diario en un entorno colaborativo. El tiempo de inversión de una nota y los propósitos de su discurso son quizá la principal distinción con el diarismo, pues en este caso el periodismo tradicional reacciona de forma inmediata con fines de informar y no tanto de analizar. La relación con el diarismo tradicional es ambivalente, pues es visto como una práctica que se basa en notas de corto plazo, que no requieren de una investigación profunda, pero se asume que el diarismo y el periodismo de datos conviven y te sitúan frente a una realidad.

> "Conviven, es una forma distinta de contar historias y el insumo principal que tenemos para hacer nuestro trabajo es el diarismo, sin el diarismo pues dónde estás parado, cómo sabes dónde está la realidad." (Reportero 1 E2:178)

Notas a corto y a largo plazo

Como ya hemos visto, las expectativas del diario es que se produzcan notas a corto plazo, de forma rápida y con mucha resonancia mediática. Sin embargo, en el caso del siguiente participante hay un desmarque del periodista tradicional y de la lógica de la producción inmediata

> "Es publicación, corto plazo, longitudinaria, este, rapidez, que sea un gran putazo periodístico y no es así, entonces es un círculo vicioso porque nosotros no les vamos a ofrecer los que ellos esperan de un periodista tradicional, entonces que eso que te encasilla, que nosotros tengamos que publicar cosas más pequeñas con menor profundidad más hechas al calor." (Reportero 1 E2:83)

Por su lado las notas a largo plazo requieren de periodos prolongados de tiempo para el desarrollo de varios procesos, como lo son las solicitudes de información, la búsqueda de bases de datos, la limpieza y la visualización, todo integrado en una planeación.

> “Ahorita todavía tenemos muy buen tiempo para empezar a hacer solicitudes de información, para buscar bases de datos, para analizarlas, si hay que limpiarlas, para ver con quién trabajamos para visualizar los datos, o sea, como ya hacer una planeación más aterrizada.” (Editora E1:21)

Subyace la distinción de cómo se invierte el tiempo y qué tipo de producto se elabora. Esto define en buena medida la forma en que se realiza periodismo: desde lo tradicional o bien desde la investigación.

El deber ser del periodismo de datos

El principal requisito que debe reunir una nota de periodismo de datos, desde la ortodoxia de la práctica, es la capacidad de convertir datos complejos y sistematizarlos en un producto entendible para un amplio espectro de lectores (Reportero 1 E2: 156). Se habla de contar una historia de forma atractiva y elevar el discurso entre los lectores; así como de mostrar o evidenciar datos; o de crear nuevos datos e interpretaciones a partir de cruzar información.

> “Una buena nota de periodismo de datos es que al final esos datos tú puedes convertir que pueden ser complejos en algo completamente entendible y asimilable.” (Analista 3 E2:156)

> “Sinceramente yo creo que te da la posibilidad de llegar a revelaciones que te comentaba hace rato, que no son comunes, puedes llegar a revelaciones que no pensarías que están ahí y eso es rico profesionalmente.” (Analista 1 E2: 174)

También forma parte de este deber ser, el generar productos significativos para los lectores, a partir del manejo de grandes volúmenes de datos, pues se les brinda la posibilidad de aterrizar esta información a una realidad concreta.

> “A mí una cosa que me gusta que te da el periodismo de base de datos es de generar este tipo como de historias en las cuales la persona que las lee crea su propia historia, porque cuando trabajas con datos muy grande eh... no sé por ejemplo hablando del agua y que la gente pueda bajar esas historias donde ellos viven, la historia para mí que vivo en Iztapalapa es diferente para quien vive en la Condesa.” (Analista 2 E2:159)

El periodismo de datos, al hacer asimilable la información para el público amplio facilita la comprensión de temas complejos y al darles esa claridad, también los empodera para la toma de decisiones.

Producto intuitivo

La funcionalidad de una nota interactiva depende de qué tan intuitiva es, en el sentido de que no requiera instrucciones de uso y sea fácil de navegar o explorar.

> "Funcionalidad de repente sabes que esto no parece botón hay que hacerlo como que más obvio, cositas así, es como trabajar mucho para que el usuario lo tenga fácil, inclusive a veces no necesite instrucciones, aunque hay empresas donde se trabaja te exigen que tengan instrucciones." (Diseñador 1 E3:44)

Apertura al periodismo de datos

El tipo de productos que desarrolla el equipo del E2 comienza a abrirles espacios en la versión impresa del diario. Se habla de publicar en doble página o doble plano, lo cual es visto como un espacio de reconocimiento para darle notoriedad a un producto noticioso.

> "Recientemente hicimos una en doble plano, entonces los espacios o la extensión que nos permiten utilizar empieza a abrirse por tanto en el medio impreso como en una sola plana, como bien mencionaba, tenemos la oportunidad de presentar notas que puedan ocupar la doble plana, y de manera similar hay espacios que se empiezan a abrir como ya lo dije." (Diseñador E2: 64)

Críticas al periodismo de datos

De acuerdo con uno de los participantes la visión que tienen otros medios de lo que se desarrolla en el E2 no se asemeja al periodismo tradicional, pues no integra el trabajo de campo.

“Usualmente nos critican… porque no tiene trabajo de campo. Nos critican de amigos. De que nos falta trabajo de campo. Si fuéramos periodistas. Sí es algo que estamos bien cojos. […] yo he hecho unas siete u ocho entrevistas, cuando mucho. Depende el tipo de trabajo.” (Editor y Analistas E3:45)

PROCESOS COGNITIVOS

Para asociar los procesos cognitivos con los procesos de valor agregado es necesario retomar los flujos y los usos de los datos y de la información en el contexto de las redacciones, las cuales proveen de las situaciones, tareas, responsabilidades y las condiciones intervinientes, descritas en la Parte III, que se deben sortear para la elaboración de una nota de periodismo de datos.

Por ello, cabe preguntarse si los procesos cognitivos tienen relación con las estructuras organizacionales de las redacciones; con el perfil y funciones de cada uno de sus integrantes; así como son sus competencias y experiencia para solucionar problemas desde sus propias trincheras profesionales.

En el Cuadro 4 y en la Parte I de este documento se presentaron las características profesionales, los roles y las formas de trabajo de los actores que trabajan en redacciones del periodismo de datos. En la Figura 24 se muestran al conjunto de actores que participaron en este estudio y que dan forma al producto noticioso de periodismo de datos.

FIGURA 24: PROFESIONALES QUE AGREGAN VALOR A LOS DATOS

Como actores externos se tienen a los expertos que ayudan a validar las notas, a las personas que operan como enlaces en las diferentes dependencias donde se generan datos, y a la audiencia, entre otros. Hay una multiplicidad de actores cuya actividad depende de la relación que establecen con los datos, pues se pueden ceñir al acopio, a la sistematización, a la limpieza, a la visualización e incluso a la creación de datos. La colaboración y el trabajo en equipo son características intrínsecas a estas redacciones, en las que se comparten, valores, exigencias, disposiciones, espacios y retos; por ello el proceso de elaboración de una nota, es en esencia un proceso en el que convergen diferentes profesionales que realizan procesos intelectuales interrelacionados y desarrollan acciones de co-construcción.

Acorde con lo anterior la identificación de los procesos cognitivos en el estudio ha sido un proceso complejo debido a la dinámica y las diferentes especialistas que intervienen en la construcción de la nota de periodismo de datos. La aproximación a estos datos tiene su base en la estructura de categorías y códigos creados previamente y constó de dos etapas:

1) Se identificaron y codificaron los segmentos que hacían referencia a verbos relacionados con actividades intelectuales; los cuales se relacionaban con conceptos y pensamientos que no son directamente observables y que se infieren a partir de: a) del análisis del lenguaje expresado en los 792 segmentos de entrevistas y b) el contexto que les da significado. Se agruparon de forma lógica estos procesos considerando la taxonomía de Bloom (1990) para establecer una primera clasificación que permitiera sistematizar la información. Las categorías utilizadas fueron: conocimiento, comprensión, aplicación, análisis, síntesis y evaluación.
2) Al compartir estos resultados, se recibieron observaciones de que no todos los verbos que se identificaron debían corresponder con los propuestos en la taxonomía de Bloom, por lo que era necesario crear una categoría por separado en la que se rescatarán los matices y

particularidades de los periodistas de datos, tal como se esperaría de un estudio que utilice la TF. A esta categoría se le denomino “In vivo” (ver, Figura 26) e incluye los verbos a los que los participantes hacen referencia de forma implícita o explícita, sin restringirse a los procesos caracterizados por Bloom. Para su ubicación se tuvo que hacer una nueva lectura completa de las transcripciones de los tres escenarios.

Como se puede apreciar las interrelaciones establecidas entre las condiciones causales, las estrategias de acción, las condiciones intervinientes, las consecuencias y el fenómeno mismo de este estudio, están interrelacionadas con procesos cognitivos de los distintos actores que conforman el ensamble para la creación de una nota de periodismo de datos (ver, Figura 25). En este estudio la caracterización de estos procesos cognitivos se basa en las entrevistas aplicadas y en ese sentido retoma “el decir” de los participantes. En ese sentido la fotografía puede no ser tan fiel por su origen, pues se trata de reconstrucciones verbales, sin embargo, representa una forma de aproximarse a estos procesos con evidencia empírica.

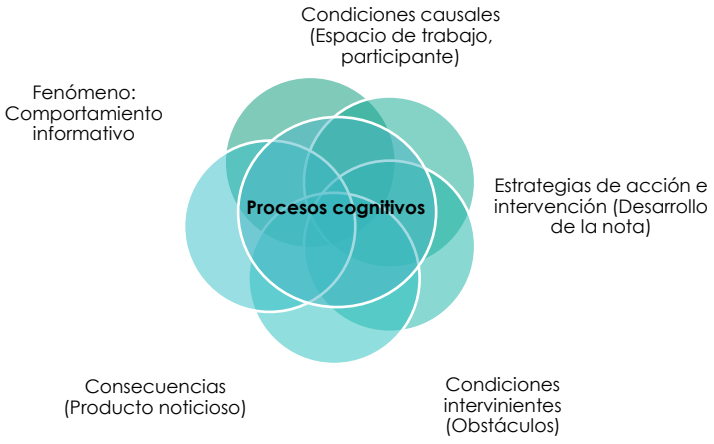


FIGURA 25: PROCESOS COGNITIVOS Y SU RELACIÓN EN TORNO A LAS RELACIONES LÓGICAS AXIALES

Balance general

Los procesos cognitivos se insertan de manera central en la base del proceso de valor agregado principalmente en lo que se refiere a la búsqueda, la explicación y la resolución de problemas. No hay un patrón único para los profesionales que se dedican a co-construir una nota de periodismo de datos. Acorde con ello solo hay dos procesos que se repiten entre todos profesionales, que son buscar y visualizar.

Se observa que los procesos se llevan a cabo de manera colaborativa lo que permite identificar las interacciones entre los profesionales y los datos para crear un constructo que tenga sentido para crear un producto con narraciones comprensibles para una audiencia.

Al no haber un patrón único, es necesario reconocer que cada profesional involucrado lleva a cabo procesos cognitivos que son parte del rol que está desempeñando. Es decir, si es el Editor desarrolla procesos de planeación organización o estructuración; mientras que los reporteros se enfocan en el análisis y en contar historias.

Estos procesos a su vez se vinculan con el tratamiento de datos, lo cual se lleva a cabo en periodos limitados de tiempo. Se observa que la síntesis no es un proceso que sea fácil de identificar, e inclusive de expresar.

La articulación de las actividades siendo tan diferentes, recaen en el editor quien en última instancia opera como se dijo antes, como un organizador y planeador.

En el caso de la categoría *Aplicación,* la mayor parte de los verbos empleados se ciñen a resolver, encontrar, explicar y construir. Esto supone poner en juego procesos de abstracción en situaciones específicas de resolución de problemas. Por lo que corresponde a la categoría de *Conocimiento,* la mayor frecuencia de verbos segmentados corresponde al código buscar, seguido de definir y mostrar, lo cual habla de trabajar bajo fórmulas establecidas, emplear términos incorporados, comunicar principios o generalizaciones.

En cuanto a la categoría de *Comprensión*, los códigos explicar, revisar y comprobar reúnen el mayor número de segmentos codificados de este apartado. Se

percibe un mayor nivel de pensamiento abstracto en esta categoría, pues implica procesos como reformular, parafrasear y verificar la información. Esto es congruente con el objetivo expresado de desarrollar notas comprensibles para todo público, como se verá más adelante.

La categoría de *Evaluación* enmarca la repetición del código revisar entre los más frecuentes. Esto parece natural, pues se trata de un proceso paralelo que implica comprender y evaluar la información que se va obteniendo.

Estructura de categorías y códigos

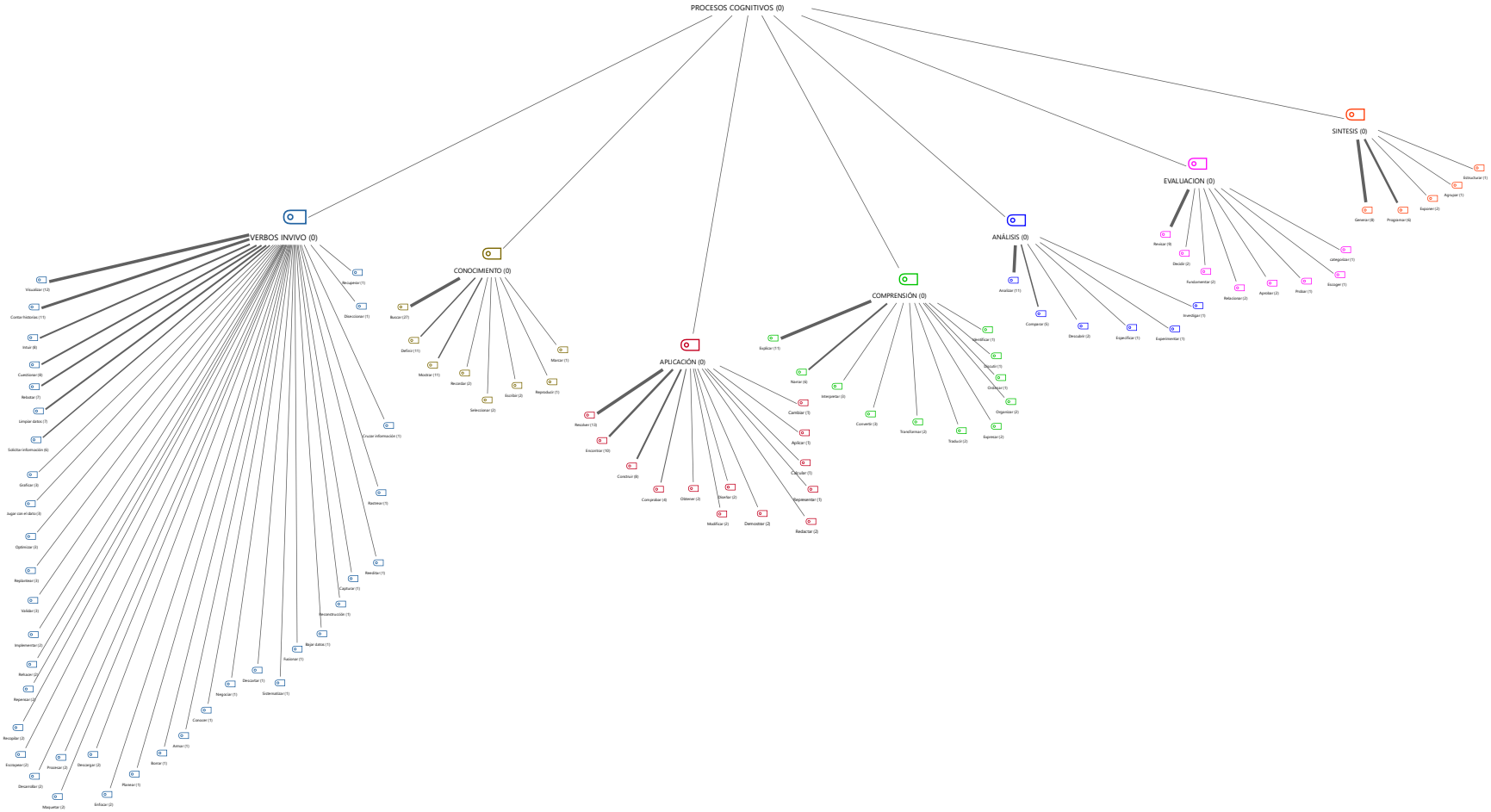


FIGURA 26: FRECUENCIA DE VERBOS SEGMENTADOS POR PROCESO COGNITIVO

La Figura 26 permite observar la estructura total de códigos derivados/utilizados de/para el análisis de las entrevistas, se presentan de izquierda a derecha de acuerdo a la frecuencia de segmentos codificados acumulados por categoría. Se trata de una estructura que consta de 7 categorías y 89 códigos distribuidos de la siguiente forma: In Vivo (38 códigos); Aplicación (13 códigos); Comprensión 11 (códigos); Evaluación (8 códigos); Conocimiento (8 códigos); Análisis (6 códigos) y Síntesis (5 códigos). En este caso es notoria la cantidad de códigos naturales de la categoría “In vivo”, lo que permite revalorar el hecho de trabajar con una perspectiva que pone en primer plano a los datos derivados del propio análisis, en lugar de trabajar con elementos a priori.

