



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
FACULTAD DE QUÍMICA

Representaciones visuales de espacios y multiversos químicos para la
divulgación de conocimiento científico y expresión artística

TESIS

que para obtener el título de:

INGENIERA QUÍMICA

presenta:

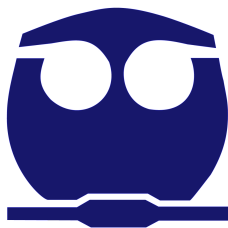
Daniela Gaytán Hernández

Asesor del tema:

Dr. José Luis Medina Franco

Supervisora técnica:

M en C. Ana Luisa Chávez Hernández



Ciudad Universitaria, CD. MX. 2024



Universidad Nacional
Autónoma de México



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.

Indice

1. Antecedentes.....	4
1.1. La química y el arte.....	4
1.2. Métodos computacionales para la creación de arte.....	6
1.3. Química computacional y quimioinformática.....	8
1.4. Reducción de dimensiones.....	9
1.4.1 Análisis de componentes principales (PCA).....	9
1.4.2 t-SNE.....	9
1.4.3 T-MAP.....	10
1.5. Representación molecular.....	10
1.6. Espacio químico y multiverso químico.....	10
2. Objetivos.....	14
2.1. General.....	14
2.2. Específicos.....	14
3. Novedad y significancia.....	15
4. Metodología.....	16
4.1. Bases de datos.....	16
4.2. Curado de bases de datos.....	19
4.3. Cálculo de descriptores moleculares.....	20
4.4. Generación de multiversos químicos.....	20
5. Resultados.....	21
6. Discusión de resultados.....	26
6.1. Ejemplos de espacios químicos y multiversos relacionados con el arte.....	26
6.2. Inteligencia artificial y arte digital.....	29
7. Conclusiones.....	31
8. Referencias.....	33
ANEXO.....	37

1. Antecedentes

1.1. La química y el arte

Históricamente, la relación entre la ciencia y el arte se ha dado desde que el hombre creó el arte ¹. El arte puede ser considerado como el conjunto de actividades y productos del ser humano con objetivos estéticos, éticos y de comunicación que impactan a los individuos o sociedades. Su impacto puede buscar la transmisión de ideas, necesidades, inquietudes o valores ².

La Química es una disciplina científica que ha tenido una relación simbiótica con el arte, y respectivamente, ha tenido influencia en su evolución. La Química es definida por la Real Academia Española como la “Ciencia que estudia la estructura, propiedades y transformaciones de la materia a partir de su composición atómica”. Algunos ejemplos donde la Química ha tenido contribuciones en el arte son el desarrollo y análisis de pigmentos para pinturas ^{3,4}. Es innegable que la invención y la disponibilidad de nuevos pigmentos químicos influyeron en el uso del color en el arte. La historia de la pintura ha estado influida por la disponibilidad o no de determinados pigmentos, y los descubrimientos científicos se han reflejado directamente en la paleta del artista REF. Desde que los pintores comenzaron a plasmar sus sueños y visiones en imágenes han recurrido a conocimientos y habilidades técnicas para proveerse de materiales ⁵. Los pigmentos no son “sencilla y únicamente color”, sino sustancias con propiedades y atributos específicos, sin olvidar entre éstos el precio. Los artistas impresionistas y sus seguidores, como Van Gogh, Matisse, Gauguin y Kandinsky, exploraron con gran energía las novedosas posibilidades de colores proporcionadas por los avances químicos. Su audiencia no solo quedó sorprendida por desafiar las convenciones artísticas, como la desviación de la paleta de colores “naturalista”, sino también por la presentación de colores nunca antes presentados en el lienzo: tonos de naranja sumamente brillante, púrpuras suaves, verdes vibrantes y nuevos ⁵.

Entre los materiales artísticos orgánicos se encuentran los aglutinantes y barnices empleados como capas protectoras. En general, la naturaleza (proteica o aceitosa) de los aglutinantes tradicionalmente se ha identificado por métodos de tinción selectivos usados en histoquímica. La caracterización de aglutinantes naturales en capas de pinturas a pesar de la cantidad de métodos existentes, aún es extremadamente difícil. Debido a los procesos de degradación inducidos por el envejecimiento natural y la contaminación, así como interferencias entre otros componentes presentes como los pigmentos ⁶.

La ciencia es la herramienta del arte, la que posibilita la materialización de las ideas y delimita el ideario de los artistas. Lo importante de la ciencia no es solo que haya servido para que la obra se pueda ejecutar; lo fundamental es que permitió que se pudiese imaginar, además, el conocimiento científico permite hacer una interpretación más profunda de las obras de arte. Conocer la ciencia del azul de Klein es clave para interpretar toda la complejidad de la obra. El azul de Klein es un ejemplo en el que la ciencia ha propiciado nuevos materiales, y a su vez nuevas técnicas para satisfacer las necesidades de los artistas ². Además el conocimiento científico da pie a una interpretación más profunda e inmersiva de las obras de arte.

En 2019, el año Intencionalidad de la Tabla periódica se realizó la publicación del libro “La Tabla de los elementos” el cual incluye un catálogo de las obras de 120 artistas plásticos que plasmaron las representaciones de todos los elementos de la tabla periódica ⁷. Cada artista asignado con un molde y el nombre de un elemento químico se adentro en explorar y descubrir maneras de comunicar el mensaje, la emoción específica que ese elemento les transmitía. Los materiales empleados abarcan desde los tradicionales como óleo, acrílico y encáustica, hasta cerámica, lámina de plata, varios metales, sal, carbón, grafito, tinta, caseína, plomo, ceniza, papel e incluso componentes reciclados de equipos informáticos. Estas obras se complementan con pequeñas narraciones o cédulas las cuales van más allá de la descripción física y material del elemento. Son relatos que nos hablan desde la historia, la literatura y el arte: historias, poemas y reflexiones acerca de los elementos y su manera de existir en este mundo, de entrelazarse con nosotros (y nosotros con ellos).

Otras aplicaciones de la Química y el arte son el uso de métodos químicos analíticos, entre ellos espectroscopia Raman y fluorescencia de rayos X, en la identificación de los materiales implicados en la elaboración de una pintura ⁸. Además, la Química ha participado en la creación, protección, restauración y visualización del arte. La presencia de pigmentos en pinturas de una sola capa y en manuscritos puede ser identificada por análisis de FRX (Fluorescencia de Rayos X) no destructivos. La fluorescencia de rayos X portátil es ideal para el examen de pinturas en los museos ya que permite identificar los elementos característicos de la mayoría de los pigmentos.

En 2018, un retrato hecho por un algoritmo de inteligencia artificial (IA) abrió nuevas interrogantes respecto a la creación y concepción de lo que consideramos arte. La obra Edmond de Belamy generó numerosas críticas que cuestionan la capacidad creativa (o imitativa) de los modelos de IA ⁹.

1.2. Métodos computacionales para la creación de arte

El proceso de producción de obras artísticas involucra varios factores, aunque se destacan aspectos esenciales. A lo largo de la historia, los artistas han utilizado la tecnología disponible para crear obras, a veces innovando o aprovechando herramientas y materiales. Por ejemplo, las esculturas antiguas se tallaban en mármol con cincelos. En la actualidad, las herramientas digitales ofrecen posibilidades vanguardistas ¹⁰.

Una obra de arte nos da la oportunidad de expresar un concepto que se puede concretar bidimensional o tridimensionalmente, creando un objeto artístico que es perceptible por medio de los sentidos de la vista (visión) y del tacto. De la misma manera, hay manifestaciones artísticas que se pueden percibir por medio de los otros sentidos el oído (audición), el olfato y el gusto ¹⁰.