La Figura 27 muestra la totalidad de las categorías y códigos que se identificaron en la revisión de las entrevistas de los tres escenarios y en el reporte final de una nota elaborada por el equipo del E1. Como se observa la categoría con la zona más amplia corresponde a la categoría *In vivo* (110 segmentos); seguido de *Conocimiento* (57 segmentos), *Aplicación* (49 segmentos); *Comprensión* (34 segmentos); *Análisis* (21); *Evaluación* (20) y *Síntesis* (18). Una lectura global del gráfico permite observar que el proceso cognitivo con mayor presencia es el de *Buscar* (27 segmentos), seguido de *Resolver* (13 segmentos), *Visualizar* (12), *Analizar* (11), *Explicar* (11) y *Contar historias* (11). Estos procesos denotan una disposición imbricada y manifiesta a lo largo del estudio, se trata del interés de contar historias de una forma sencilla para explicar temas complejos por medio de la búsqueda, análisis y visualización de datos. En estos procesos es constante que los participantes encuentren obstáculos como los ya descritos, mismos que tienen que resolver para dar forma a una nota de periodismo de datos.

A continuación, se analiza cada una de las categorías que agrupa a los códigos relacionados con procesos cognitivos y se ofrecen ejemplos que las ilustren. Se tratan también las diferencias sustantivas, cuando las hay, entre los contextos y la prevalencia de procesos por tipo de participante.

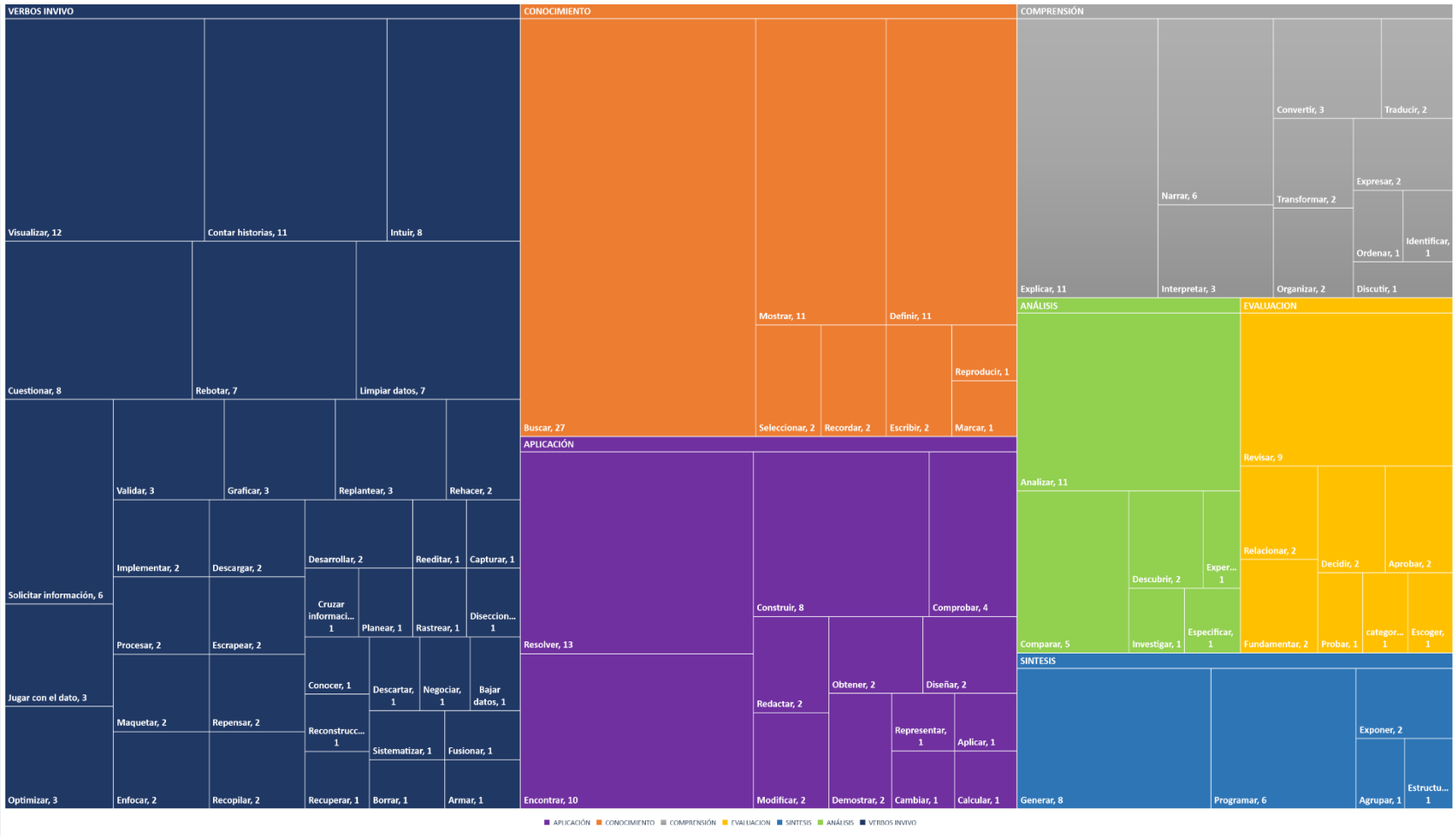


FIGURA 27: FRECUENCIA DE CÓDIGOS RELACIONADOS CON PROCESOS COGNITIVOS Y SU AGRUPACIÓN EN CATEGORÍAS

Guía de colores: Azul obscuro = In vivo; Morado = Aplicación; Naranja = Conocimiento; Gris = Comprensión; Amarillo = Evaluación; Azul claro = Síntesis; Verde = Análisis.

Para apreciar la estructura total de categorías y códigos y su interrelación con los procesos cognitivos el lector puede remitirse al Anexo 2. Para la elaboración de este esquema se consideraron las intersecciones entre cada código; es decir, los segmentos de las entrevistas compartidos, y se excluyeron los códigos que no presentaron ninguna intersección.

In vivo

Esta categoría consta de 40 códigos, la distribución por contexto y por profesionales que los mencionan se muestra en el Cuadro 8.

CUADRO 8: PRESENCIA DE CÓDIGOS DE PROCESOS COGNITIVOS INVIVO EN ESCENARIOS Y POR TIPO DE PARTICIPANTES

		Escenarios en donde se presenta							
	Código	E1	E2	E3	Editor	Analista	Reportero	Diseñador	Programador
1	Visualizar (12)	■	■	■	■	■	■	■	■
2	Contar historias (11)	■	■	■		■	■	■	
3	Intuir (8)	■	■		■	■	■		
4	Cuestionar (8)	■		■		■	■		
5	Rebotar (7)			■		■		■	■
6	Limpiar datos (7)	■		■	■		■		■
7	Solicitar información (6)	■	■	■	■	■	■		■
8	Jugar con el dato (3)	■		■			■		■
9	Validar (3)	■		■			■		■
10	Optimizar (3)			■				■	
11	Graficar (3)			■				■	
12	Replantear (3)	■		■			■	■	
13	Rehacer (2)	■		■		■	■		
14	Escrapear (2)	■					■		
15	Repensar (2)	■					■		

		Escenarios en donde se presenta							
	Código	E1	E2	E3	Editor	Analista	Reportero	Diseñador	Programador
16	Enfocar (2)	■			■		■		
17	Procesar (2)		■	■		■			■
18	Descargar (2)	■	■			■	■		
19	Implementar (2)			■				■	
20	Recopilar (2)	■		■			■	■	
21	Desarrollar (2)	■			■		■		
22	Cruzar información (1)		■						
23	Rastrear (1)		■			■			
24	Armar (1)	■			■				
25	Recuperar (1)	■			■				
26	Borrar (1)	■					■		
27	Maquetar (1)	■					■		
28	Bajar datos (1)	■					■		
29	Diseccionar (1)	■			■				
30	Reeditar (1)			■		■			
31	Capturar (1)	■					■		
32	Fusionar (1)	■			■				
33	Planear (1)	■			■				
34	Negociar (1)			■				■	
35	Descartar (1)		■			■			
36	Sistematizar (1)	■			■				
37	Reconstruir (1)	■					■		
38	Conocer (1)	■			■				

En este cuadro se pueden observar que los verbos que se presenta en los tres escenarios es el de visualización, el cual también se presenta en los diferentes profesionales que intervienen en el desarrollo de la nota. El escenario 1 fue el que

más procesos “In vivo” concentra y fue el que aportó la mayor cantidad de acciones específicas en la elaboración de una nota de periodismo de datos por ejemplo dos ejemplos claros son *jugar con datos*, *escrapear* y *maquetar*, entre otras.

Los reporteros son los que manifiestan un mayor número de procesos relacionados con la categoría “In vivo”; seguidos del editor y el analista. Se observa que las acciones con las que se les relaciona están vinculadas con sus roles en el desarrollo de la nota.

En el caso del editor se presentan acciones como *planear, armar, fusionar y sistematizar*, lo que nos lleva a pensar que sus acciones son más de integración y de trabajo con todo el grupo. Mientras que los analistas y reporteros tienen procesos cognitivos más vinculados con el desarrollo de la nota *(contar historias, solicitar información, rehacer*); En el caso de los diseñadores y programadores tienen poca similitud en sus procesos, ya que solo comparten los de *visualizar y rebotar*.

Aplicación

Esta categoría consta de 13 códigos, en donde la mayor parte de los verbos empleados se ciñen a resolver, encontrar, explicar y construir. Esto supone poner en juego procesos intelectuales para la resolución de problemas (ver, Cuadro 9).

CUADRO 9: PRESENCIA DE CÓDIGOS DE PROCESOS COGNITIVOS DE APLICACIÓN EN ESCENARIOS Y POR TIPO DE PARTICIPANTES

		Escenarios en donde se presenta							
	Código	E1	E2	E3	Editor	Analista	Reportero	Diseñador	Programador
1	Resolver (13)	■	■	■	■	■		■	■
2	Encontrar (10)	■	■	■	■	■	■		■
3	Construir (8)	■	■	■	■	■	■		■
4	Comprobar (4)		■	■		■			■
5	Demostrar (2)	■	■			■	■		
6	Diseñar (2)			■					■

7	Modificar (2)			■				■	
8	Obtener (2)		■	■		■		■	
9	Redactar (2)	■		■		■	■		
10	Calcular (1)			■		■			
11	Aplicar (1)			■	■				
12	Cambiar (1)	■					■		
13	Representar (1)			■				■	

Lo que se observa en el cuadro es que los procesos de resolver, encontrar y construir son los que se repiten en todos los escenarios y en varios de los profesionales. Los verbos encontrar y construir son compartidos por el editor, analista, reportero y el programador. El verbo comprobar se presenta en los segmentos de: proceso inicial, conjetura inicial, proceso de elaboración, falta de información, trabajar con datos y jugar con datos. Algunos verbos exclusivos son: diseñar para el programador; representar y modificar para los diseñadores, cambiar para los reporteros, aplicar para el editor; calcular para el analista. Lo que nuevamente nos lleva a que los roles y procesos cognitivos están vinculados.

Conocimiento

Esta categoría consta de 8 códigos, en donde la categoría que sobresale es la de buscar datos o información; esto tiene referencia a un constante planteamiento de necesidades de información cuya satisfacción da pie a la toma de decisiones, al replanteamiento de rutas o a la integración de datos (ver, Cuadro 10).

CUADRO 10: PRESENCIA DE CÓDIGOS DE PROCESOS COGNITIVOS DE CONOCIMIENTO EN ESCENARIOS Y POR TIPO DE PARTICIPANTES

		Escenarios en donde se presenta							
	Código	E1	E2	E3	Editor	Analista	Reportero	Diseñador	Programador
1	Buscar (27)	■	■	■	■	■	■	■	■
2	Mostrar (11)	■		■	■		■	■	

3	Definir (11)								
4	Recordar (2)								
5	Escribir (2)								
6	Seleccionar (2)								
7	Reproducir (1)								
8	Marcar (1)								

Como se mencionó antes el verbo buscar es uno de los que se presenta con mayor frecuencia en todos los escenarios y en todos los profesionales. Este proceso, además se presenta en 21 familias de códigos, entre las que se encuentran: *valores* del área, proceso inicial, exploración de temas, conjetura inicial, estrategias de validación, cambio de ruta, falta de información, trabajo con datos y jugar con el dato, entre otros. Esto nos lleva a trazar la idea de que a lo largo de la elaboración de un producto noticioso las necesidades de información se muestran dinámicas y transversales en todo el proceso y para todos los involucrados.

En cuanto al proceso definir se presenta en tres profesionales, que son el analista, el reportero y el diseñador. Los verbos que se presentaron en un solo tipo de profesional fueron seleccionar y reproducir mencionados por los programadores.

Comprensión

Esta categoría consta de 11 códigos, de los cuales explicar es el que presenta un mayor número de segmentos (11), las familias de códigos en los que se presentaron son: trayectoria personal, conceptualización del lector, cambio de ruta, visualización, soluciones propuestas y elaboración de videos. Este, sin embargo, no se presenta en todos los escenarios y tampoco en todos los profesionales, solo fueron mencionados por el editor, el diseñador y el programador (ver, Cuadro 11).

CUADRO 11: PRESENCIA DE CÓDIGOS DE PROCESOS COGNITIVOS DE COMPRENSIÓN EN ESCENARIOS Y POR TIPO DE PARTICIPANTES

		Escenarios en donde se presenta							
	Código	E1	E2	E3	Editor	Analista	Reportero	Diseñador	Programador
1	Explicar (11)	■		■	■			■	■
2	Narrar (6)	■	■		■	■			
3	Convertir (3)		■			■			
4	Interpretar (3)		■					■	
5	Expresar (2)			■				■	
6	Organizar (2)	■			■				
7	Transformar (2)		■			■			
8	Traducir (2)		■					■	
9	Discutir (1)			■		■			
10	Identificar (1)			■				■	
11	Ordenar (1)	■			■				

El analista es quien presenta verbos únicos, tales como convertir y transformar. Por parte del editor, es el único que habla de organizar. En el caso del diseñador habla de interpretar, expresar y traducir.

Evaluación

Esta categoría consta de 8 códigos, de los cuales el mayor es revisar con 9 segmentos que se presentaron en las entrevistas del editor y el reportero. El verbo que se presenta en los tres escenarios es relacionar. En este caso ninguno de los verbos se presentó en todos los profesionales. El programador es quien tiene más procesos cognitivos asociados en este código, estos son: aprobar, fundamentar, decidir y escoger. Los que también presentan códigos exclusivos son el editor con categorizar y el analista con probar (ver, Cuadro 12).

CUADRO 12: PRESENCIA DE CÓDIGOS DE PROCESOS COGNITIVOS DE EVALUACIÓN EN ESCENARIOS Y POR TIPO DE PARTICIPANTES

		Escenarios en donde se presenta							
	Código	E1	E2	E3	Editor	Analista	Reportero	Diseñador	Programador
1	Revisar (9)	■	■		■		■		
2	Relacionar (2)	■	■	■	■	■			
3	Aprobar (2)			■				■	■
4	Fundamentar (2)			■					■
5	Decidir (2)			■		■			■
6	Categorizar (1)	■			■				
7	Escoger (1)			■					■
8	Probar (1)			■		■			

El proceso revisar se presenta en 10 familias de códigos, entre los que están: organización, proceso inicial, selección de temas, búsqueda de soluciones y bases de datos sucias. Esto nos lleva a acciones más técnicas en la elaboración de la nota.

Síntesis

Esta categoría consta de 5 códigos, de los cuales ninguno se presenta en los tres escenarios. Sólo programar y exponer se presentan en dos escenarios. En el caso de los profesionales, programar se presenta en reporteros, diseñador y programador. Los verbos que se presentan en un solo profesional son estructurar en el diseñador y agrupar en el caso del editor (ver, Cuadro 13).

CUADRO 13: PRESENCIA DE CÓDIGOS DE PROCESOS COGNITIVOS DE SÍNTESIS EN ESCENARIOS Y POR TIPO DE PARTICIPANTES

		Escenarios en donde se presenta							
	Código	E1	E2	E3	Editor	Analista	Reportero	Diseñador	Programador
1	Generar (8)		■	■		■			■
2	Programar (6)	■		■			■	■	■
3	Exponer (2)	■		■	■				■
5	Estructurar (2)			■				■	
6	Agrupar (1)	■			■				

El verbo generar se presenta en 10 familias de códigos, entre los que están, características de la nota, conceptualización del lector, producto de acuerdo con una audiencia, contar una historia, conceptualizar el dato y jugar con el dato. Mientras que programar, se presentan en organización, proceso de elaboración y, herramientas de programación, entre las principales.

Análisis

Esta categoría consta de 6 códigos, donde el verbo analizar se presenta en tres de los profesionales: el Editor, el Analista y el Reportero. En el caso del Editor se presentan de forma exclusiva los verbos descubrir y especificar, y para el programador, los verbos investigar y experimentar (ver, Cuadro 14).

CUADRO 14: PRESENCIA DE CÓDIGOS DE PROCESOS COGNITIVOS DE SÍNTESIS EN ESCENARIOS Y POR TIPO DE PARTICIPANTES

		Escenarios en donde se presenta							
	Código	E1	E2	E3	Editor	Analista	Reportero	Diseñador	Programador
1	Analizar (11)	■	■		■	■	■		
2	Comparar (5)	■	■	■		■	■		
3	Especificar (1)	■			■				

4	Descubrir (2)	■			■				
5	Investigar (1)			■					■
6	Experimentar (1)			■					■

El proceso de analizar se presenta en 13 segmentos, los cuales corresponden a organización, proceso inicial, selección de temas, conjetura inicial, cambio de ruta, proceso de elaboración, entre otros.

Códigos con mayor frecuencia

Para finalizar la exposición de los procesos cognitivos, se presenta el Cuadro 15, en la que se enlistan los códigos con mayor frecuencia. Los de mayores segmentos codificados corresponden a las categorías de conocimiento y comprensión, lo que nos indica que los profesionales que se dedican a elaborar notas constantemente necesitan entender los datos y la información que tienen en sus manos, pero también requieren reducir la incertidumbre que les causa la ausencia de datos o información.

CUADRO 15: CÓDIGOS DE PROCESOS COGNITIVOS CON MAYOR FRECUENCIA

		Escenarios en donde se presenta							
	Código	E1	E2	E3	Editor	Analista	Reportero	Diseñador	Programador
1	Buscar (27)	■	■	■	■	■	■	■	■
2	Resolver (13)	■	■	■	■	■		■	■
3	Visualizar (12)	■	■		■	■	■	■	■
4	Explicar (11)	■		■	■			■	■
5	Analizar (11)	■	■		■	■	■		
6	Mostrar (11)	■		■	■		■	■	
7	Contar historias (11)	■	■	■		■	■	■	

8	Definir (11)								
9	Encontrar (10)								
10	Revisar (9)								
11	Cuestionar (8)								
12	Intuir (8)								
13	Rebotar (7)								
14	Limpiar datos (7)								
15	Generar (8)								
16	Solicitar información (6)								
17	Construir (8)								

Algunos verbos de la categoría In vivo, se presentan entre los 20 más frecuentes, entre ellos destacan:

- Intuir: proceso que, como se mencionó antes, actúa como una brújula y que está presente en el proceso inicial de elaboración de una nota, la conjetura inicial y jugar con el dato.
- Rebotar se presenta como un proceso importante en la elaboración de la nota y de jugar con el dato, es una analogía de la iteración en cual se relaciona también con la visualización.
- Limpiar del dato el cual se vincula con el proceso de elaboración de la nota, las bases de datos sucias, y es vista como una tarea previa para el análisis de los datos.

Procesos cognitivos por tipo de participantes

La Figura 28 ofrece una perspectiva global del tipo de procesos cognitivos que ponen en juego los integrantes de los equipos de periodismo de datos. Se observa en primera instancia la presencia transversal del término *buscar* el cual tiene alusiones por parte de todos los participantes involucrados. No sólo se trata del código con mayores alusiones diferenciadas por participante, sino también el más referido. El mayor número de segmentos codificados con el verbo buscar provienen

de reporteros, editores y analistas. Recordemos que en primera instancia los que comienzan a plantear temas y quienes tienen un mayor contacto e interacción con la información son los editores, reporteros y analistas, de hecho, también son quienes tienen un mayor contacto con los datos e información.

En un segundo plano se ubica el verbo resolver en este caso los segmentos provienen principalmente de los diseñadores y de la editora. Lo cual es congruente con lo referido por la editora, pues tiene que hacer frente a distintos obstáculos en términos de limitaciones de tiempo, de comprensión de la información o datos con los que se interactúan. En el caso de los diseñadores, los principales puntos de resolución se ciñen a la traducción de los datos en una visualización.

En lo que corresponde al código de contar una historia, las alusiones provienen principalmente del reportero y en menor medida del analista y diseñador. Son este grupo los principales en darle forma al producto en términos de integrar una historia que se cuente con datos, ya sea desde el análisis, la integración de un ángulo para un tema y la visualización de los datos y la información textual que la acompaña. El análisis particular sobre cada tipo de participante nos arroja la siguiente información:

- Reportero: la mayor parte de procesos que pone en juego se relacionan con las tareas de buscar, contar una historia, revisar, analizar y definir. Realiza también acciones como solicitar información y acciones determinadas con los datos como jugar con el dato, limpiar, capturar, maquetar, construir y *escrapear*.
- Analista: por su perfil las tareas que realiza este tipo de participante se decantan principalmente a generar, cuestionar e intuir. Su relación con los datos es palpable en acciones como obtener, convertir, rebotar, comparar, relacionar y narrar una historia con estos.
- Editor: sus actividades se vinculan con la gestión y el control de calidad de la nota que se produce, pues busca, analiza, explica, resuelve, revisa, planea y organiza.
- Diseñador: la búsqueda que realiza este tipo de participante se relaciona con su necesidad de ofrecer soluciones y para ello revisa otros medios para saber

cómo le dieron solución a un problema de representación visual de datos. En ese sentido, se entiende las alusiones al verbo resolver y al verbo buscar. Su relación con los datos se vincula con el propósito de traducirlos a otro idioma, el visual, por ello hay alusiones a verbos como visualizar, rebotar, graficar, optimizar, interpretar, implementar, transformar y traducir.

- Programador: pone en juego procesos cognitivos como buscar, encontrar, resolver, seleccionar, diseñar y comprobar. El rol que desempeña está muy vinculado con los procesos que realiza, en este caso es predecible que aparezca el verbo programar, sin embargo, es el único que hace alusión al proceso de experimentar
- Becario: hace alusiones aisladas al verbo cruzar información y resolver.

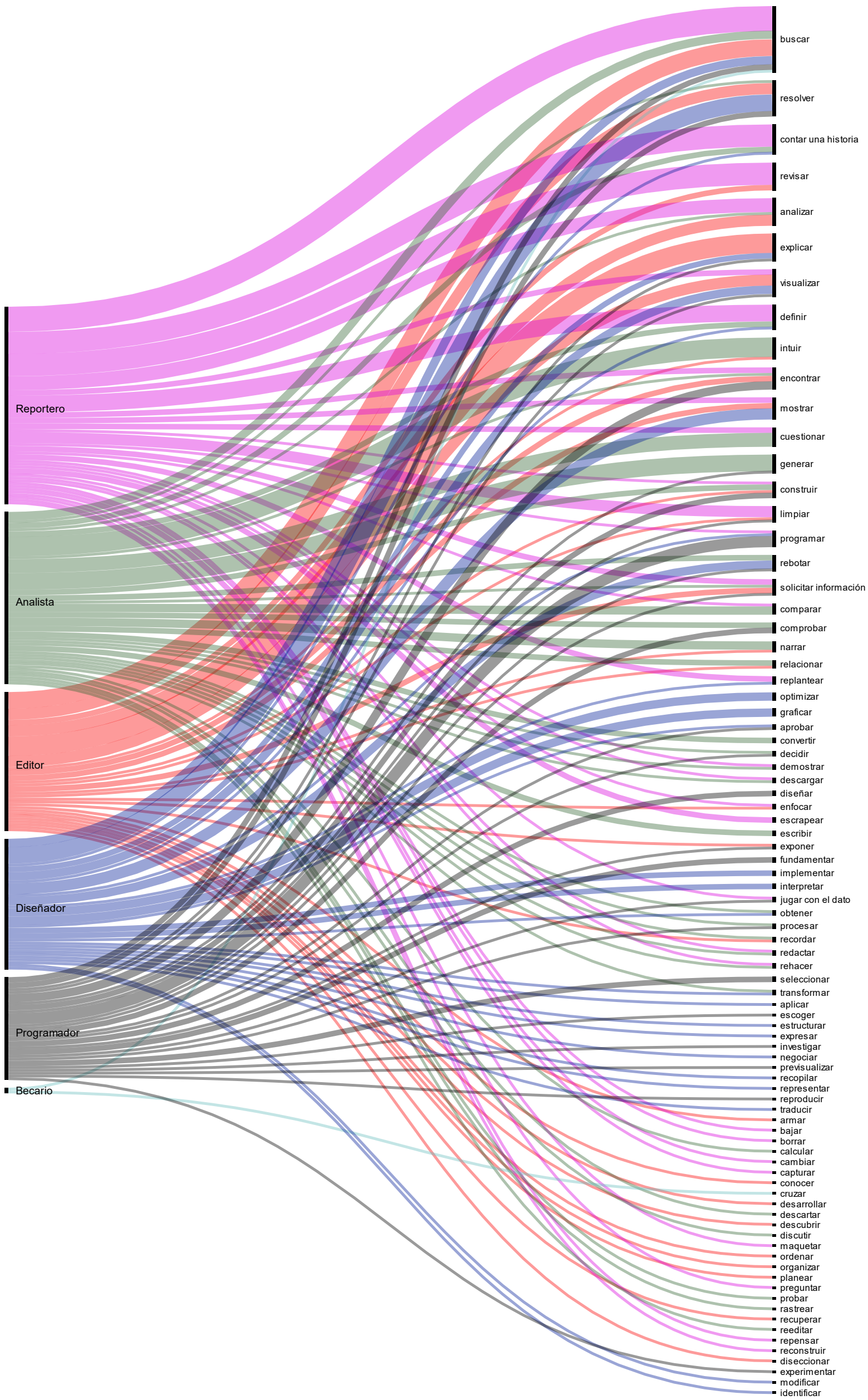


FIGURA 28: PROCESOS COGNITIVOS POR TIPO DE PARTICIPANTE

Nota: En este gráfico se excluyen los segmentos codificados del Informe sobre la elaboración de una nota proporcionado en el E1, por ello hay diferencias en las frecuencias de segmentos por código.

Procesos cognitivos por escenarios

Al hacer el comparativo de los procesos cognitivos expresados de forma implícita o explícita en los tres escenarios (ver, Figura 29), encontramos que los escenarios con el mayor espectro de procesos cognitivos es el E1 y el E3, pues muestran un abanico amplio de códigos verbos de los cuales el proceso de buscar es el más repetido. Este verbo también está presente en los tres escanerios al igual que resolver, contar una historia,visualizar y construir. Lo cual nos habla de procesos comunes.

Los escenarios E1 y E2 refieren a contextos similares, pues se tratan de redacciones de diarios; sin embargo, son pocos los procesos coincidentes. En este caso sólo hay coincidencias significativas en los verbos: buscar, contar una historia, definir, intuir, construir y solicitar información; los demás son procesos aislados, casi exclusivos, lo cual habla mucho de las formas de trabajo, la constitución de los equipos y los flujos de información.

Los escenarios E1 y E3 son también los que incluyen el mayo número de procesos cognitivos esclusivos y tienen coincidencias en los verbos buscar, resolver, visualizar, explicar, encontrar, mostrar, cuestionar, programar, relacionar y rehacer.

En cuanto a los escenarios E2 y E3 hay coincidencias en los procesos de buscar, resolver, visualizar, encontrar, generar, construir, comprobar y comparar. A continuación se hace una caracterización de cada escenario en términos de procesos cognitivos que ponen en juego los miembros de sus equipos de periodismo de datos.

> E1: el mayor número de procesos mencionados se concentran en buscar, revisar, analizar, visualizar, explicar, definir, encontrar, mostrar, limpiar, escrapear y solictiar información. Buena parte de estos procesos los realizan dos personas; la editora y la reportera con el equipo de becarios. Sus verbos exclusivos son: conocer, armar, cambiar, desarrollar, borrar, bajar, redactar, ordenar y organizar, entre otros
>
> E2: los procesos que expresan se ciñen a contar una historia, buscar, revisar, intuir y generar. Los códigos exclusivos son: convertir, interpretar, descartar, traducir y cruzar información.
>
> E3: La mayor parte de los procesos presentes son; buscar, resolver, visaulizar, explicar, encontrar, mostrar, cuestionar, programar, comparar y programar. Algunos

de los verbos aislados son: aplicar, discutir, escoger, experimentar, identificar, representar.

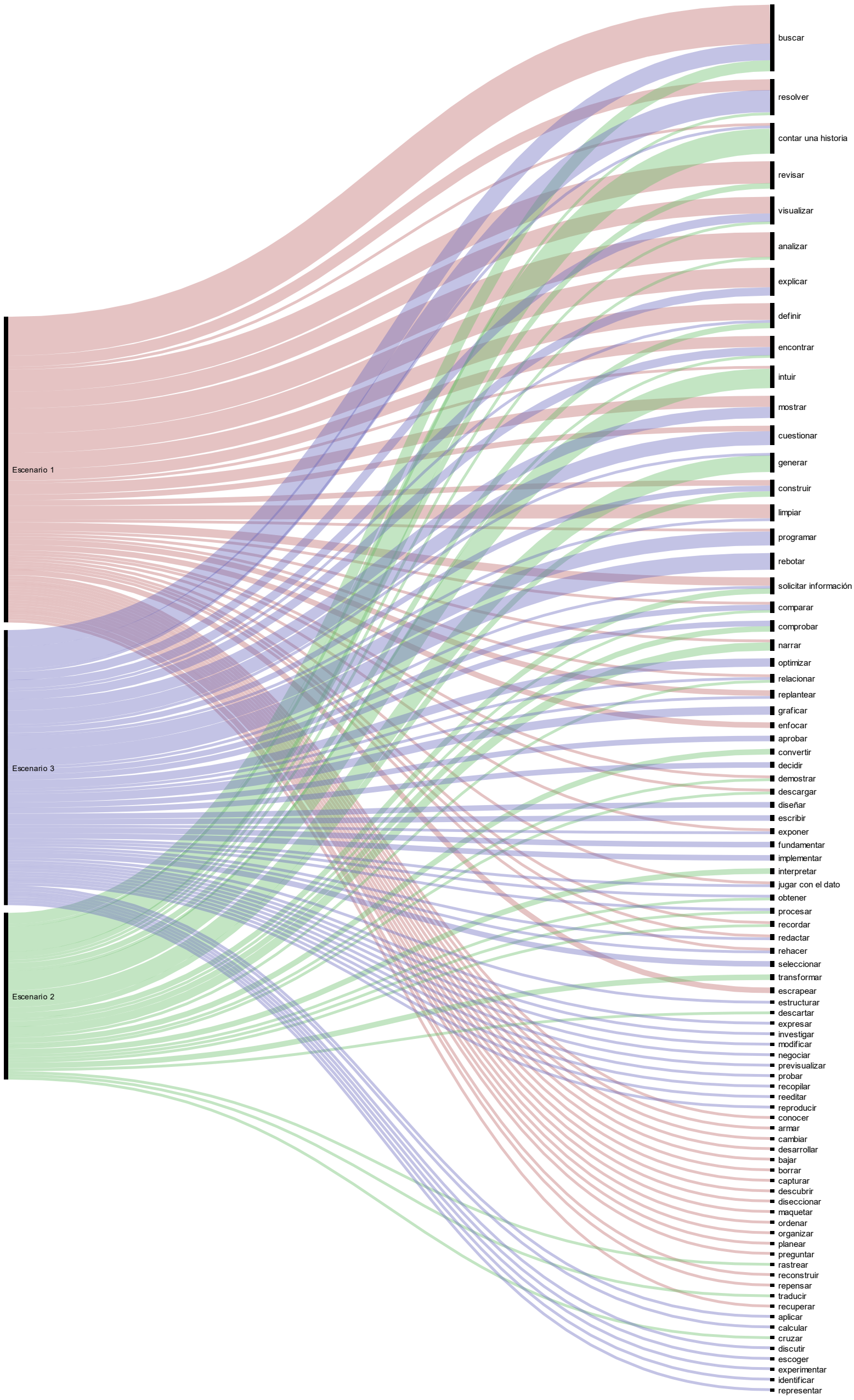


FIGURA 29: PROCESOS COGNITIVOS POR ESCENARIO

Procesos cognitivos por valor agregado

Desarrollo de una nota

El origen de una nota

El proceso inicial para el desarrollo de una nota comienza con la búsqueda de datos o de información, en un proceso donde se involucra también a la revisión de recursos, junto con la intuición, el cuestionamiento, el contar una historia. Se exploran ideas de forma paralela a este proceso, principalmente en el proceso de búsqueda. Al seleccionar un tema se analiza la información, se busca contar una historia, se definen el tipo de información que se requerirá para el desarrollo de la nota.

La conjetura inicial viene de un proceso en el que la comparación, la comprobación y la intuición tienen un papel central, junto con el cuestionamiento de la información o datos recabados. Se busca, además, soluciones similares a la forma en que se han contado determinados temas, en este caso se revisa, analiza la información o los datos y comienza a perfilarse su visualización.