Las obras de arte generadas por computadora involucran una amplia gama de herramientas de IA, así como paradigmas de interacción humana de IA. Algunos utilizan algoritmos innovadores e interactivos para la creación y curado de obras. Por otro lado, se han creado plataformas para que artistas y profesionales utilicen modelos de IA, como RunwayML, GANPaint y DeepAngel. Además de los GAN's, se han explorado muchos otros algoritmos generativos visuales, como la transferencia neuronal de estilo, la Estética Computacional, Fractal Flame, la evolución adversarial potenciada por el aprendizaje profundo y los métodos híbridos ¹¹. Nos encontramos inmersos en un entorno digital y tecnológico en nuestro día a día, donde las innovadoras tecnologías de comunicación e información (TIC's) se hacen presentes en el ámbito artístico. Como resultado, están emergiendo nuevas formas de expresión fundamentadas en las TIC, lo que amplía las formas de crear arte. De esta evolución surge el concepto de arte digital como una categoría entre las artes contemporáneas ¹⁰. Algunos ejemplos del uso de IA en el Arte son [DALL-E Artificial Intelligence](#) ¹¹, [Midjourney](#), herramientas que utilizan algoritmos de IA para generar nuevas obras de arte o imágenes desde cero a partir de una base de datos de obras de arte o imágenes de referencia. DALL-E puede crear imágenes y arte originales, realistas a partir de una descripción textual. Combinando conceptos, atributos y estilos. Un ejemplo es utilizando la plataforma DALL-E, capturando como descripción "*Científico generando una idea*" el cual

nos da como resultado la obra de la Figura 1. Otro ejemplo es empleando la plataforma Canva, usando como palabra clave “Multiverso técnica acuarela” consiguiendo la obra digital de la Figura 2.

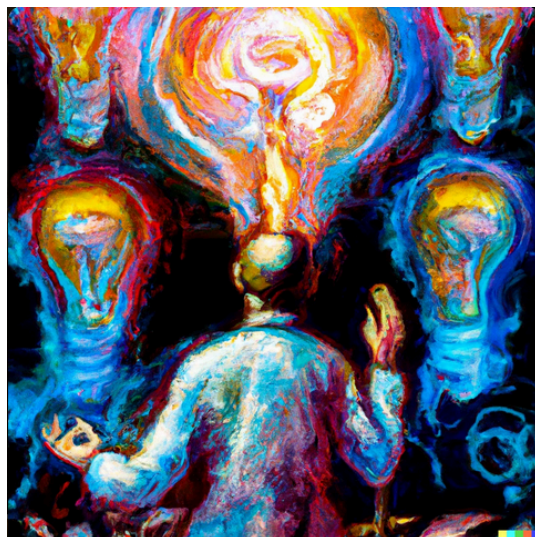


Figura 1. Obra de arte digital generada por la plataforma DALL-E.



Figura 2. Obras digitales generadas por la plataforma gratuita Canva (<https://www.canva.com/>) usando como palabras clave “multiverso” y “acuarela”.

Otro ejemplo que utiliza los algoritmos de IA, la química y el arte son las llamadas “moléculas musicales” (en la referencia original, “music molecules”)¹² que demuestra que la música es un medio de almacenamiento de información altamente dimensional que puede utilizarse para codificar la estructura molecular. El método resultante permite escuchar una estructura molecular como una composición musical, en la que la tonalidad de la música se basa en las propiedades moleculares y la melodía en la

disposición de átomos y enlaces. Transformar moléculas en música proporciona una rica colección de datos musicales que pueden utilizarse para entrenar modelos de generación de música, como se ha visto con MusicVAE. Si fuese posible representar las moléculas en forma de música, se abrirán posibilidades para transformar la experiencia sensorial de lo visual a lo auditivo. Otras formas de expresión artística digital y que también se encuentran en la naturaleza es el arte fractal.

1.3. Química computacional y quimioinformática

La Química Computacional es una disciplina de la Química que comprende todos los aspectos de la investigación en Química que se benefician de la aplicación de las computadoras ¹³. Los métodos computacionales han ayudado a analizar datos de manera más eficiente, elegir colecciones de compuestos para seleccionar moléculas para evaluación experimental, desarrollar hipótesis que ayuden a comprender el mecanismo de acción de fármacos y a desarrollar nuevas estructuras químicas. La Química Computacional ha tenido un gran impacto en la sociedad y cada vez hay un mayor número de aplicaciones que se ven reflejadas en un incremento de artículos y publicaciones científicas ¹⁴.

La quimioinformática es una disciplina esencial en el descubrimiento de fármacos que fusiona la tecnología, la gestión de la información y la informática. Su enfoque radica en transformar datos en información y, a su vez, convertir esta información en conocimiento valioso. Esto permite tomar decisiones más informadas y rápidas en la identificación y organización de compuestos farmacológicos. Engloba la creación, organización, análisis y aplicación de información química, empleando métodos informáticos para resolver desafíos químicos. Además, involucra la representación de moléculas como objetos en un espacio químico multidimensional. En definitiva, la quimioinformática agiliza el proceso de descubrimiento de medicamentos al unir tecnología, gestión y conocimiento químico ^{15,16}. La quimioinformática es una aplicación específica del aprendizaje automático, en la que se modela la estructura y propiedad de compuestos químicos. A diferencia de otras áreas, la complejidad química y la naturaleza de los compuestos requieren el uso de teoría de grafos y nuevos métodos de aprendizaje automático. En lugar de vectores numéricos, se emplean grafos (e hipergrafos) para representar estos objetos químicos, lo que exige enfoques capaces de tratar datos discretos estructurados.

Dentro de las técnicas empleadas se encuentran los algoritmos de IA que tienen gran utilidad para extraer información de un conjunto de datos y aportar conocimientos prácticos o útiles, compuestos por los algoritmos de aprendizaje de máquina (del inglés, *machine learning*,) y aprendizaje profundo (del inglés, *deep learning*)^{17,18}. Los principios y métodos empleados en la quimioinformática permiten gestionar, visualizar y analizar sistemáticamente la información química y facilita el análisis y visualización de manera eficiente de miles o millones de datos. Los métodos de Quimioinformática permiten administrar y organizar información, visualizar el espacio químico, realizar minería de datos y establecer relaciones matemáticas entre estructura y actividad. Un ejemplo clásico son los métodos para cuantificar relaciones estructura-actividad biológica y relaciones estructura-propiedad en general¹⁸.

1.4. Reducción de dimensiones

1.4.1 Análisis de componentes principales (PCA)

El análisis de componentes principales es un método estadístico que reduce las dimensiones de un espacio muestral con p dimensiones, donde p es el número de variables a medir en una muestra. Por muestra nos referimos a una molécula¹⁹. Cada componente principal (PC) es un vector que sigue la dirección en el que las observaciones varían más. Así el PC1 recupera la mayor varianza y el PC2 menor varianza que PC1 y mayor que el PC3. La condición es que estos vectores no están correlacionados entre sí, es decir que sean perpendiculares. Los PC se obtienen por una combinación lineal de sus variables originales.

1.4.2 t-SNE

Los algoritmos de reducción de la dimensionalidad tienen un valor incalculable para la inspección cualitativa de datos de alta dimensión y se utilizan ampliamente en todas las disciplinas científicas. En términos generales, estos algoritmos transforman una entrada de alta dimensión en una incrustación confiable de menor dimensión. El objetivo de esta incrustación es preservar las similitudes entre los puntos, que a menudo se miden por las distancias en los espacios correspondientes²⁰. Los métodos de reducción de la dimensionalidad y aprendizaje múltiple, como la incrustación estocástica de vecinos

distribuida en t (t-SNE), se utilizan con frecuencia para mapear datos de alta dimensión en un espacio bidimensional con el fin de visualizarlos y explorarlos.