La búsqueda de información o datos es un proceso presente en tres etapas el proceso inicial, la exploración de temas y la conjetura inicial. Le sigue el proceso de comparación que tiene presencia en la selección de temas y la formulación de la conjetura inicial. El análisis por su parte es parte de las etapas de selección de temas, la conjetura inicial y la búsqueda de soluciones similares. Finalmente, la intuición tiene un papel importante durante el proceso inicial y la formulación de una conjetura de inicio. En la Figura 30 se puede observar la concentración de procesos en determinadas etapas, en este caso la conjetura inicial es la que ofrece un mayor rango de procesos cognitivos.

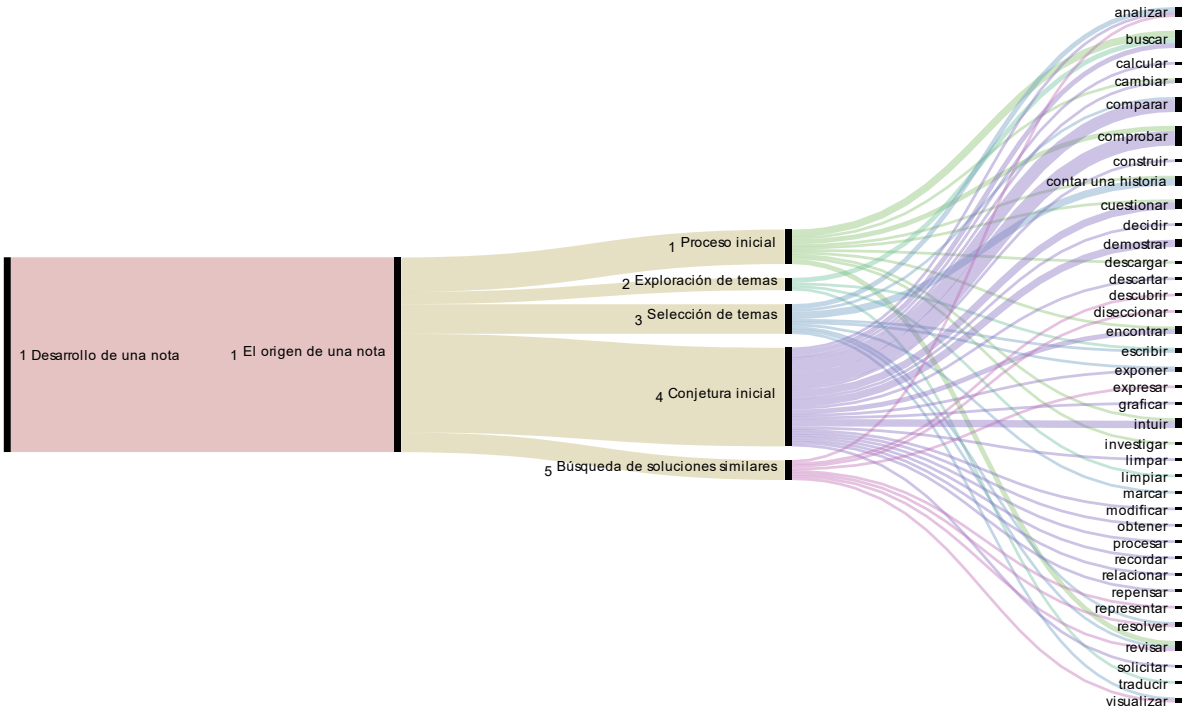


FIGURA 30: PROCESOS COGNITIVOS PRESENTES EN EL ORIGEN DE UNA NOTA

La construcción de una nota

En la Figura 31 se puede apreciar que los procesos cognitivos presentes durante la construcción de una nota tienen su mayor rango de presencia en la elaboración o desarrollo de una nota, pues además de la búsqueda de información o datos, hay procesos como comparar, analizar, investigar, limpiar datos, rebotar ideas, redactar, revisar, sistematizar, por mencionar algunos. Hay un trabajo iterativo que consiste en buscar un ángulo o idea y para ello se rebotan ideas. Como se ha mencionado, la propia información o los datos pueden llevar al participante a replantear el tema, cambiar de ruta y en esta etapa los procesos de explicar, demostrar, rehacer, replantear y validar están presentes.

Por último, se encuentran las etapas de integrar la parte humana en una nota, esto significa integrar entrevistas a expertos o involucrados en una temática en particular, en esta etapa se habla de buscar lo cual refiere a la importancia de considerar a las personas como fuentes de información para cubrir una necesidad

en particular de información. Por otra parte, se observan procesos aislados en las etapas de relacionar la coyuntura y agenda, y las estrategias de validación.

Nuevamente el proceso de búsqueda de información o de datos está presente en la mayor parte de los códigos, junto con analizar y visualizar.

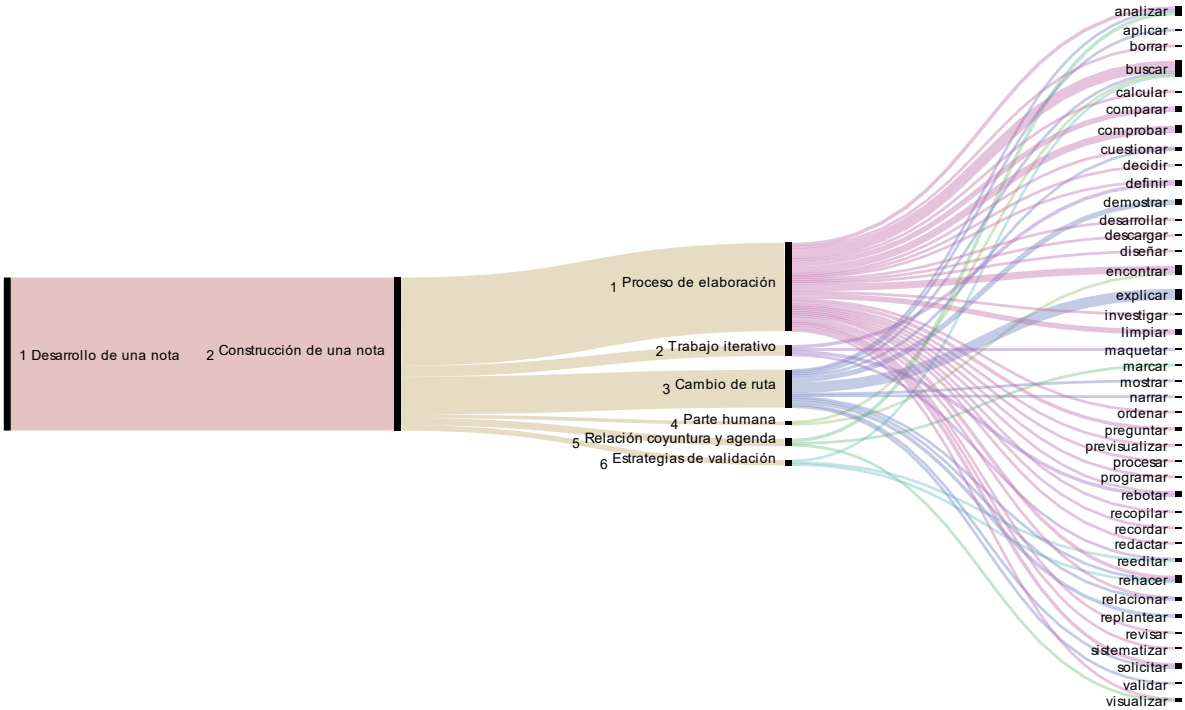


FIGURA 31: PROCESOS COGNITIVOS PRESENTES EN LA CONSTRUCCIÓN DE UNA NOTA

Contar una historia

El rango de procesos presentes en este código es menor, se habla de generar una nota contando una historia a partir del cruce de información (ver, Figura 32).

FIGURA 32: PROCESOS COGNITIVOS PRESENTES EN CONTAR UNA HISTORIA

Obstáculos

En la Figura 33 se puede observar que en términos de frecuencia el mayor rango de procesos cognitivos presentes en obstáculos se concentra en el tiempo. En este caso los procesos que se llevan a cabo están asociados con este código por la inversión que requieren, pues se habla de que se requiere tiempo para buscar información o datos, analizar, comprobar, encontrar, limpiar datos, solicitar información, procesar, revisar y redactar.

Las solicitudes de información, por su parte, son vistas como un obstáculo, pues su obtención implica una labor burocrática que implica invertir más tiempo a una nota e incluso, en aquellos casos en que la información o los datos no existen se habla de construirlos. En ese sentido, la falta de información lleva a procesos de búsqueda, comprobación, solicitud de información, rastreo, limpieza de datos y relación.

Las bases de datos sucias son aquellas que no ofrecen datos sistematizados y consistentes, la forma de resarcir esta carencia es el proceso de limpieza de datos que consiste en un proceso complejo que de suyo involucra, análisis, comparación y comprobación. En este caso se habla de procesos de revisión, limpieza, localización y conversión.

En cuanto a la falta de recursos humanos se analizan y buscan alternativas, como apoyo externo. En la incomprensión de un dato el mayor obstáculo está en el proceso de interpretación, pues puede ocurrir que la información no es clara o no se tienen los presupuestos suficientes para comprenderlos.

Una lectura global de los procesos presentes nos permite observar la presencia, en distintos obstáculos, de los procesos de buscar, encontrar, solicitar información y limpiar.

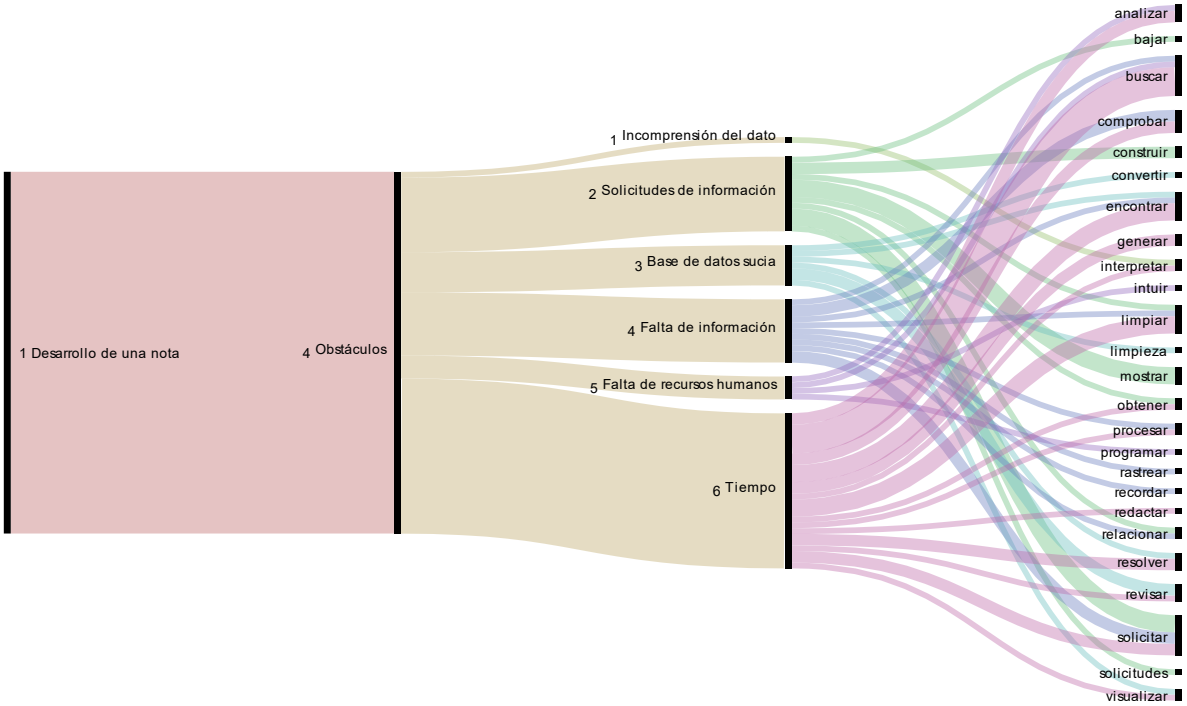


FIGURA 33: PROCESOS COGNITIVOS PRESENTES EN OBSTÁCULOS

Interacción con los datos

Identificación de necesidad de datos

La Figura 34 nos muestra que, al no contar con un dato en específico para comprobar una hipótesis o conjetura, se realizan procesos de localización, limpieza, procesamiento, y relación. Para delimitar lo que se requiere se realizan procesos de búsqueda y cuestionamiento en el que se habla de enfocar la indagación en lo que se requiere en particular.

Si el dato no existe, se continua con el complejo proceso de construir el dato, lo cual conlleva realizar procesos como comprobar, *escrapear*, limpiar, revisar, procesar, rastrear, rehacer, definir, entre otros, este es uno de los procesos más exigentes por lo que implica en términos de la complejidad de los procesos y su inversión de tiempo.

Se observa en general que en este caso no hay un proceso transversal y que la carencia de un dato y su construcción, tienen procesos coincidentes tales como la comprobación, la limpieza de datos, el procesamiento y la realización de

solicitudes de información. Esto habla de su paralelismo, es decir, cuando se identifica la carencia de un dato se da una prospectiva de que se debe hacer para construirlo.

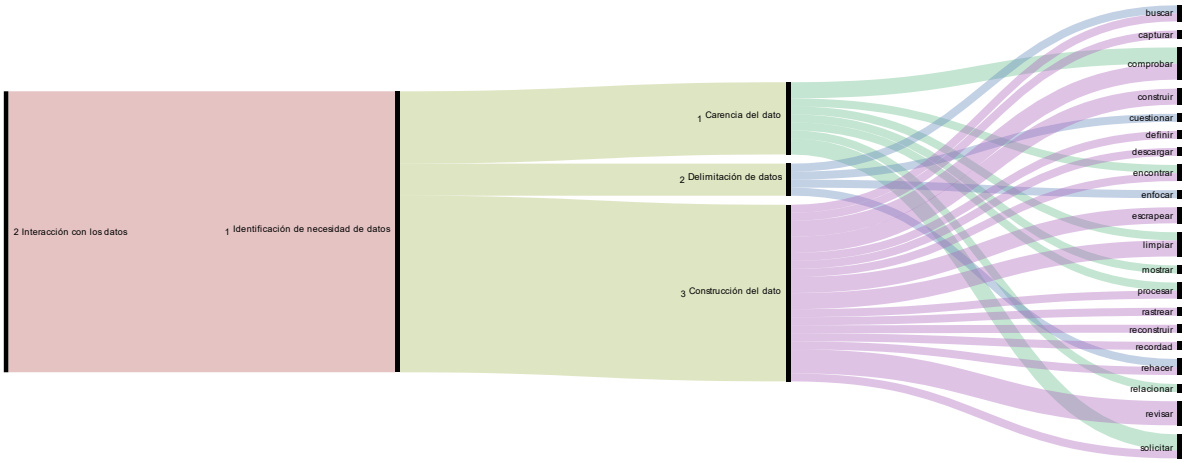


FIGURA 34: PROCESOS COGNITIVOS PRESENTES EN LA IDENTIFICACIÓN DE NECESIDAD DE DATOS

Trabajar con datos

La conceptualización del dato lleva a los participantes a pensar en su uso, pues se habla de demostrar, cruzar y de contar una historia a partir de éstos. Para su extracción y limpieza, hay procesos de *escrapeo*, sistematización, recopilación, desarrollo y limpieza, principalmente. Sin embargo, la mayor parte de los procesos en esta etapa se concentran en jugar con el dato, pues el trabajo constante con los datos en esta dinámica requiere de procesos de búsqueda, análisis, cuestionamiento, definición, comprobación, intuición, juego, validación y otros procesos, que se realizan de forma individual o colectiva con el propósito de darle forma a un producto específico a partir de la negociación del significado de lo localizado y lo que se pretende desarrollar (ver, Figura 35) .

En este proceso, se puede advertir que los datos pueden ser más relevantes que la coyuntura por la forma en que se cuestionan o por su potencial para demostrar un tema. Se habla de que un dato no da lo suficiente cuando no permite mostrar, relacionar o validar una conjetura. Por otro lado, los datos también se valoran en cuanto a sus propiedades y arreglo para construir productos interactivos a partir de éstos.

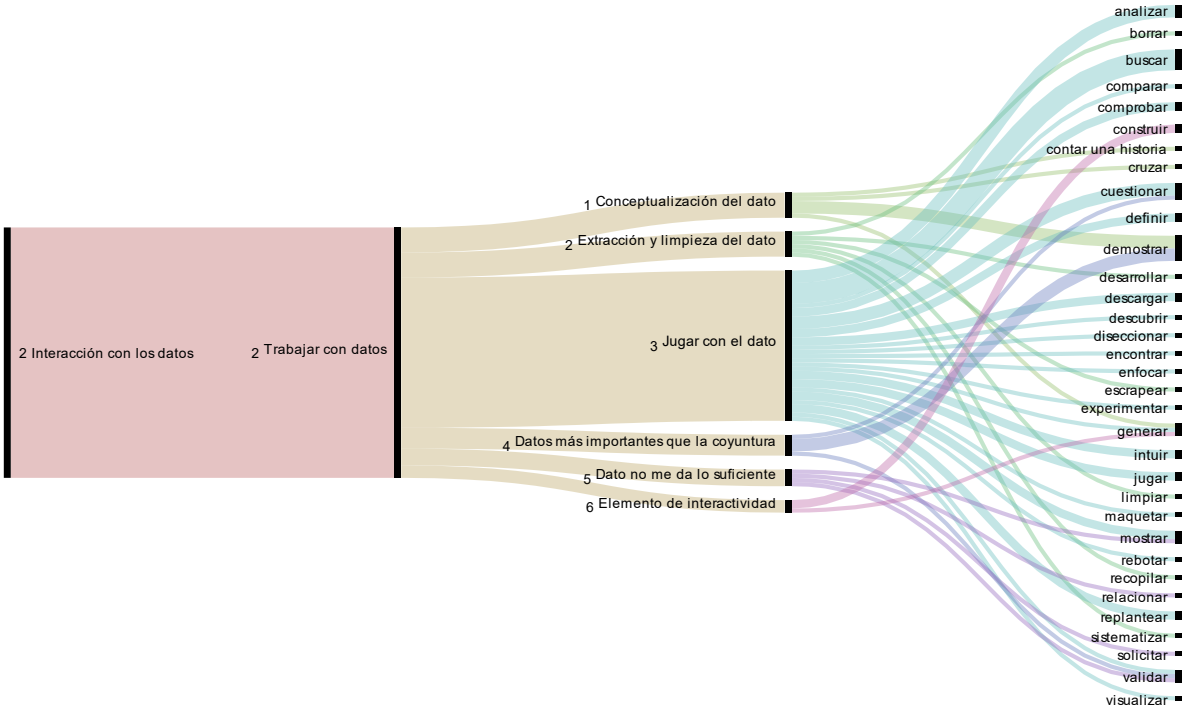


FIGURA 35: PROCESOS COGNITIVOS PRESENTES EN TRABAJAR CON DATOS

Visualización

El proceso de construcción de una visualización involucra la capacidad de resolver, definir y darle forma visual a los datos (ver, Figura 36). Sin embargo, la mayor parte de los procesos en esta etapa se concentran en jugar con el dato, pues el trabajo constante con los datos en esta dinámica requiere de procesos de búsqueda, análisis, cuestionamiento, definición, comprobación, intuición, juego, validación y otros procesos, que se realizan de forma individual o colectiva con el propósito de darle forma a un producto específico a partir de la negociación del significado de lo localizado y lo que se pretende desarrollar.

. La maquetación requiere de un trabajo de definición de lo que se pretende visualizar. La forma en que se pretende representar en gráfico un conjunto de datos lleva a rebotar e identificar opciones.

Una buena visualización es la que consigue explicar, mostrar y contar una historia por medio de un lenguaje visual claro. Esto implica para el participante, ponerse en el lugar del receptor y orientar los insumos al logro de estos fines.

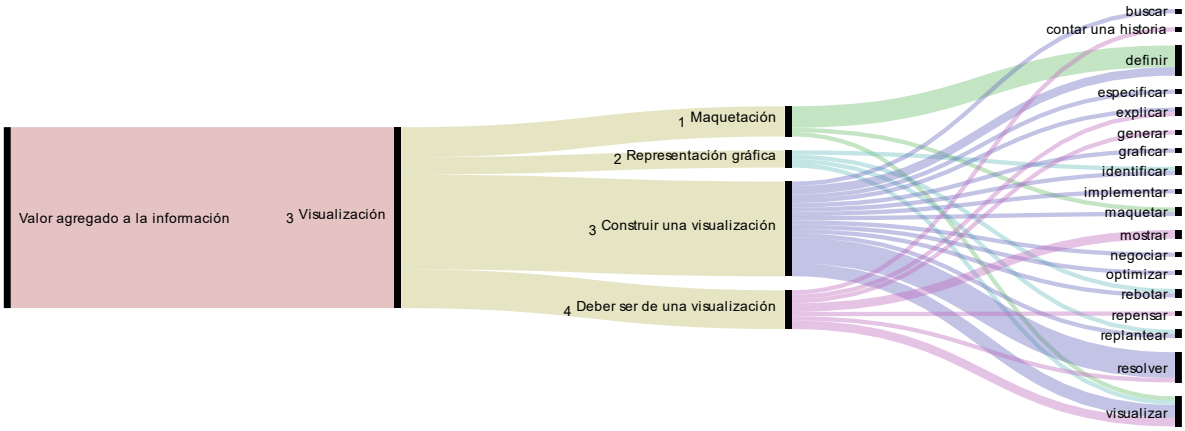


FIGURA 36: PROCESOS COGNITIVOS PRESENTES EN VISUALIZACIÓN

Programación

La programación es otro ángulo o perspectiva para los datos, en este caso se habla de su formato o acomodo para implementarlos u optimizarlos para su integración en sitios o portales. Gran parte de las soluciones que propone esta área se ciñen al objetivo de que la nota explique un tema complejo, en ese sentido orientan su trabajo en la optimización de los datos por medio de la programación. Por último, se habla de la programación y los procesos que ponen en juego al utilizar herramientas propias de esta área, en este caso: experimentar, jugar, mostrar y, claro, programar.

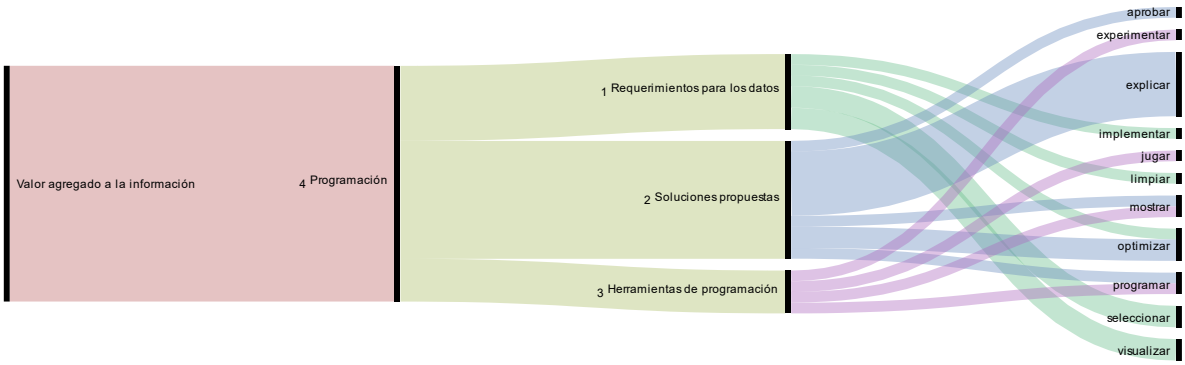


FIGURA 37: PROCESOS COGNITIVOS PRESENTES EN PROGRAMACIÓN

Producción multimedia

La producción de vídeos o contenidos multimedia, de acuerdo con los participantes, involucran procesos de explicar y generar contenidos. Tómese en cuenta que para este estudio no se consideró al personal que trabaja directamente en esta área.

FIGURA 38: PROCESOS COGNITIVOS PRESENTES EN PRODUCCIÓN MULTIMEDIA

DISCUSIÓN

El primer punto de discusión a partir de los resultados obtenidos es que los reporteros y editores entrevistados; hacen alusión a procesos y actividades que se relacionan con el uso de diferentes recursos y fuentes de información, prevaleciendo de manera importante las fuentes primarias y la búsqueda de conjuntos de datos. Esto nos llevaría a cuestionar las diferencias y similitudes entre los conceptos de comportamiento informativo (CI) y comportamiento en el uso de datos (CUD).

En lo general, podemos afirmar que ambos conceptos se basan en necesidades de información, tiene un proceso para satisfacer dichas necesidades y al final se obtiene un producto o se realiza una acción. Sin embargo, al hablar de datos, no se trata sólo de una necesidad que se basa en la localización de un documento sino en ubicar datos, con determinadas características y para determinados fines. Esto implica que para utilizarlos se deben someter a tratamientos de análisis, limpieza o sistematización, lo cual no garantiza que la necesidad sea atendida, pues no se trata de un insumo del que se dispone de forma inmediata y por lo mismo no hay claridad sobre su alcance para satisfacer una necesidad en concreto.

Es un hecho que el aspecto personal y el contexto son los factores que inciden en el comportamiento informativo de los sujetos. En el caso particular de

este estudio se identifica que los participantes se debaten entre su agenda personal; las temáticas demandadas por la organización y la coyuntura social. De esta forma la selección de la temática a abordar no es personal como ocurre con otros profesionales, sino que se deriva del análisis y toma de decisiones grupal.

El participante permanentemente está buscando ideas que sean susceptibles de ser narradas con datos y de forma visual. Para ello retoma las buenas prácticas de periódicos nacionales o extranjeros; e inclusive adapta la forma de mostrar los datos para replicar notas con datos locales, que le parecen interesantes.

Sin embargo, a diferencia de otros profesionales, además de buscar fuentes de información confiables, está buscando errores, omisiones, inconsistencias en la información que se publica en diferentes medios; y lo retoma para producir y publicar notas, en las que se valide la información, se reinterpreten los datos a partir de perspectivas novedosas e innovadoras. También busca información o temas que sean poco comprensibles para hacerlos legibles a su audiencia y dar elementos para la reflexión.

La búsqueda de datos y de información son actividades permanentes que permean todo el proceso de elaboración de la nota, pero también el antes y el después de la misma. Esto plantea cuestionamientos a modelos como el de Belkin (1982) y Atffield (2005) en el que hay un esquema de Situación-Vacío-Uso; y lo que se observa en este caso es Vacío-Procesamiento de datos-Uso-Vacío-Escritura-Vacío; es decir, el vacío es una situación de incertidumbre gradual durante el desarrollo de una nota, que busca disiparse por medio de la búsqueda o la construcción de datos.

La incertidumbre, como se puede percibir en los modelos de CI como los de Kuhlthau (1991), Dervin (1983) y Wilson (1999), juega un papel relevante en el proceso de elaboración de una nota de periodismo de datos, pues no hay una idea clara de lo que se quiere obtener en las primeras interacciones con los datos. Lo que sí hay es una disposición tácita del periodista a buscar información y a ubicar oportunidades para generar o desarrollar una nota a partir de lo hallado. En ese sentido, se busca reducir de forma gradual el sentido de incertidumbre, tal como ocurre en el estudio de Attfield, Blandford y Dowell (2003b). Se buscan otros

ángulos, se revisan los datos de solicitudes previas, se presenta en los participantes un estado de alerta constante que actúa como un olfato en busca de temas, así como una disposición permanente a contar historias con datos. Se puede apreciar procesos que son transversales durante el desarrollo del producto, éstos son la búsqueda de información y la elaboración de la nota. En cuanto a la selección de temas puede ser abierta, pero también puede ser asignada por un área editorial o un cliente, como en el caso del E3.

Hay prácticas comunes en el consumo de fuentes de información, como la lectura de la competencia y de medios que son considerados referentes. Entre los propósitos de lectura se ubica el conocimiento de los temas de coyuntura, la emulación de una nota o el descubrimiento de soluciones a problemas afines.

Los participantes tienden a tomar decisiones a partir de lo que encuentran en los datos; cuestionan los datos. Si tienen una idea o una hipótesis, ésta se replantea o se reformula con base en lo hallado.

En el plano personal destacan tres elementos: la intuición, las motivaciones y las disposiciones. Estas son los elementos que permiten construir una conjetura, lo cual detona la necesidad de información y conduce la búsqueda de datos e información. La intuición, que es la percepción que motiva a comprobar o descartar datos, funge como un criterio para la identificación de temas o situaciones que deseablemente pueden tener un impacto mediático. La intuición, es vista como un elemento de rasgos imprecisos cuyo propósito es encontrar una idea o tema que subyace en los datos con los que se interactúa. No siempre es certera en sus suposiciones, pero es motivada por la idea de comprobar o desechar una hipótesis.

Estas condiciones enmarcan la elaboración o construcción de una nota; siendo el procesamiento del dato (limpieza, procesamiento y análisis) lo que determina la calidad de los datos, la búsqueda de datos alternos y la definición del nivel de detalle de los datos que se utilizarán. Se consideran también las interacciones con otras áreas que contribuyen a dar forma a la nota; es por esto por lo que se habla de la práctica de *rebotar ideas* para armar una propuesta integral de texto e imagen. Posteriormente surge una dinámica iterativa en la que se negocia lo que es y debería ser el producto final. A decir de Attfield, Blandford y Dowell

(2003b), el periodista está involucrado en un proceso continuo de “movimientos” o ajustes, que se basan en evaluar cada resultado y realizar nuevos ajustes en un continuum en el que no hay “movimientos predeterminados”. En otras palabras, en el desarrollo de una nota de periodismo de datos, no hay recetas probadas.

La construcción de la nota es dinámica porque la idea de lo que es o debería ser el producto final tiende a cambiar mientras se avanza en su elaboración. La información disponible y la interacción con ésta marcan el rumbo del producto producido. En ese camino hay cambios de ruta; es decir, replanteamientos o ajustes que pueden ser negativos o positivos de acuerdo con la factibilidad de demostrar o respaldar con datos una hipótesis. Estos ajustes pueden derivar en el replanteamiento de la idea de la nota, o bien descartarla por completo. Como está visto durante el proceso de búsqueda de información y elaboración de la nota se pueden encontrar temas o ideas que desplazan el planteamiento inicial. Lo cual es visto como un proceso desestabilizador que resulta de la búsqueda de información, el hallazgo de nueva información y el proceso de elaboración de la nota, lo cual coincide con lo hallado por Attfield y Dowell (2003a).

Algunos de los obstáculos que afrontan los participantes durante el desarrollo de una nota de esta naturaleza son: restricciones de acceso a la información pública; limitaciones para trabajar con datos considerados como confidenciales; inexistencia de los datos requeridos; demora en las solicitudes de información a instancias gubernamentales; imposibilidad de manipular archivos que fueron entregados en formatos cerrados, y recopilación de información que se encuentra desperdigada en distintas instancias.

Como se puede apreciar, la disponibilidad de los datos no significa siempre la posibilidad de trabajar en su análisis de forma inmediata. Ejemplo de ello son las bases de datos sucias o inconsistentes, pues requieren de un tratamiento que involucra una gran inversión de tiempo y de recursos humanos y técnicos para su análisis. Otro obstáculo en ese sentido es la incomprensión de los datos en dos vertientes: (1) incomprensión de la base de datos al no contar con un glosario o estar sucia (con inconsistencias en los datos que concentra) y (2) incomprensión de la temática por su nivel de especialización.

Otro obstáculo que señalan los participantes es el tiempo, pues opera como un condicionante de lo que es o no factible de hacer, considerando los esfuerzos y las tareas de otras áreas, pero también las exigencias al interior de las instituciones, las cuales demandan la publicación de notas día a día. De ahí que el tiempo se asuma como un parámetro elemental para determinar el alcance de una nota. Por último, se habla de la plataforma de salida como un obstáculo por sus restricciones en términos de espacio, número de caracteres, formato y plataforma de interacción.

Uno de los principales hallazgos de la tesis es la categoría de *interacción con los datos*. Se trata de uno de los hallazgos repentinos que emergen de la contrastación de los datos cualitativos y del empleo de la TF. En este caso el dato es visto por los participantes como un insumo, un argumento, una fuente de información cuya validez reside en su origen (en su mayoría de instancias gubernamentales), pues, a decir de los participantes, el utilizar datos gubernamentales les da mayor peso a los productos noticiosos en términos de validez.

Los datos son vistos como elementos que aportan grados de objetividad, confiabilidad y transparencia a una nota. La interacción con éstos permite comprobar, refutar o replantear una nota. Su disponibilidad, como ya se ha señalado, no significa la satisfacción de una necesidad, pues en ocasiones requieren trabajo adicional de limpieza y procesamiento. En ocasiones se construye a partir de distintas fuentes, no es una información dada, requiere, para su uso, de la intervención del participante o de un equipo que trabaja desde distintos ángulos el tratamiento del dato, se captura, se limpia, se analiza, se traduce, se interpreta con distintas motivaciones que van de la homologación a la visualización y publicación de estos.

La interacción con los datos es llamada también de forma coloquial por los participantes como “jugar con el dato”, lo que supone el cuestionamiento de éstos, la búsqueda de patrones o tendencias, hacer y deshacer la idea de la nota. Esta categoría mantiene un vínculo con la intuición, pues opera motivada por la percepción de una idea formada (o por formarse) de lo que será el producto noticioso.

Las primeras interacciones con los datos permiten estimar su alcance para respaldar una nota. Se habla de que en ocasiones el dato “no da lo suficiente”, pues a pesar de que hay claridad sobre el rumbo de una nota, el dato ubicado no es suficiente para desarrollarla. En ese contexto, la idea de la nota o tema que desea desarrollarse es un regulador al momento de valorar los datos que se obtienen. En otras palabras, la interacción con los datos está condicionada por la hipótesis que se busca demostrar, de tal modo que esta hipótesis se desecha o se desplaza por un nuevo ángulo o hallazgo de mayor fuerza. Esta prospectiva de la nota cuestiona los datos que se tienen y la idea de lo que se quiere producir a partir de éstos. La prospectiva de lo que será un producto noticioso no es exclusiva de un rol determinado sea de editor, reportero, analista o diseñador.