1.4.3 T-MAP

El algoritmo, denominado TMAP (*Tree MAP*) genera y distribuye visualizaciones intuitivas de grandes conjuntos de datos del orden de hasta 10^7 con dimensionalidad arbitraria en un árbol. Este trazado basado en árboles facilita la inspección visual de los datos con una alta resolución al visualizar explícitamente la distancia más cercana entre clusters y la estructura detallada de los clusters a través de ramas y subramas ²¹.

1.5. Representación molecular

Tradicionalmente, las moléculas se representan como diagramas de estructura con enlaces y átomos, sin embargo, para el procesamiento computacional de estructuras químicas en quimioinformática se necesitan otras representaciones. Una representación molecular es cualquier forma de codificar un compuesto químico, dentro de esta se encuentran las representaciones lineales y los descriptores moleculares. Una representación lineal codifica la conectividad entre pares de átomos de una molécula²². Algunos ejemplos de representaciones lineales son *Simplified Molecular Input Line Entry System* (SMILES) ²³, *SMILES arbitrary target specification* (SMARTS) y InChIKey. Por otra parte, un descriptor molecular codifica características fisicoquímicas y estructurales ²⁴. Los descriptores moleculares son la representación molecular más utilizada en las visualizaciones de espacio químico y más reciente, el multiverso químico. Ambos conceptos se describen en la siguiente sección.

1.6. Espacio químico y multiverso químico

El espacio químico o espacio molecular puede definirse como un arreglo multidimensional de estructuras químicas, en donde los renglones corresponden a estructuras químicas y las columnas o dimensiones están definidos por un número de descriptores moleculares ²⁵. El espacio químico hace referencia a todas las moléculas posibles y a los espacios conceptuales multidimensionales que representan sus propiedades estructurales y funcionales ²⁶. En otras palabras, el espacio químico, es

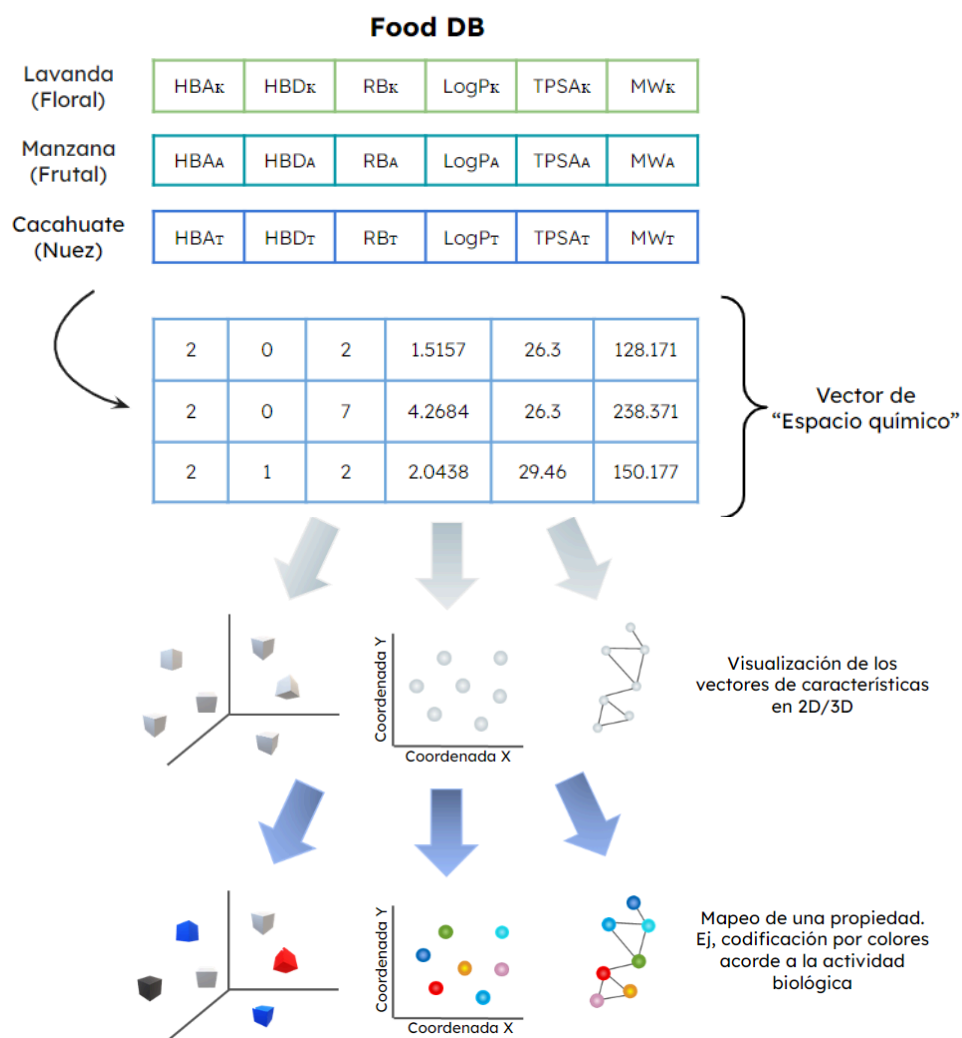
una base de datos molecular en donde cada molécula está definida por un número “n” de descriptores. El espacio químico, además de ser la base conceptual de la disciplina de la Quimioinformática, representa una forma sistemática de organizar, manejar, almacenar y transmitir información química. Un aspecto central para establecer un espacio químico es el tipo y número de descriptores que se emplean para caracterizar a las moléculas. Debido a que cada grupo de “n” descriptores puede dar origen a diferentes espacios químicos, un “multiverso químico” reúne a un grupo de espacios químicos, cada uno definido por descriptores diferentes ²⁷.

Multiverso “una colección hipotética de universos observables potencialmente diversos, cada uno de los cuales comprendería todo lo que es experimentalmente accesible por una comunidad conectada de observadores” ²⁸.

El espacio químico es un concepto esencial de la quimioinformática y tiene numerosas aplicaciones prácticas en el descubrimiento de fármacos, síntesis orgánica, química en alimentos, ciencia de material entre otras áreas de la química. Las aplicaciones del espacio químico incluyen la selección y diseño, la diversidad y el análisis de relaciones estructura-propiedad (actividad)²⁵. En general, la visualización de datos en gráficos y figuras de dos dimensiones es una forma común de analizar datos para generar y transmitir información. En este contexto, la representación visual del espacio y multiversos químicos se utiliza ampliamente en el análisis de bases de datos moleculares en la industria farmacéutica y, más recientemente, en la industria de los alimentos ²⁶.

Los espacios químicos y los multiversos químicos, como muchos otros tipos de análisis, se analizan con frecuencia mediante técnicas de visualización de datos (Figura 4). De hecho, la visualización de datos se utiliza ampliamente en la ciencia y en otras áreas para resumir y comunicar datos de forma eficaz con el fin de producir información y, en última instancia, conocimiento. Se han publicado extensas revisiones sobre la visualización de espacios químicos ²⁹⁻³⁰.

A



B

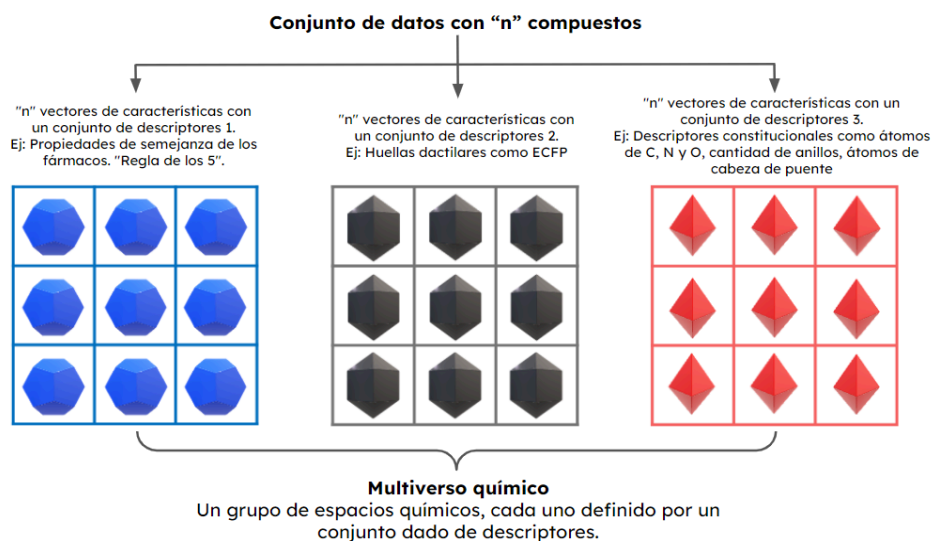


Figura 4. Concepto esquemático de **A)** Espacio químico y su representación visual en dimensiones reducidas. **B)** Representación esquemática de multiverso químico para un hipotético conjunto de datos de "n" compuestos: descriptores de diferente diseño (propiedades continuas, huellas moleculares, descriptores estructurales, etc.) pueden dar lugar a espacios químicos alternativos para el mismo conjunto de datos.

En este proyecto se utilizarán herramientas establecidas de visualización del espacio químico no solamente para graficar datos y transmitir información química, sino también como mecanismo de expresión artística que esté motivada por herramientas de la quimioinformática. Se busca que este proyecto contribuya a la divulgación de la Química a través del Arte y fomentar el uso continuo de la Química no solamente como un disciplina que sirva para el Arte sino que sea una forma de expresión artística por sí misma.

2. Objetivos

2.1. General

Generar representaciones visuales del espacio y multiversos químicos como mecanismos de divulgación de información química y expresión artística (“*Chemical-Art-Gallery*”).

2.2. Específicos

- I. Reunir y curar o estandarizar los compuestos de la base de datos de compuestos químicos derivados de alimentos (FoodBD) ³¹.
- II. Generar espacios químicos para cada grupo de moléculas mediante el cálculo de descriptores moleculares de diferente tipo (e.g., crear multiversos químicos para cada conjunto de moléculas).
- III. Aplicar diferentes métodos de visualización de datos multidimensionales a los espacios y multiversos químicos generados (TSNE, TMAP, PCA y SOM).
- IV. Automatizar el proceso de visualización de datos químicos utilizando protocolos en software libre como Colab. De esta manera desarrollar un “ChemArt Generator”.
- V. Generar la primera versión de una galería digital de visualizaciones de espacios y multiversos químicos: Museo del Espacio Químico y/o Galería del Multiverso Químico.

3. Novedad y significancia

El proyecto es novedoso porque representa la primera aplicación de la quimioinformática para generar representaciones visuales del espacio químico y los multiversos químicos no solamente para transmitir información científica, sino también artística y expresión propia del autor. El proyecto es relevante porque servirá como mecanismo de divulgación de la ciencia, acercando a la química a la población general que gusta del arte, especialmente, la pintura (por ejemplo, nuevas generaciones de niñas y niños y la juventud, entre otros sectores de la población). Se presenta una idea general para crear representaciones visuales del espacio químico y múltiples universos químicos como una forma de comunicación en el campo de la química, con la finalidad de generar nuevas experiencias y expresiones artísticas. Para ejemplificar esta propuesta, hemos creado visualizaciones del espacio químico de cuatro categorías de sabores utilizando una amplia base de datos pública de sustancias químicas alimentarias conocida como FooDB ³¹. Empleamos diversos descriptores y huellas moleculares para llevar a cabo esta tarea. Las cuatro categorías de sabores se describen en detalle en la sección Métodos. Este concepto también puede potenciar la intersección entre el arte y la quimioinformática, y tiene el potencial de extenderse a otras disciplinas vinculadas a la información, como la bioinformática.

4. Metodología

4.1. Bases de datos

En este caso, utilicé compuestos químicos alimenticios para generar representaciones visuales del espacio químico como obras de arte digital. Las características de los alimentos, como sus sabores, tonalidades, consistencias y aromas, se asocian generalmente a los grandes placeres de la vida y, por ello, han sido fuente de inspiración en el mundo del arte. No obstante, no se ha explorado una perspectiva a nivel estructural de las moléculas hasta el momento. La versión actual de FooDB³¹ contiene 70.477 compuestos, y tras la normalización o curado del conjunto de datos (descrita en la sección 4.2) cuenta con 52.856 moléculas. Esta brinda información sobre macronutrientes y micronutrientes, incluidos muchos de los componentes que dan a los alimentos su sabor, color, gusto, textura y aroma. Cada entrada química de la base de datos FooDB contiene más de 100 campos de datos independientes que abarcan información detallada sobre composición, bioquímica y fisiología (obtenida de la bibliografía). Esto incluye datos sobre la nomenclatura del compuesto, su descripción, información sobre su estructura, categoría química, sus datos físicoquímicos, su fuente(s) alimentaria(s), su color, su sabor, su efecto fisiológico, presuntos efectos sobre la salud (a partir de estudios publicados) y concentraciones en diversos alimentos ³¹. De la base de datos FoodDB, se identificaron 4964 número de saborizantes naturales derivados de compuestos alimenticios de la base de datos FoodDB, se identificaron 27 categorías de sabores y el número de compuestos que contiene cada categoría. La Figura 5 resume las 7 categorías con mayor número de compuestos.

Categorías de Sabores

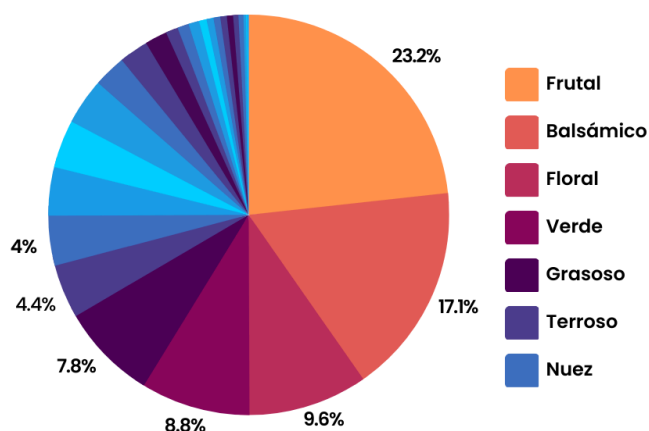


Figura 5. Distribución de los porcentajes de compuestos obtenidos para algunas de las categorías de sabores de la base de datos Food DB.