Al hablar de visualización de los datos, se alude a los requisitos previos para que el personal de visualización comience a “jugar” con éstos: limpieza, consistencia y congruencia estructural de los datos son indispensables para empezar a trabajar con ellos. Se elaboran maquetas o borradores de visualizaciones, en ocasiones se utilizan datos ficticios con el propósito de ilustrar una idea. Al igual que en el trabajo textual, durante este proceso se replantea lo que se quiere desarrollar en función de los datos y de lo que se pretende transmitir. Las visualizaciones están condicionadas a su vez por el público al que están dirigidas; es la audiencia la que determina su complejidad.

A la pregunta de cómo debe ser una visualización, se describe que ésta debe ser directa, elaborada con información validada, clara, que permita interactuar con los datos, memorable, legible, enriquecida, de fácil comprensión y visualmente atractiva.

El análisis entre valor agregado y procesos cognitivos hace referencia a una serie de verbos que expresan acciones que se llevan a cabo en el plano intelectual. De acuerdo con los resultados obtenidos, la construcción de una nota para el periodismo de datos pone en juego procesos que son claramente identificables, tales como la búsqueda de información o datos, la resolución de problemas, contar una historia, rebotar ideas, analizar entre otros. El proceso de limpieza de datos

refiere al objetivo de homologar, sistematizar y darle consistencia a un conjunto de datos por medio de herramientas tecnológicas, para hacer comparable los datos y visualizarlos. Otro de los verbos que merece atención es el de “traducir”, el cual apunta hacia la comprensión del dato y su expresión clara, esta traducción implica el uso de diccionarios, de definición de variables, e inclusive de la reconstrucción en la forma en la que el dato fue compilado e incluido por el organismo que la emite.

El esquema propuesto, a partir de los resultados obtenidos, para dar valor agregado a la información incluye cuatro categorías de procesos: *Desarrollo de la nota*, *Interacción con los datos*; *Visualización* y *Programación*; de acuerdo con los resultados obtenidos en el contexto de la organización, no son procesos lineales y su desarrollo depende de los obstáculos de personal, tiempo y de los recursos y herramientas disponibles; inclusive puede depender de la coyuntura social.

En el Anexo 2 se puede percibir la riqueza y distribución de los verbos que denotan el proceso cognitivo en la elaboración de la nota. Comparando este proceso con el propuesto por Taylor (1966) se puede afirmar que la complejidad de la vida real supera a los modelos. Hay verbos como buscar y analizar que se presentan en todos los procesos de valor agregado, mientras que hay otros que son exclusivos de una categoría como el de “recordar”. Esto, sin embargo, no significa que no se presente, sino que no fue expresado por el sujeto y que no se podía inferir su presencia a partir de lo expuesto por el entrevistado.

A continuación, en la Figura 39 se caracterizan los procesos observados durante el desarrollo de un producto noticioso a partir de la evidencia empírica sistematizada a lo largo del estudio.

.

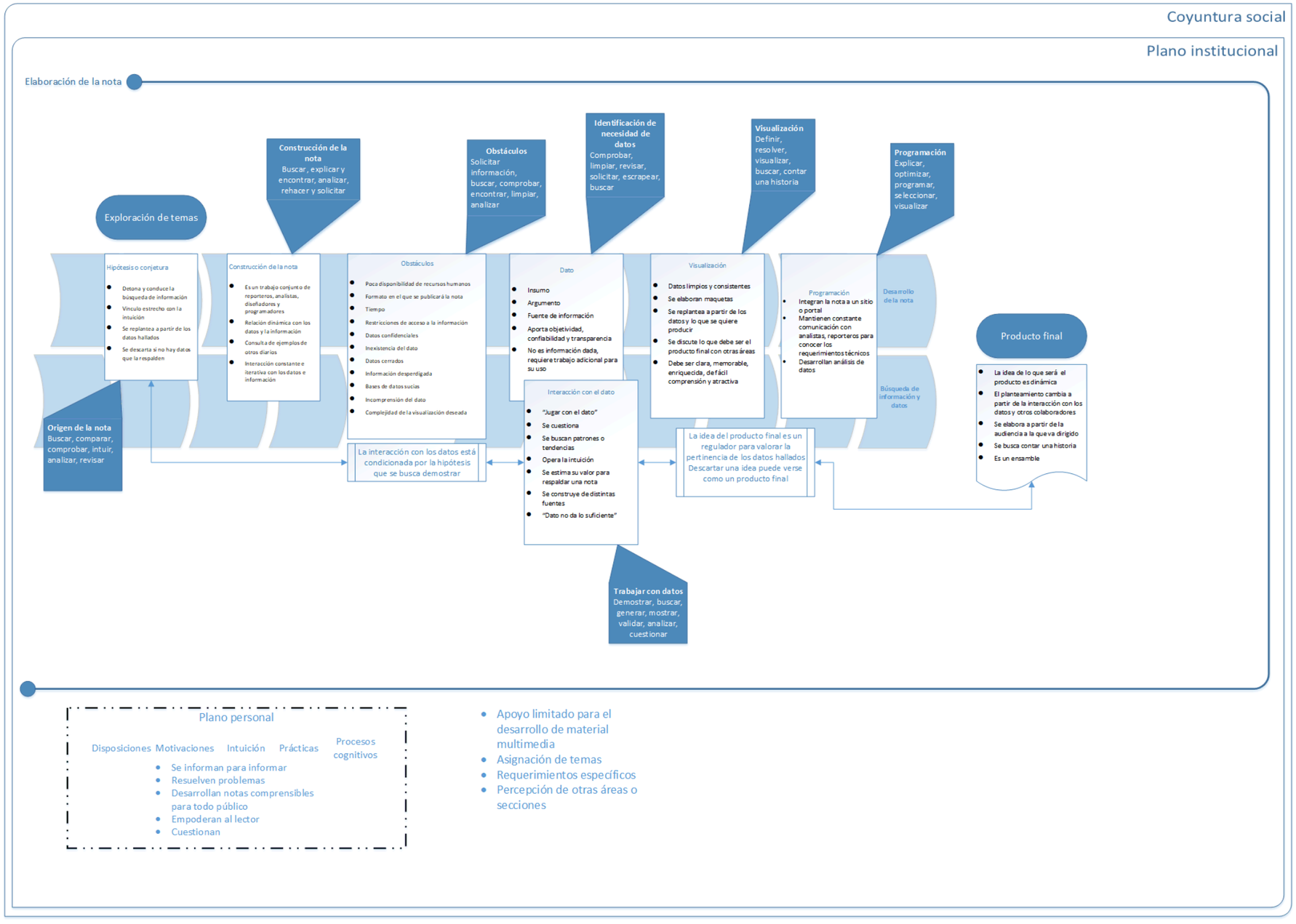


FIGURA 39: DESARROLLO DE LA NOTA, CARACTERIZACIÓN DE PROCESOS

- Coyuntura social: es el espacio del entorno, de los hechos y de las expectativas sociales en términos de temas de coyuntura o de agenda.
- Plano institucional: es el contexto de la empresa, el espacio de intercambio y formación, el espacio de los requerimientos o exigencias institucionales y sus limitaciones.
- Plano personal: es el espacio que conduce las prácticas del participante a partir de sus disposiciones, motivaciones e intuición. Es también el plano en que se ponen en juego los procesos cognitivos. En su conjunto son elementos que operan como una brújula de acciones hechas y por hacer; así como de formas de pensar.
- Elaboración de la nota:
    - Exploración de temas: la constante exposición a medios de información, referentes y a la competencia por parte del participante, le permite aproximarse a temas o ideas para el desarrollo de una nota. Este acercamiento también está presente durante el desarrollo de la nota, pues siempre se mantiene un estado de alerta o apertura para plantearse nuevos temas o cambiar el que se trabaja. Se ponen en juego procesos de búsqueda, comparación, comprobación, análisis, revisión e intuición.
    - Los procesos de desarrollo de la nota y la búsqueda de información son paralelos y transversales presentes a lo largo de cada proceso de elaboración de la nota.
    - Hipótesis o conjetura: detona y conduce la búsqueda de información o de datos y guarda un vínculo estrecho con la intuición. Si no hay suficientes datos que ayuden a respaldarla se descarta o se replantea. Condiciona en buena medida la interacción con los datos.
    - La construcción de la nota: es un trabajo conjunto de reporteros, analistas, diseñadores y programadores que ponen en juego procesos de búsqueda, explicación, localización, análisis, solicitudes de

información y reformulación. Hay una relación dinámica con los datos que oscila entre la obtención, transformación y despliegue.

- Obstáculos, se ponen en juego procesos de búsqueda, comprobación, localización, limpieza, análisis y se realizan solicitudes de información, para darle solución a los siguientes aspectos:
    - Poca disponibilidad para reaccionar de forma inmediata o realizar trabajos de mayor profundidad, pues en algunos casos no hay suficiente personal.
    - El formato de salida es visto como un obstáculo pues el contenido debe ajustarse a los requerimientos de espacio y compatibilidad del entorno de salida, sea impreso o digital.
    - El tiempo es un recurso y un obstáculo que sirve como un parámetro para estimar el alcance o nivel de profundidad de una nota que se trabaja. No sólo es el tiempo del periodista de datos, sino también del equipo de diseñadores o programadores involucrados en el proceso de desarrollo de una nota.
    - Restricciones de acceso a la información: se habla de datos no proporcionados por instancias principalmente gubernamentales.
        - Datos confidenciales
        - Datos cerrados
        - Información dispersa en distintas instituciones o instancias
    - Inexistencia del dato: el dato deseado o requerido para el trabajo de una nota no existe y por lo tanto se toma la decisión de descartar la idea o construir el dato a partir de la consulta de distintas fuentes y su procesamiento.
    - Bases de datos sucias: no hay consistencia ni homogenización de los datos concentrados, lo cual supone un trabajo de limpieza e inversión de tiempo.

    - Incomprensión del dato: no se cuenta con los suficientes conocimientos para interpretar los datos obtenidos o bien la forma en que están estructurados los datos no son claros.
    - Complejidad de la visualización deseada: los requerimientos técnicos de los datos, la complejidad del requerimiento, las posibilidades de los programas informáticos o las capacidades personales impiden obtener la visualización deseada.
  - Dato: es visto como un argumento, un insumo, una fuente de información que aporta objetividad, confiabilidad y transparencia. Para su uso requiere de trabajo previo. La relación en términos de procesos cognitivos, que se establece con los datos son de extracción, comprobación, limpieza, revisión y la propia solicitud de estos.
  - Interacción con los datos: es el punto en el que se cuestionan, comparan, construyen o complementan los datos. Se habla de “jugar con el dato”, buscar patrones o tendencias, se valoran los datos en función de la idea de los que será la nota o bien se concluye que un dato “no da lo suficiente” y se descarta la idea (lo cual puede verse también como un resultado final). Los procesos presentes en esta etapa corresponden a demostrar, buscar, generar, mostrar, validar, analizar y cuestionar.
  - Visualización: es el momento en que comienza a trabajarse en la representación gráfica de los datos. Para llegar a este punto se requieren datos limpios y consistentes. En algunos casos se elaboran maquetas con datos ficticios para ilustrar y discutir, con otras áreas, lo que sería el producto final. Se busca que sea clara, memorable, enriquecida, atractiva y de fácil comprensión. Hay procesos cognitivos presentes como: definición, resolución de problemas, visualización, búsqueda y contar historias.
  - Producto final: es una idea dinámica de lo que será el resultado y está presente a lo largo de todo el proceso. Puede cambiar o descartarse a partir de la interacción con los datos. Su nivel de complejidad es

acorde con la audiencia a la que va dirigido. Puede verse como un ensamble o como un producto colectivo en el que participan profesionales con distintas ópticas y propósitos.

CONCLUSIONES

Después de analizar los resultados de la investigación se puede observar que la relación entre el usuario, la información y los datos es dinámica y está enmarcada por una necesidad y una expectativa sobre lo que se pretende desarrollar con los datos. El análisis de la información obtenida permite observar la complejidad de trabajar con grupos de diversas disciplinas que buscan darle sentido a un producto en común, lo cual nos remite a la construcción social de conocimiento por medio de la negociación de significados o bien al área del comportamiento informativo colaborativo que supone la confluencia de distintos actores que realizan una tarea concreta y que requiere de habilidades o roles diferenciados para el beneficio intencional y mutuo.

En este caso, la hipótesis o la conjetura de una nota es un punto de partida que conduce a la búsqueda de información y, en función de ello, el participante selecciona o descarta información. Es a partir de lo descubierto en esta etapa que se plantean temas, los cuales se sopesan con otros durante el proceso de búsqueda para, finalmente, determinar el alcance de las tareas que implicará su desarrollo y si derivarán en un producto o nota de corto o largo plazo.

Se infiere que el periodismo de datos es un conjunto de saberes y propensiones que operan como una forma de cultura que se reproduce por los

participantes a través de prácticas concretas, como informarse para informar; cuestionar los datos hallados; buscar irregularidades, anomalías o delitos; generar notas amigables o entendibles para el público en general y, con todo ello, difundir o evidenciar una problemática.

Con la determinación de descubrir historias, ganar la confianza del lector y empoderarlo para que se informe y actúe, los participantes realizan distintas actividades en su día a día: buscan que sus notas sean claras y con un gran impacto visual; consultan otros medios para revisar su agenda y la forma en que tratan determinados temas. Se mantienen en un continuo estado de alerta para encontrar posibles temas a desarrollar. Monitorean las actividades de otros medios que son referentes en soluciones narrativas o visuales que puedan aplicar a su trabajo. Su labor implica también el reconocimiento del entorno social de los lectores a la que van dirigidos sus productos, y en función de ello definen qué tan compleja será su nota, sin perder de vista la idea de que el éxito de una nota radica en la facilidad de su comprensión y en sus posibilidades de ser vista como un contenido significativo en el pensar y actuar de sus lectores.

Los participantes marcan constantemente su diferencia con el diarismo y, por esta razón, se alejan del periodismo basado únicamente en declaraciones, o en la difusión de boletines. Esto modifica en buena medida la dinámica de trabajo en relación con su comportamiento informativo y el uso de datos , así como en la naturaleza de las notas que producen. Justifican la inversión que hacen del tiempo cuando se trata de notas a largo plazo en términos de la profundidad que dedican a una investigación, la recopilación de datos, su tratamiento y limpieza, su análisis y síntesis, así como la integración de entrevistas.

El periodismo de datos puede ser visto también como una cultura de colaboración en la que se comparten datos y experiencias entre especialistas, como diseñadores, analistas y programadores que intervienen en un proceso de sistematización colectiva basado en estructuras formales e informales de intercambio. Estas redes permiten a los participantes resolver problemas como la falta de personal especializado o la obtención y el tratamiento de datos. Es común

que se hable de proyectos colaborativos con equipos de trabajo multidisciplinarios que trabajan de forma simultánea bajo un sistema común de valores.

La aproximación que se obtuvo en esta investigación a la figura del periodista de datos permite conocer la perspectiva de un tipo de usuario que trabaja de forma intensa con datos y que, por lo mismo, afronta obstáculos muy particulares; por ejemplo: la poca consistencia o sistematización de los datos, la incomprensión de estos a falta de un glosario o por su nivel de tecnicismo, el nivel de especialización requerido para utilizar programas especializados para el tratamiento de datos. Encara también simulaciones de acceso a la información pública, pues en ocasiones la información que reciben se ubica en archivos cerrados que impiden trabajar directamente con los datos, pues se trata de tablas entregadas en formato de imagen que, al procesarlas para reconocimiento de caracteres, pierden sus propiedades. Este tipo de retos propician el intercambio de saberes en canales informarles, pues se comparten información técnica para afrontar este tipo de obstáculos.

Por otro lado, la analogía del artesano y el periodista de datos es ilustrativa, pues antes de trabajar el barro el artesano prepara sus insumos, para luego moldearlos y darle forma a un producto final. Lo mismo ocurre con el periodista que asegura que los datos cumplan con determinadas condiciones para su sistematización y comparación. El proceso de moldeado, en este caso, consiste en hacer y rehacer un producto a partir de los datos, se formulan preguntas y se buscan vetas o tendencias. La relación con los datos, en este contexto, se puede agrupar en procesos de obtención, transformación y despliegue.

- Obtención: se busca el acceso a un conjunto de datos principalmente gubernamentales. En caso de que la información no exista o no sea facilitada el dato se construye a partir de la sistematización de datos desperdigados o la captura de datos de forma manual.
- Transformación: los datos se someten a tratamientos de sistematicidad y limpieza, por medio de programas especializados, para que los datos sean consistentes, comparables y susceptibles de visualizarse.

- Despliegue. se busca la mejor forma de visualizar los datos acorde con un objetivo y público especifico.

Es comprensible entonces que, debido a los obstáculos presentes en el ejercicio de este tipo de periodismo, en los estudios revisados sobre comportamiento informativo de periodistas pocas veces se trata el tema de la interacción con los datos. En el caso de esta investigación, la aproximación a la perspectiva de los participantes revela que el dato es percibido como un detonante de un tema, un insumo, un argumento, un elemento de objetividad y transparencia que, a diferencia de la información, es un recurso no dado; es decir, se trata de un insumo que debe someterse previamente a criterios de sistematicidad y consistencia, principalmente por medio de programas informáticos especializados, con la finalidad de comprobar o descartar una hipótesis. La potencialidad del dato se basa en su posibilidad de respaldar una hipótesis o conjetura y en las opciones de su tratamiento en términos de limpieza y sistematización.