Para obtener las categorías de sabores con sus compuestos asociados:

1. Se recopilaron tres bases de datos diferentes (La base de datos Food DB curada previamente, Flavor DB y Compounds-Flavor DB)
2. Se seleccionaron las columnas necesarias para lograr realizar la intersección entre las 3 bases de datos.
 - a. Food DB curada: ID, NAME, SMILES
 - b. Flavor DB: ID, NAME, FLAVOR_GROUP
 - c. Compounds-Flavor: FLAVOR_ID, COMPOUND_ID
3. Se obtuvo una base de datos “unificada” con la siguiente información: [FLAVOR_ID, NAME_FLAVOR, FLAVOR_GROUP, COMPOUND_ID, NAME_FOOD_DB, NEW_SMILES] Por lo tanto fue posible saber los compuestos asociados con cada una de las categorías de sabor existentes en la base de datos “FLAVOR DB”
4. Posteriormente partiendo de las veintisiete categorías de sabores, se definieron cuatro nuevas categorías de sabores: 1) Sabores de la tierra, 2) Matices de vino, 3) Contraste de picante y graso, 4) Remedios naturales. La Tabla 1 muestra el número de compuestos en cada una de las nuevas categorías. La categoría 1. *Sabores de la tierra*, está conformada por los sabores terrosos, herbáceos y verdes. Los sabores de *matices de vino* se componen de sabores afrutados y florales. La categoría de 3. *Contraste entre picante* y graso se compone de los

sabores picantes y grasos. Por último la categoría 4. *Remedios naturales* se integran los sabores balsámicos, medicinales y químicos, que son sabores característicos que se encuentran en ungüentos, alcohol y jarabes. La Figura 6 muestra el número de compuestos únicos y superpuestos en cada una de las nuevas categorías.

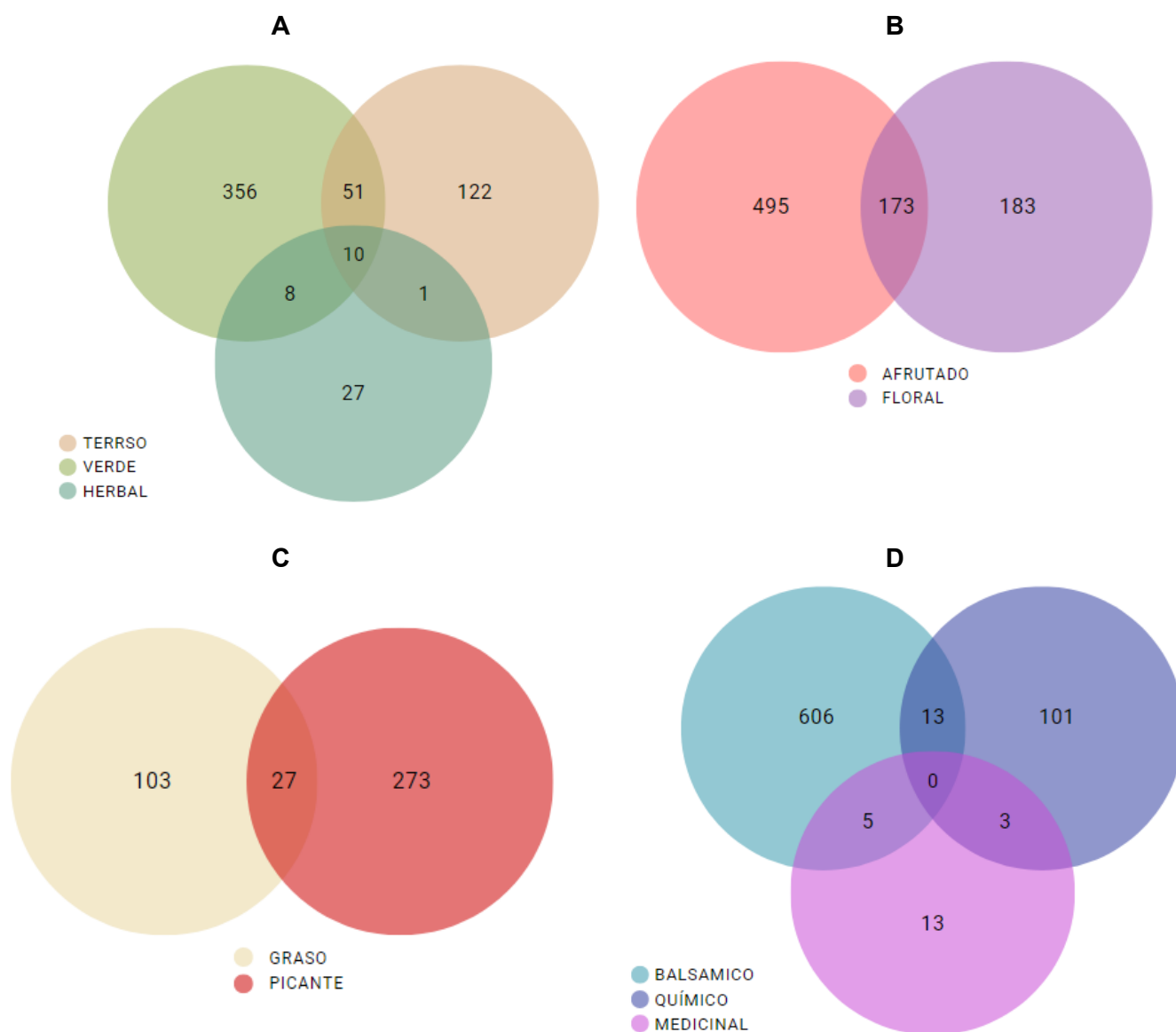


Figura 6. Estructuras únicas y superpuestas de las cuatro categorías de sabores propuestas de FoodDB³¹. A) Sabores de la tierra, B) Matices de vino, C) Contraste entre picante y graso, D) Remedios naturales.

Tabla 1. Número de compuestos de cada categoría de sabor 1) Sabores de la tierra, 2) Matices de vino, 3) Contraste de picante y graso, 4) Remedios naturales.

Notas de sabores	Categoría de sabor	Sabores ejemplares	Número de compuestos
Terroso	Sabores de la tierra	Hongos (Terroso) Ebúrneas (Verde) Clavo (Herbaceo)	655
Verde			
Herbaceo			
Afrutado	Matices de vino	Plátano, piña (Afrutado) Lavanda, Rosas (Floral)	1024
Floral			
Graso	Contraste entre picante y graso	Queso, mantequilla (Graso) Cumarina (Picante)	430
Picante			
Balsámico	Remedios naturales	Vainilla, Canela (Balsámico) Alcohol (Químico) Chavicol (Medicinal)	762
Químico			
Medicinal			

4.2. Curado de bases de datos

Se seleccionaron cadenas SMILES²³ sin información estereoquímica como representación molecular de los compuestos. En este trabajo no se tuvo en cuenta la información estereoquímica porque no todos los compuestos de los tres conjuntos de datos contienen una estereoquímica definida. Todo el proceso de preparación se realizó con el kit de herramientas quimioinformáticas de código abierto RDKit³² (<http://www.rdkit.org>), versión 2019.09.1 y las funciones Standardizer, LargestFragmentChoser, Uncharger, Reionizer y TautomerCanonicalizer implementadas en la herramienta de validación y estandarización de moléculas MolVS³³. Los compuestos se estandarizaron y los que constaban de múltiples componentes se dividieron y se retuvo el componente más grande. Los compuestos formados por cualquier elemento distinto de H, B, C, N, O, F, Si, P, S, Cl, Se, Br e I, así como los compuestos con errores de valencia, se eliminaron del conjunto de datos. Los compuestos restantes se neutralizaron y reionizaron para generar posteriormente un tautómero canónico. También se eliminaron de cada base de datos.

4.3. Cálculo de descriptores moleculares

Para cada molécula en las bases de datos reportadas en la Tabla 1 se calcularon descriptores moleculares como propiedades fisicoquímicas y huellas digitales moleculares utilizando el lenguaje de programación de Python y la biblioteca RDKit. Las propiedades fisicoquímicas de que se calcularon fueron átomos aceptores de enlace de hidrógeno (HBA), átomos donadores de enlace de hidrógeno (HBD), enlaces rotables (RB), coeficiente de reparto octanol/agua (logP), área topológica superficial (TPSA) y peso molecular (MW). Las huellas digitales moleculares que se calcularon fueron de conectividad extendida, ECFP, (del inglés *Extended Connectivity Fingerprint*) con 1024-bits y otra de diccionario (*MACCS keys*) con 166-bits ³⁴. Para ECFP ³⁵ se utilizaron las variantes de diámetro 4 (ECFP4) y diámetro 6 (ECFP6). Se puede utilizar prácticamente cualquier otro descriptor, como se comenta en la sección de discusión de resultados.

4.4. Generación de multiversos químicos

Para cada espacio químico, se generaron gráficos bidimensionales mediante métodos de reducción de variables como PCA, t-SNE, tMAPs y SOMs. PCA, t-SNE genera gráficos que organizan los compuestos, los compuestos similares forman grupos y los compuestos distintos se distancian entre sí. PCA es una técnica de reducción de la dimensionalidad lineal para transformar datos con muchas dimensiones (es decir, descriptores) en un espacio dimensional más bajo y mantener las diferentes relaciones entre los puntos de datos tanto como sea posible ¹⁹. El PCA se generó a partir de seis descriptores completos de moléculas (MW, HB, HBA, SlogP, TPSA y RB). Los TMAP permiten visualizar muchos compuestos químicos a través de la distancia entre clusters y la estructura detallada de éstos a través de ramas y subramas. El hashing sensible local permite agrupar jerárquicamente cada compuesto según subestructuras comunes utilizando huellas moleculares. Después, cada compuesto químico se codificó mediante el algoritmo MinHash. Para generar las visualizaciones con tMAPS el número de vecinos más próximos se fijó en $k = 50$, y el factor utilizado por el algoritmo de consulta aumentada, $kc = 10^{21}$.

5. Resultados

Las Figuras 7-10 muestran ejemplos de las “Galerías de arte” o “Cuadros Digitales” compuestas por las visualizaciones del espacio químico de diferentes categorías de compuestos químicos alimenticios. Las representaciones del espacio químico se generaron con t-SNE (Figuras 7 y Figura 8), PCA (Figura 9) y TMAPs (Figura 10). En la descripción de cada imagen se presenta información básica de la “técnica” (método de visualización, alusivo a las técnicas utilizadas en los cuadros), descriptores y compuestos químicos (que sería información significativa para que una persona orientada a la química entendiera los datos presentados). Cada representación visual del espacio químico u obra de Arte incluye un "Título" que recuerda el nombre de la obra de arte o pintura digital.

A**TÍTULO: *Mindful taste I***

Técnica: t-SNE

Descriptores: MW, HB, HBA, SlogP, TPSA, and RB.

Compuestos químicos: Flavor food-chemical compounds.

Categorías de sabor: Verdes, Terrosos y Herbáceo

B**TÍTULO: *Mindful taste II***

Técnica: t-SNE

Descriptores: MW, HB, HBA, SlogP, TPSA, and RB.

Compuestos químicos: Flavor food-chemical compounds.

Categorías de sabor: Afrutado y Floral

C**TÍTULO: *Mindful taste III***

Técnica: t-SNE

Descriptores: MW, HB, HBA, SlogP, TPSA, and RB.

Compuestos químicos: Flavor food-chemical compounds.

Categorías de sabor: Graso y Picante

D**TÍTULO: *Mindful taste IV***

Técnica: t-SNE

Descriptores: MW, HB, HBA, SlogP, TPSA, and RB.

Compuestos químicos: Flavor food-chemical compounds.

Categorías de sabor: Balsámico, Químico, y Medicinal

Figura 7. Cuatro categorías de sabor y todos los compuestos de FooDB. Las categorías de sabor son **A)** Sabores de la Tierra (655 compuestos), **B)** Matices de vino (1024 compuestos), **C)** Contraste entre picante y graso (430 compuestos), y **D)** Remedios Naturales (762 compuestos). El sentido que se quiso lograr con estas representaciones de la Figura 7 es la alusión o comparación con una imagen de una resonancia magnética funcional del cerebro y la identificación que tienen las áreas cerebrales que se involucran en algún proceso cognitivo, motor o perceptual del cerebro.

A**TÍTULO: Umami**

Técnica: t-SNE

Descriptores: MW, HB, HBA, SlogP, TPSA, and RB.
 Compuestos químicos: Flavor food-chemical compounds.

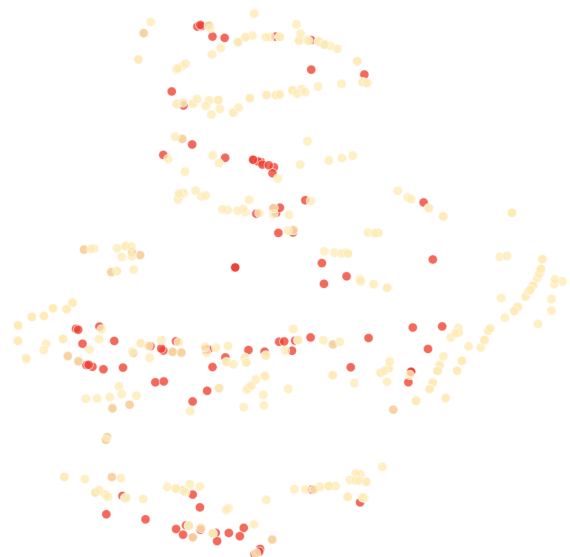
Categorías de sabor: Verdes, Terrosos y Herbáceo

B**TÍTULO: Sweet**

Técnica: t-SNE

Descriptores: MW, HB, HBA, SlogP, TPSA, and RB.
 Compuestos químicos: Flavor food-chemical compounds.

Categorías de sabor: Afrutado y Floral

C**TÍTULO: Sour**

Técnica: t-SNE

Descriptores: MW, HB, HBA, SlogP, TPSA, and RB.
 Compuestos químicos: Flavor food-chemical compounds.

Categorías de sabor: Graso y Picante.

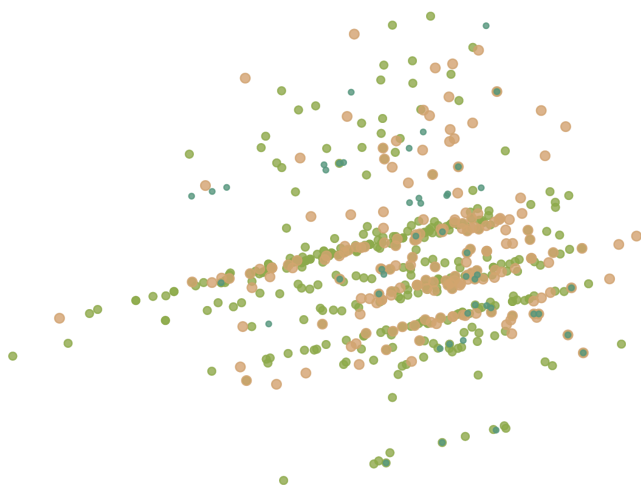
D**TÍTULO: Bitter**

Técnica: t-SNE

Descriptores: MW, HB, HBA, SlogP, TPSA, and RB.
 Compuestos químicos: Flavor food-chemical compounds.

Categorías de sabor: Balsámico, Químico, y Medicinal.