Profundizar en la relación entre el usuario, la información y los datos hizo posible descubrir que más que hablar de un comportamiento informativo, como esta investigación planteó al inicio, prevalecen elementos que conducen a un comportamiento en el uso de datos, como lo son la concepción del dato; la identificación de la necesidad de datos; los obstáculos que se afrontan para trabajar con ellos; la delimitación, obtención y construcción del dato; el trabajo con los mismos; su tratamiento y comprensión, así como la interacción del dato con la hipótesis establecida y el producto final.

En este punto es posible exponer que el producto final representa la suma de saberes interdisciplinarios y que, al ser un producto que integra elementos de diseño, visualización e interacción con los datos en un entorno digital, los involucrados participan en una negociación entre ellos sobre lo que es o debería ser el producto y sus añadidos. En otras palabras, el producto final es un ensamble, una concurrencia de saberes interdisciplinarios que oscilan en procesos de búsqueda de información o de datos y la elaboración de una nota.

Emplear la TF en este estudio representó un reto como investigador, pues es un enfoque que exige una sensibilidad teórica y un distanciamiento o autorregulación con respecto a las preconcepciones de otras investigaciones, con la dificultad de enriquecer lo hallado. Es clara la postura de que el investigador es a su vez una herramienta de investigación con una carga subjetiva inherente a sus juicios, análisis, preferencias y omisiones. En mi caso, mis prejuicios de comunicólogo especializado en prensa, ofrece una visión más cercana a los participantes, pero a su vez se trata de una visión sesgada o parcial del universo de significados que ponen en juego en cada acción. A la distancia, considero que esta ventaja me permitió traducir estas prácticas a otra óptica, la de los estudios sobre comportamiento informativo lo que me permitió aportar, de forma mínima, a esta área de estudio con los preceptos de otra disciplina. No obstante, hubo de por medio una exigencia de distanciarse de estos presupuestos para ubicar puntos ciegos en el acopio e interpretación de los datos. Hubo, además, en el camino de esta investigación, diversos retos que se vinculan con la aproximación a los datos, por ejemplo, durante el análisis de los resultados se cuidó no desvincular los datos del contexto en el que se obtuvieron, para no desvirtuar sus significados.

El empleo de la TF permitió cumplir con el objetivo de analizar y caracterizar el comportamiento informativo de los profesionales que desarrollan periodismo de datos, así como sus prácticas, disposiciones y procesos cognitivos que ponen en juego durante la elaboración de una nota. En el caso de los procesos cognitivos, es necesario señalar que una limitación en este caso fue que el análisis se basa en lo dicho por los participantes, no en lo directamente observado durante la elaboración de una nota. A pesar de ello, la información obtenida permite observar la importancia de procesos cognitivos como la búsqueda de información o de datos, la resolución de problemas, el contar una historia, así como revisar y analizar, a lo largo de la elaboración de una nota. Así como establecer diferencias entre escenarios y participantes, lo cual hace evidente la diferenciación de roles en la forma de relacionarse con los datos.

Entre las condiciones del estudio se resaltan los diferentes niveles de *granularidad* de los datos, pues se parte de un conjunto de información obtenida de

las visitas y entrevistas a los involucrados, un proceso de codificación y recodificación a lo largo de todo el estudio, que obligo a crear, suprimir y fusionar códigos y, con ello, rehacer gráficas y ajustes en la interpretación y análisis de los datos. En este punto hay una forma de correlato entre el objeto de estudio y lo que se realizó en esta investigación, pues al igual que los profesionales involucrados en el desarrollo de periodismo de datos, en esta investigación hay etapas de acopio de información, sistematización, tratamiento, limpieza y visualización de los datos. Lo cual implica determinado nivel de empatía con los participantes. El uso de programa MAXQDA facilitó el acopió sistemático de las entrevistas, así como la elaboración de categorías y códigos. Permitió realizar cruces de traslapes entre procesos de desarrollo de una nota y procesos cognitivos. Se puede decir que este programa ayuda a organizar el trabajo del investigador, pero nunca sustituye su juicio.

El análisis fue un proceso constante y longitudinal en este estudio pues, desde la formación de códigos, su aplicación y agrupación en categorías, así como en la interpretación y recodificación, el análisis estuvo presente. Esto obligó muchas veces a reacomodar la narrativa y rehacer constantemente las gráficas, las estadísticas y principalmente la descripción de la información obtenida.

La metáfora sobre la distancia necesaria para ver un cuadro en su totalidad o para apreciar los detalles ayuda a describir el trabajo realizado durante la investigación. Si bien, al inicio se parte de la abundancia y saturación de los datos, conforme se avanza en la codificación y la categorización de estos es posible establecer relaciones a nivel global y particular con respecto a lo obtenido. Para la comprensión de los datos hay un tratamiento y sistematización de la información obtenida, una codificación que va dando forma a una estructura estructurante, es decir una estructura que se conforma de categorías y códigos que opera como un dispositivo de acopio, sistematización, agrupación, condensación y análisis.

El trabajo de análisis durante la codificación y contrastación de datos es semejante a un proceso artesanal, ya que para que el dato cualitativo sea útil debe ser descrito y sustentado en la evidencia. Esto supone una constante relectura de lo obtenido y lo interpretado. El *yo* investigador, en ese contexto, es cuestionado de

manera permanente en su papel de observador, herramienta de investigación, intérprete de datos y analista.

Sin duda el mayor reto de la investigación radicó en definir la forma de narrar y analizar la información obtenida. La revisión de manuales especializados en TF, no siempre aportó claridad al respecto, pues son muy generales y poco esclarecedoras. Hubo varios planteamientos intermedios que fueron descartados, pues fueron vistos como aproximaciones densas a la información empírica. Estos planteamientos, a pesar de ofrecer mayor detalle dificultaban la lectura del documento. Se valoró la importancia de darle mayor peso a una lógica narrativa secuencial de los procesos descritos, sobre un planteamiento basado únicamente en la estructura de categorías y códigos. Se revisaron trabajos similares y se observó que la narración o descripción de los resultados es vista como un reto, pues no hay una regla común dada la singularidad, la naturaleza de los datos y los propósitos de cada investigación. A pesar de ello, se encontraron similitudes en la forma de narrar, por ejemplo, la inclusión de los esquemas de categorías y códigos para presidir el análisis de cada apartado; la inclusión de segmentos de códigos intercalados con apartados descriptivos y analíticos; y la comparación entre escenarios en los casos en los que había suficiente evidencia. Así fue que la lógica de la narración de los resultados de esta investigación, paso de la densidad descriptiva a la densidad conceptual,

El papel de la perspectiva externa en el proceso de esta investigación ha sido crucial para su buen término. Las observaciones de quienes han leído las primeras aproximaciones y planteamientos dejan ver la importancia de la mirada externa, pues la inmersión que realiza el investigador en los datos en ocasiones se torna en un fanatismo por lo que dicen o dejan de decir los datos, y la mirada externa nos advierte sobre los puntos ciegos, o nos desmiente sobre alguna fijación o sobreinterpretación. La inmersión no siempre es un buen referente, si bien ayuda a construir, contextualizar e interpretar, se requiere de un espacio para alejarse o descontextualizarse para buscar interpretaciones alternas y de nuevo, recomenzar.

Entre las bondades del enfoque cualitativo empleado en este trabajo destaca la caracterización de las motivaciones o disposiciones presentes entre los

participantes, las cuales se ven reflejadas en sus prácticas. La noción de compartir información o experiencias con otros colegas, así como el colaborar en iniciativas de datos abiertos son acciones que ponen en práctica estas disposiciones incorporadas. Al respecto se considera pertinente utilizar para futuras investigaciones el concepto de *habitus*, planteado por Bourdieu (2016), pues sintetiza el papel de las disposiciones, motivaciones e intereses personales en las prácticas hechas y por hacer en un campo determinado, en este caso se habla de integrar estos elementos en el estudio del comportamiento informativo, como lo han hecho ya autores como Dervin (1999b).

La perspectiva cualitativa permitió además traspasar la mirada de la persona que desarrolla el periodismo de datos en un contexto determinado y comprender sus prácticas, exigencias e intereses que lo llevan a actuar de determinada forma bajo circunstancias específicas.

La aplicación de la TF en estudios del comportamiento informativo ayuda a clarificar los procesos que se ponen en juego al interactuar con la información, sin limitarse a la caracterización de las fuentes utilizadas, se analizan las motivaciones, los obstáculos, las exigencias y limitaciones que dan forma a una instantánea holística, pero sobre todo permite comprender la mirada o situación particular de un profesional en contexto determinado.

En el caso de esta investigación se suman nuevos elementos para analizar en un futuro la forma en que otros profesionistas interactúan directamente con los datos, ya sean economistas, analistas financieros, académicos, demógrafos, sociólogos, entre otros. El comportamiento en el uso de datos cobra mayor importancia hoy en día, cuando las Ciencias de la información se plantean retos de frente a los *Big data* y a la gestión de datos de investigación. Estudios recientes como el de Harrison y Exner (2019) en el que se estudia el comportamiento en la búsqueda de datos de estudiantes de economía y la forma en que buscan y utilizan conjuntos de datos, son ejemplo de esta veta reciente en nuestro campo de estudio. Comprender la forma en que las personas buscan, construyen, interpretan y les dan significado a los datos en distintas disciplinas abre una vasta agenda de trabajo para la investigación.

REFERENCIAS

ACRL (2012). Information literacy competency standards for journalism students and professionals. *College and Research Libraries News*. 73(5), pp. 274-285.

Allen, B. (1991). Cognitive research in information science: implications for design. *Annual Review of Information Science and Technology*, 26, pp. 3-37.

Antón Bravo, A. (2013). El Periodismo de datos y la Web Semántica. *CIC Cuadernos de Información y Comunicación*, 18, pp. 99-116.

Anwar, M. y Asghar, M. (2004). Information seeking behavior of Pakistani newspapers. *Pakistan Journal of Library and Information Science*, (10), pp. 55-76.

Anwar, M., Al-Ansari, H. y Abdullah A. (2004). Information behavior of Kuwaiti journalists. *Libri*, 54(4), pp. 228-236.

Aquesolo Vegas, J. (1996). De la documentación informativa al periodismo de precisión. *Documentación de las ciencias de la información*. Madrid: [Servicio de Publicaciones de la Universidad Complutense], Universidad Complutense de Madrid, 21(19), pp. 10-40.

Attfield, S. J. (2005). Information seeking, gathering and review: Journalism as a case study for the design of search and authoring systems. Tesis de doctorado. University of London.

Attfield, S. y Dowell, J. (2003a). Information seeking and use by newspaper journalists. *Journal of Documentation*, 59(2), pp. 187-204.

Attfield, S., Blandford, A. y Dowell, J. (2003b). Information seeking in the context of writing: A design psychology interpretation of the “problematic situation.” *Journal of Documentation*, 59(4), pp. 430-453.

Ávila, A. M. (s/a). Periodismo de datos: historia y momento actual. En *Manual de periodismo de datos iberoamericano.* Disponible en: http://manual.periodismodedatos.org/ana-maria-avila.php

Bárcenas Aboites, M. (2012). La infografía periodística o la visualización de la información. *Derecho a comunicar* (6) sep-dic., pp. 21-42.

Becker, P.H. (1993). Common pitfalls in published grounded theory research. *Qualitative Health Research*, 3(2). pp. 254-260.

Belkin, N. J., Oddy, R., y Brooks, H. (1982). ASK For Information Retrieval. *Journal of Documentation*, 38(1), pp. 61-71.

Bernard, H. R. y Ryan, G. W. (2010). *Analyzing qualitative data: systematic approaches*. Thousand Oaks Calif.: SAGE.

Blaxter, L., Hughes, C., y Tight, M. (2005). *Cómo se hace una investigación*. Barcelona España: Gedisa Editorial.

Bloom, B. S. (1990). *Taxonomía de los objetivos de la educación: la clasificación de las metas educacionales*: *manuales I y II.* Buenos Aires: El Ateneo.

Bounegru, L., Chambers, L. y Gray, J. (2011). *Data Journalism Handbook*. Disponible en: http://datajournalismhandbook.org

Bourdieu, P. (2003). *Sobre la televisión*. Barcelona: Anagrama.

Bourdieu, P. (2016). *Cuestiones de sociología*. Madrid, España: Akal.

Brahme, M. E., y Garcia, C. (2010). *The differences in information seeking behavior between distance and residential doctoral students*. Tesis de doctorado. Pepperdine University Graduate.

Brookes, B. (1975). *The fundamental problem of information science. In V. Horsnell, Informatics 2: proceedings of a conference held by the Aslib coordinate indexing group*. Oxford: Aslib. pp. 42-49.

Cairo, A. (s/a). Infografías y visualizaciones que simplifican el exceso o confunden datos. *Manual de periodismo de datos iberoamericano.* Disponible en: http://manual.periodismodedatos.org/alberto-cairo.php

Calderín, M. (2009). Las necesidades de información de los periodistas frente a la conversión de los medios tradicionales en cibermedios. *GT Ciberperiodismo y Comunicación Digital*. Disponible en: http://www.invecom.org/eventos/2009/pdf/calderin_m.pdf

Campbell, F. (1997). Journalistic construction of news: information gathering. *New Library World*, 98(2), pp. 60-71.

Carlsson, H., Hanell, F. y Lindh, K. (2013). Exploring multiple spaces and practices: a note on the use of ethnography in research in library and information studies.

Information Research, 18(3) paper C17. Disponible en: http://InformationR.net/ir/18-3/colis/paperC17.html

Cavanna, M. (2013). *Un giornalismo di precisione? Storia ed evoluzione del data journalism internazionale*. Tesis de doctorado. Universita degli studi di Catania. Prova Finale.

Chaparro Domínguez, M. Á. (2014). Nuevas formas informativas: el periodismo de datos y su enseñanza en el contexto universitario. *Historia y comunicación social*, 19, pp. 43-54.

Charlotte, W. (2000). Teaching online information retrieval to students of journalism. *Aslib Proceedings: New Information Perspectives*, 52(1), pp. 39-47.

Charmaz, K. (2006). *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. Londres; Thousand Oaks Calif.: Sage Publications.

Chaudhry, A. S. y Al-Sagheer, L. (2011). Information behavior of journalists: Analysis of critical incidents of information finding and use. *The International Information & Library Review,* 43(4), pp. 178-183.

Chinn, L. (2001). *The information seeking behavior and needs of journalists in context*. Tesis de maestría. University of North Carolina at Chapel Hill.

Creswell, J. (1998). *Qualitative inquiry and research design: choosing among five traditions*. Thousand Oaks: Sage.

Delgado Guallar, J. y Cornet, A. (2004). Fuentes de información digitales en los centros de documentación de prensa: las bases de datos comerciales, *El profesional de la información*. Barcelona: Swets, Zeitlinger Ibérica, 13(2), marzo-abril, pp. 107-117.

De Mey, M. (1982). The Cognitive paradigm: Cognitive science, a newly explored approach to the study of cognition applied in an analysis of science and scientific knowledge. Dordrecht: Reidel.

Dervin, B. (1983). An overview of sense-making research: Concepts, methods, and results to date. Sense-making packet, *Papers presented at the annual meeting of the International Communication Association*, Dallas, 17 de mayo.

Dervin, B. (1992). From the mind's eye of the user: the sense-making qualitative-quantitative methodology. En J.D. Glazier, y R.R. Powell (Eds.), *Qualitative Research in Information Management*. Englewood: Libraries Unlimited. pp. 61-84.

Dervin, B. (1998). Sense-making theory and practice: an overview of user interests in knowledge seeking and use. *Journal of Knowledge Management*, 2 (2), pp. 36-46.

Dervin, B. (1999a). Chaos, order and sense-making: A proposed theory for information design. En Jacobson, R. E. (2000). *Information design*. Cambridge, Mass: MIT Press.

Dervin, B. (1999b). On studying information seeking methodologically: the implications of connecting metatheory to method. *Information Processing and Management*. 35 (6), pp. 727-750.

Dinet, J. (2014). *Information Retrieval in Digital Environments*. Nueva York: John Wiley y Sons, Inc.

Egawhary, E. y O´Murchu, C. (2010). *Data Journalism Book*. Londres: City University London; Centre for Investigative Journalism.

Eisenberg, M. y Dirks, L. (2008). Taylor's Value-Added Model: Still Relevant After All These Years., *Conference paper* (4), pp. 1–13. Disponible en: https://www.ideals.illinois.edu/handle/2142/15081

Ellis, D. (1987). *The Derivation of a Behavioural Model for Information Retrieval System Design*. Tesis de doctorado. University of Sheffield, Dept. of Information Studies.

Ellis, D. (1993). Modeling the Information-Seeking Patterns of Academic Researchers: A Grounded Theory Approach. *The Library Quarterly*, 63(4), pp. 469-486.