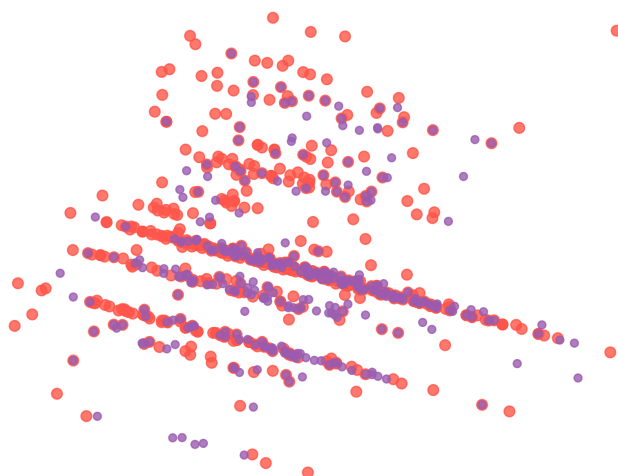
Figura 8. Cuatro categorías de sabor. **A)** Sabores de la Tierra (655 compuestos), **B)** Matices de vino (1024 compuestos), **C)** Contraste entre picante y graso (430 compuestos), y **D)** Remedios Naturales (762 compuestos). Para las representaciones de la Figura 8 poseen un nombre de los sabores básicos que percibe el ser humano, asignados acorde a las categorías de sabor de cada una de las visualizaciones, por ejemplo, la figura que tiene como título “Sweet” contiene categorías de sabor afrutados y florales, los cuales pueden ser asociados con sabores dulces.

A**TÍTULO: PANTE**

Técnica: PCA

Descriptores: MW, HB, HBA, SlogP, TPSA, and RB.
 Compuestos químicos: Flavor food-chemical compounds.

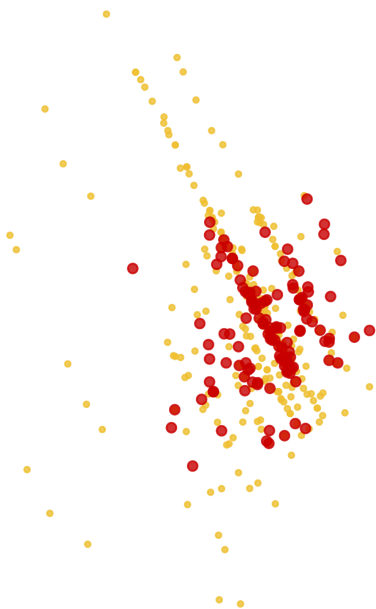
Categorías de sabor: Verdes, Terrosos y Herbáceo

B**TÍTULO: SYRAH**

Técnica: PCA

Descriptores: MW, HB, HBA, SlogP, TPSA, and RB.
 Compuestos químicos: Flavor food-chemical compounds.

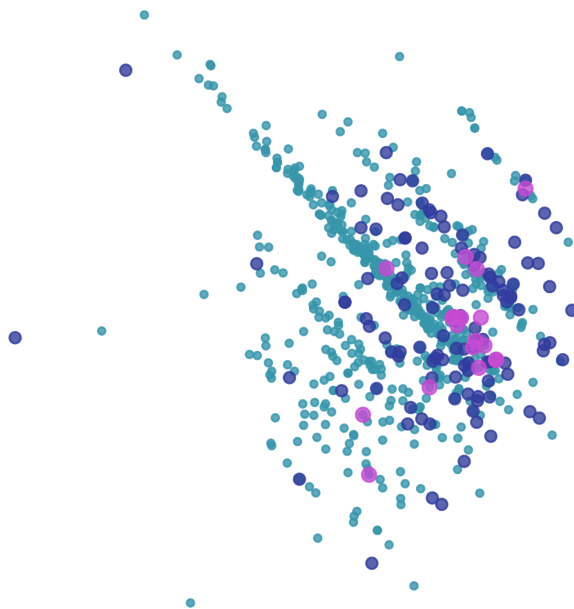
Categorías de sabor: Afrutado y Floral.

C**TÍTULO: FONTINA**

Técnica: PCA

Descriptores: MW, HB, HBA, SlogP, TPSA, and RB.
 Compuestos químicos: Flavor food-chemical compounds.

Categorías de sabor: Graso y Picante.

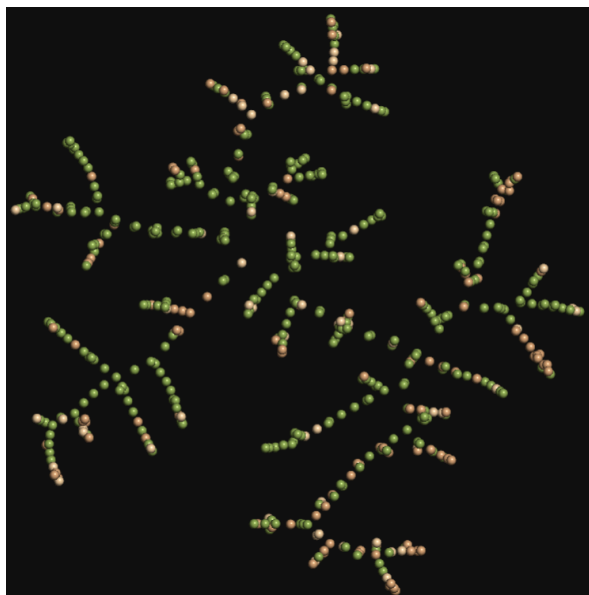
D**TÍTULO: XOCOC**

Técnica: PCA

Descriptores: MW, HB, HBA, SlogP, TPSA, and RB.
 Compuestos químicos: Flavor food-chemical compounds.

Categorías de sabor: Balsámico, Químico y Medicinal.

Figura 9. Cuatro categorías de sabor. **A)** Sabores de la Tierra (655 compuestos), **B)** Matices de vino (1024 compuestos), **C)** Contraste entre picante y graso (430 compuestos), y **D)** Remedios Naturales (762 compuestos). Los nombres de las visualizaciones de la Figura 9 se asignaron con nombres de alimentos contenidos en alguna de las categorías de sabor que se utilizaron para llevar a cabo la visualización, por ejemplo la obra que lleva por título “*PANTE*” se refiere al nombre de un Hongo mexicano, ya que los hongos se encuentran en la categoría de sabor “Earthy” o “Terrosos”.

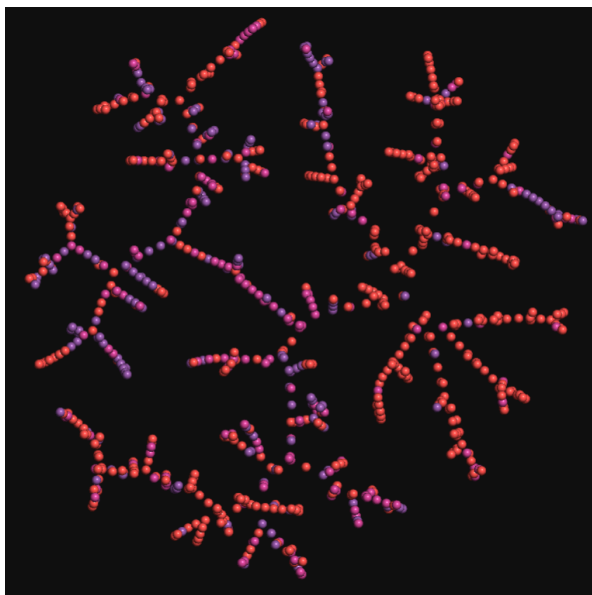
A**TÍTULO: *Ébano (Ebony)***

Técnica: TMAP

Descriptor: MACCS keys fingerprints

Compuestos químicos: Flavor food-chemical compounds.