Ellis, D. y Haugan, M. (1997). Modelling the information seeking patterns of engineers and research scientists in an industrial environment, *Journal of Documentation,* 53(4), pp. 384-403.

Ferreras Rodríguez, E. (2013). Aproximación teórica al perfil profesional del "Periodista de datos." *Revista ICONO14. Revista Científica de Comunicación y Tecnologías Emergentes*, 11(2), pp. 115-140.

Fisher, K., Erdelez, S. y McKechnie, L. (eds). (2005). *Theories of information behavior*. Medford N.J.: Published for the American Society for Information Science and Technology by Information Today.

Flores Vivar, J. y Salinas Aguilar, C. (2012). Sinergias en la construcción del Nuevo Periodismo derivadas del Data Journalism y el Transmedia Journalism. *III Congreso Internacional de Comunicación 3.0*. Salamanca.

Ford, N., y Mansourian, Y. (2006). The invisible web: An empirical study of “cognitive invisibility”. *Journal of Documentation*, *62*(5), pp. 584-596.

Forsythe, D. E. (1998). Using ethnography to investigate life scientists information needs. *Bulletin of the Medical Library Association*, 86(3), pp. 402-409.

Glaser, B. (1978). *Theoretical sensitivity: advances in the methodology of grounded theory*. Mill Valley Calif.: The Sociology Press.

Glaser, B. G. y Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine Pub. Co.

Glaser, B. y Strauss, A. (1968). *The discovery of grounded theory: strategies for qualitive research*. Weidenfeld and Nicolson.

González-Teruel, A. y Abad-García, M. F. (2012). Grounded theory for generating theory in the study of behavior. *Library and Information Science Research*, 34(1), pp. 31-36.

Grenfell, M. (2014). *Pierre Bourdieu: key concepts* (2nd ed.). Milton Park Abingdon Oxon; Nueva York: Routledge.

Guallar, J. y Cornet, A. (2004). Fuentes de información digitales en centros de documentación de prensa: las bases de datos comerciales. *El Profesional de la Información*, 13, (2), pp. 107-117.

Guthrie, J. T. (1988). Locating Information in Documents: Examination of a Cognitive Model. *Reading Research Quarterly*, 23(2), pp. 178-199.

Garrison, B., y Exner, N. (2019). Data Seeking Behavior of Economics Undergraduate Students: An Exploratory Study. *Reference & User Services Quarterly,* 58(2), pp, 103-113.

Hall, K. (1981). The economic nature of information. *The Information Society*, 1(2), pp, 143-166.

Hossain, M. A. y Shariful, I. (2012). Information seeking by print media journalist in Rajshahi, Bangladesh. *International Federation of Library Associations and institutions.* 38(4), pp. 283-288.

Howard, A. (2014). *The art and science of data-driven journalism*. Columbia University

Hsu, L.-L. (2014). *Online Information Seeking Behavior: Models of Information Source Selection and Information Seeker Satisfaction*. Tesis de doctorado. University of Colorado.

Huvila, I. (2009). Analytical information horizon maps. *Library and Information Science Research*, 31(1), pp. 18-28.

Ibarra, J. (s/a). Periodismo de datos: una práctica incipiente en México. *Manual de periodismo de datos iberoamericano.* Disponible en: http://manual.periodismodedatos.org/jesus-ibarra.php

Kelle, U. (1992). Theory Building in Qualitative Research and Computer Programs for the Management of Textual Data. *Sociological Research Online,* 2(2), pp. 1-13.

Kuckartz, U. y McWhertor, A. (2014). *Analysing Qualitative Data–But How? Qualitative Text Analysis: A Guide to Methods, Practice and Using Software*. Los Angeles: SAGE.

Kuhlthau, C. y Tama, S. L. (2001). Information search process of lawyers, a call for 'just for me' information services, *Journal of Documentation*, 57(1), pp. 25-43.

Kuhlthau, C. C. (1991). Inside the search process: seeking meaning from the users perspective. *Journal of the American Society for Information Science*, 42(5), pp. 361-371.

Leckie, G. J., Pettigrew, K. E. y Sylvain, C. (1996). Modeling the Information Seeking of Professionals: A General Model Derived from Research on Engineers, Health Care Professionals, and Lawyers. *The Library Quarterly*, 66(2), pp. 161-193.

López Yepes, A. (2006). *Los nuevos métodos de rastreo sobre textos completos en la documentación automatizada, aplicados a la investigación periodística*. Madrid: Universidad Complutense de Madrid. Disponible en: file://C:/educación/ ANDREA OEI/ DOCUMENTALES CUADERNOS/LOS N…/Alfonso.htm

López-García, X., Toural-Bran, C., Pereira-Farina, X. y Barbosa, S. (2009). Automatización de las bases de datos: Potencialidades de herramientas básicas para otro periodismo posible. *Profesional de la Información*, 18(3), pp. 308-315.

Marcos Recio, J. C. (1998). Una nueva concepción de la documentación en los medios electrónicos: retos y nuevas tareas profesionales. *Documentación de las ciencias de la información*. Madrid: [Servicio de Publicaciones de la Universidad Complutense], Universidad Complutense de Madrid, 21, pp. 113-130.

Maxwell, J. (1996). *Qualitative Research Design. An Interactive Approach*. Londres, Sage.

Mellon, C. A. (1986). Library Anxiety: A Grounded Theory and Its Development. *College & Research Libraries*, 76(3), pp. 276-282.

Meyer, P. (1973). *Precision journalism: A reporter's introduction to social science methods*. Bloomington: Indiana University Press.

Micó-Sanz, J., Masip, P. y García-Avilés, J. (2009). Periodistas que ejercen de documentalistas (¿y viceversa?). Nuevas relaciones entre la redacción y el archivo tras la digitalización de los medios. *El Profesional de la información*. 18(3), pp. 284-290.

Montesi, M. (2011). El comportamiento relacionado con la información en la literatura en lengua hispana. (Spanish). *Ibersid*, 5, 61–70.

Naumer, C. M. (2012). *A Frame-Based Theory of Information Behavior: A Grounded Theory Study.* Tesis de doctorado. University of Washington.

Nicholas, D., Williams, P., Cole, P. y Martin, H. (2000). The impact of the Internet on information seeking in the Media. *Aslib Proceedings*, 52, pp. 98-114.

Nicholas, D. y Martin, H. (1997). Assessing information needs: a case study of journalists. *Aslib Proceedings*, 49(2), pp. 43-52.

O'Reilly, K. (2012). *Key Concepts in Ethnography*. Londres: SAGE Publications Ltd.

Ortiz, V. (2007). *Los recursos de información en línea y la investigación periodística: estudio de caso.* Tesis de maestría en Bibliotecología. México: El Colegio de México.

Pace, S. (2003). Understanding the flow experiences of Web users. Tesis de doctorado. The Australian National University.

Parasie, S. y Dagiral, E. (2013). Data-driven journalism and the public good: "Computer-assisted-reporters" and "programmer-journalists" in Chicago. *New Media and Society*, 15(6), pp. 853-871.

Pettigrew, K. E., Raya, F. y Harry, B. (2001). Conceptual Frameworks in Information Behavior. *Annual Review of Information Science and Technology (ARIST)*, 35, pp. 43-78.

Pintado Navarro, R. (2013). El comportamiento informacional de los periodistas en la Región de Murcia. *Cuadernos de Gestión de Información*. 3, pp. 25-51.

Poteet, A. R. (2000). *Newspaper journalists information seeking behavior with online information sources*. Tesis de maestría. University of North Carolina.

Randall, W. S. y Mello, J. E. (2012). Grounded theory: an inductive method for supply chain research. *International Journal of Physical Distribution and Logistics Management*, 42(8), pp. 863-880.

Robson, A. y Robinson, L. (2013). Building on models of information behaviour: linking information seeking and communication. *Journal of Documentation*, 69(2), pp. 169-193.

Rodríguez Brito, A. y García Chico, L. (2013). Cartografiando el periodismo de datos Big Data y periodismo en el continente americano. Cinco casos de estudio. *Telos, Cuadernos de comunicación e innovación*. 95, pp. 57-65.

Hernández Salazar, P., Ibáñez Marmolejo, M., Valdez Angeles, G. Y. y Vilches Malagón, C. (2007). Análisis de modelos de comportamiento en la búsqueda de información. *Ciência da Informação*, 36(1), pp. 136-146.

Saúl, L. (2015) *Informe narrativo final: Fideicomisos. Las cajas negras de los gobiernos*. (mimeo).

Savolainen, R. (1993). The sense-making theory: Reviewing the interests of a user-centered approach to information seeking and use. *Information Processing & Management.* 29(1). pp.13-28

Savolainen, R. (2007). Information Behavior and Information Practice: Reviewing the “Umbrella Concepts” of Information-Seeking Studies. *The Library Quarterly*, 77(2), pp. 109-132.

Seldén, L. (2005). On Grounded Theory–with some malice. *Journal of Documentation*, 61(1), pp. 114-129.

Shah C (2014). *Collaborative information seeking.* Journal of the Association for Information Science and Technology, 65(2), pp. 215–236

Soneira, A. (2006). La "teoría fundamentada en los datos" (grounded theory) de Glaser y Strauss. En Irene Vasilachis (ed). *Estrategias de investigación cualitativa*. Barcelona: Gedisa. pp.153-175.

Sonnenwald, D. H. (1999). *Evolving Perspectives of Human Information Behavior: Contexts, Situations, Social Networks and Information Horizons*. Exploring the

Contexts of Information Behaviour: Proceedings of the Second International Conference on Research in Information Needs, Seeking and Use in Different Contexts, pp. 197-204. Disponible en: http://eprints.rclis.org/archive/00006960

Sonnenwald, D. H., Wildemuth, B. M. y Harmon, G. L. (2001). A Research Method to Investigate Information Seeking using the Concept of Information Horizons: An Example from a Study of Lower Socio-economic Students' Information Seeking Behavior. *Association of Library and Information Science Education* (ALISE), 2, pp. 65-86.

Taylor, R. S. (1982). Value-Added Processes in the Information Life Cycle. Journal of the American Society for Information Science, 33(5), pp. 341-46.

Taylor, R. (1986). *Value-added processes in information systems*. Norwood: Ablex.

Taylor, R. (1991). Information use environments. In B. Dervin (Ed.), *Progress in Communication Sciences*. Nueva Jersey: Ablex, pp. 217-225.

Todd, R. J. (1999). Back to our beginnings: information utilization, Bertram Brookes and the fundamental equation of information science. *Information Processing and Management*, 35, 6, pp. 851-870.

Tylor, J. (2014). An examination of how student journalists seek information and evaluate online sources during the newsgathering process. *New Media & Society*, 17(8), pp. 1277-1298.

Usher, N. (2016). *Interactive journalism : hackers, data, and code*. Urbana: University of Illinois Press.

Urquhart, C. (2013). *Grounded Theory for Qualitative Research: A Practical Guide*. Londres: Sage, 2013

Williams, P. y Nicholas, D. (1997). Journalists, news librarians and the Internet. *New Library World,* 98, pp. 11-37.

Wilson, T. D. (1980). On information science and the social sciences. *Social Science Information Studies*, 1(1), pp. 5-12.

Wilson, T. (1981). On user studies and information needs. *Journal of Documentation. Journal of Documentation*, 37(1), pp. 3-15.

Wilson, T. (1984). The cognitive approach to information-seeking behaviour and information use. *Social Science Information Studies*. 4, (2-3). pp. 197-204.

Wilson, T. D. (1990). “Object or participant? The information user in information research”, *Svensk Biblioteksforskning*, 3, pp. 5-16.

Wilson, T. (1997). Information behaviour: an interdisciplinary perspective. *Information Processing & Management.* 33 (4), pp. 551-572.

Wilson, T. (1999). Exploring models of information behaviour: the “uncertainty”project. *Information Processing & Management*. 35, (6). pp. 839-849.

Wilson, T. (2006). A re-examination of information seeking behaviour in the context of activity theory. *Information Research*, 11(4), 260. Disponible en: http://InformationR.net/ir/11-4/paper260.html

ANEXO 1 GUIÓN DE ENTREVISTA SEMIESTRUCTURADA

Objetivo: Reunir información sobre el contexto, la forma de trabajo de los participantes, su visión, motivaciones, flujos de trabajo, tareas, interacciones con la información; así como los elementos que ponen en juego durante el desarrollo de un producto noticioso; el significado de sus acciones; la influencia del contexto en sus acciones y los elementos que toma en cuenta para la toma de decisiones.

Antes de cada entrevista se solicitó a los participantes su autorización para registrar en audio la entrevista, de la misma forma se les platicó sobre la confidencialidad de la información que compartieran y el objetivo general de la investigación.

El comienzo de cada entrevista concentra las preguntas generales, principalmente lo relacionado con la trayectoria del participante, su formación y funciones del área en el que se desempeña. La idea de hacerlo de esta forma favorece el establecimiento de *rapport* con los participantes. No hay un orden secuencial para formular las preguntas, pues en muchos casos los tópicos que se querían plantear se fueron resolviendo sin la necesidad de formular la pregunta respectiva.

Principales tópicos de la entrevista

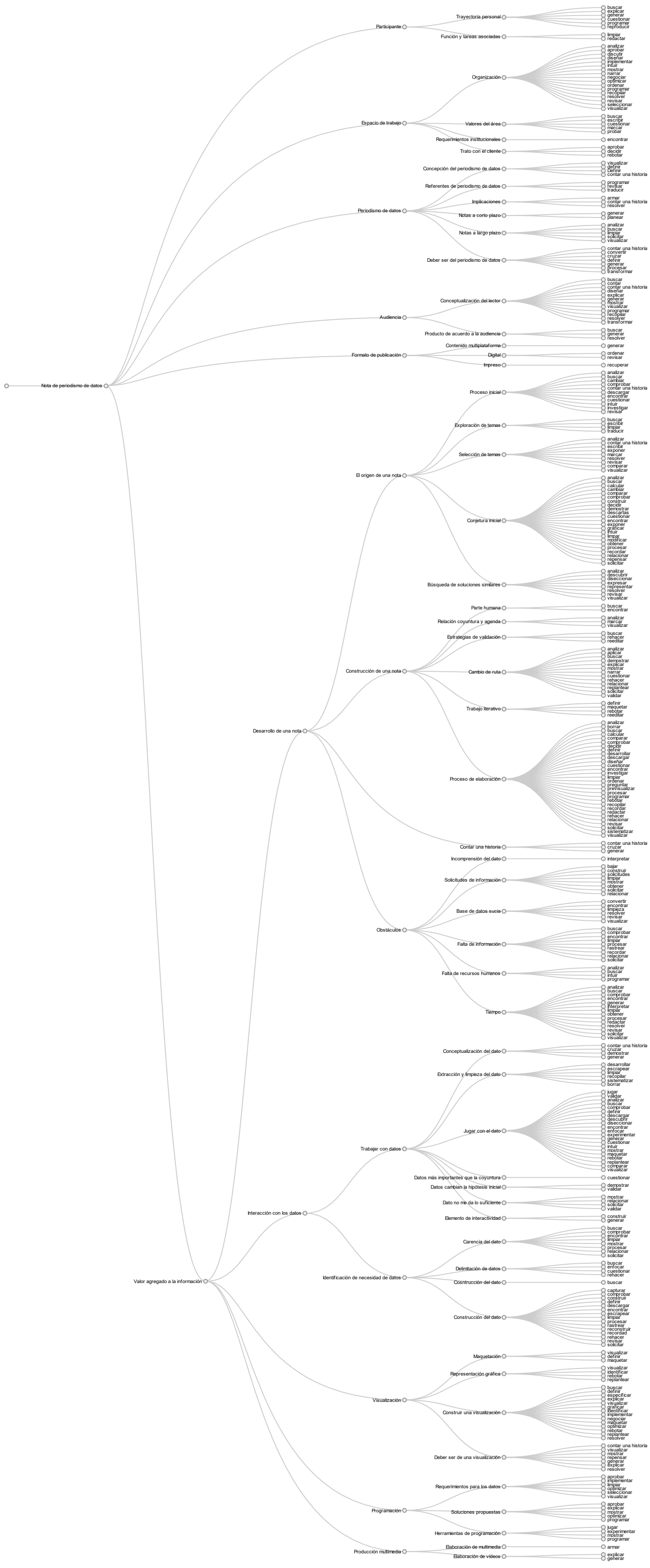
- Trayectoria personal
  -Formación
  -Forma en que se incorporan a las áreas
  -Tareas
  -Motivaciones

- Contexto
- (origen del área de trabajo)
- Objetivo del área

- Dinámica de trabajo
- Flujo de trabajo
- Condiciones y obstáculos

- Roles

- Reconstrucción del proceso de elaboración de un producto
  -Origen de un tema
  -Agenda y coyuntura
  -Desarrollo
  -Interacción con la información
  -Relación con otras áreas

- Periodismo de datos
  -Conceptualización
  -El periodismo de datos
  -Referentes de periodismo de datos

ANEXO 2 ESTRUCTURA DE CATEGORÍAS Y CÓDIGOS Y SU INTERRELACIÓN CON PROCESOS COGNITIVOS