Categorías de sabor: Verdes, Terrosos y Herbáceo

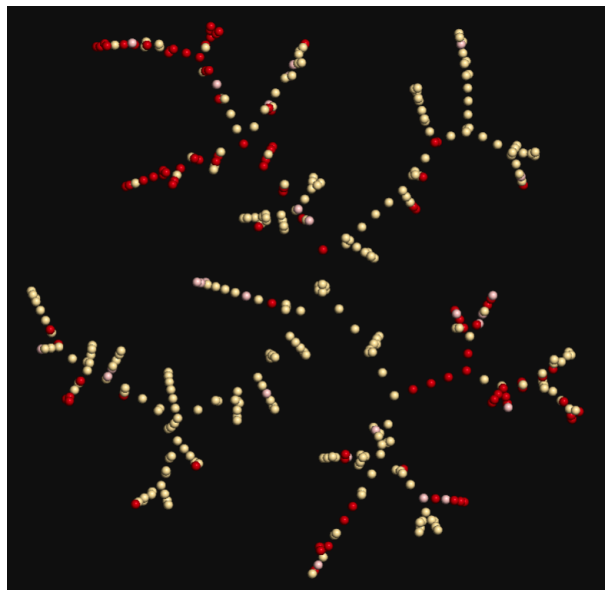
B**TÍTULO: *Flor de corazón (Heart flower)***

Técnica: TMAP

Descriptor: MACCS keys fingerprints

Compuestos químicos: Flavor food-chemical compounds.

Categorías de sabor: Afrutado y Floral

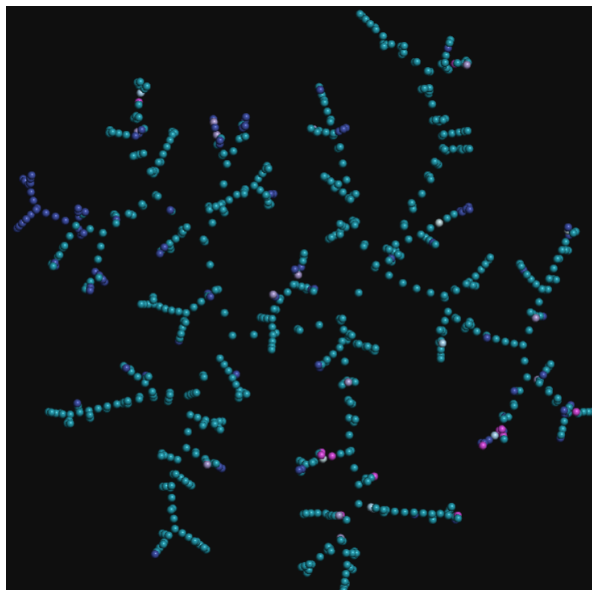
C**TÍTULO: *Amaranto (Amaranth)***

Técnica: TMAP

Descriptor: MACCS keys fingerprints

Compuestos químicos: Flavor food-chemical compounds.

Categorías de sabor: Graso y Picante

D**TÍTULO: *Huele a miel (Honey's smell)***

Técnica: TMAP

Descriptor: MACCS keys fingerprint

Compuestos químicos: Flavor food-chemical compounds.

Categorías de sabor: Balsámico, Químico y Medicinal

Figura 10. Cuatro categorías de sabor. **A)** Sabores de la Tierra (655 compuestos), **B)** Matices de vino (1024 compuestos), **C)** Contraste entre picante y graso (430 compuestos), y **D)** Remedios Naturales (762 compuestos). La idea para otorgar nombre a estas obras digitales (Figura 10) fue optar por nombres de árboles mexicanos y asignarlos de acuerdo a los colores representativos de cada obra. Por ejemplo en la visualización que lleva el título de “Amaranto” o “Amaranth” se observa una mayoría de pequeñas esferas de color “crema” que hacer referencia al amaranto.

6. Discusión de resultados

La aplicación extensiva de la quimioinformática se ha utilizado principalmente en el descubrimiento de fármacos. No obstante, sus aplicaciones abarcan un espectro mucho más amplio dentro del ámbito químico. Esto es evidente en su creciente utilización en la química alimentaria, como lo demuestra el surgimiento de nuevas áreas de investigación como la informática química alimentaria y la informática alimentaria ³⁶⁻³⁷. Además, su utilidad se extiende a la investigación en productos naturales ³⁸⁻³⁹, la caracterización de polímeros y materiales, entre otros ¹⁶. En este contexto, propongo expandir las aplicaciones de la quimioinformática a través de la representación visual del espacio químico contenido en conjuntos de datos de compuestos. Esta idea se ejemplifica en nuestro enfoque en sustancias químicas alimentarias, donde buscamos crear "obras de arte" emblemáticas. La conexión o sinergia entre la quimioinformática y el arte presenta un potencial significativo para unir al menos dos segmentos de la población que, en otras circunstancias, podrían permanecer desconectados. Además, desde una perspectiva educativa - un componente esencial en el campo de la quimioinformática - esta combinación no solo puede hacer que conceptos científicos complejos sean más accesibles y comprensibles para el público en general, sino que también puede despertar el interés y la curiosidad en jóvenes estudiantes y niños hacia el mundo de la química mediante la expresión artística.

La colaboración entre artistas y científicos emerge como un catalizador de enfoques innovadores y singulares para la explicación de conceptos científicos. Esta sinergia tiene el potencial de captar la atención de audiencias variadas y heterogéneas. Las representaciones artísticas, al dejar una impresión profunda en la mente de las personas, desempeñan un papel fundamental en la retención de información científica. Como resultado, esta fusión creativa no solo despierta la curiosidad, sino que también establece una conexión emocional con el conocimiento científico, generando un impacto duradero en aquellos que se involucran en esta experiencia única.

6.1. Ejemplos de espacios químicos y multiversos relacionados con el arte

Los ejemplos de representación visual del espacio químico como representaciones artísticas (Sección 5) se centran en en compuestos químicos alimenticios y descriptores moleculares adecuados para representar dichos compuestos químicos. Sin embargo, el número de técnicas de visualización

establecidas, descriptores moleculares y el número de estructuras químicas disponibles, es vasto. Por lo tanto, existe una multitud de formas, literalmente miles o millones, de crear obras artísticas basadas en el espacio químico. La Tabla 2 resume ejemplos de visualización del espacio y multiverso químico impulsada por la quimioinformática. La tabla da ejemplos de conjuntos de datos que podrían emplearse para representar, desde una perspectiva artística, su amplitud, complejidad, y diversidad. Hay muchos otros conjuntos de datos y una variada gama de combinaciones posibles de descriptores y técnicas de visualización. No obstante, como en cualquier manifestación artística, el valor real radica en la capacidad de contar historias o transmitir mensajes, que a veces permanecen ocultos.

Tabla 2. Ejemplos de obras digitales potenciales basadas en la visualización del espacio químico.

Base de datos	Concepto artístico	Nombre de la obra
Compuestos químicos aleatorios	Las moléculas aleatorias representan la inmensidad de nuestro cosmos y de nuestra rutina diaria. Aunque interactuamos constantemente con muchas sustancias químicas, no nos fijamos en su complejidad y caos inherente que existe en nuestro universo y en nuestra vida cotidiana.	“Caos” “Chaos”
Base de datos diversas	La diversidad ofrece múltiples colores, sabores, gustos y experiencias. En la naturaleza, la diversidad (en todos los sentidos) es una de las características habituales.	“Diversity” “Dispar” “Pride”
Productos Naturales Marinos	El mar permanece más allá de nuestra comprensión. Tiene vida, muerte, color y oscuridad. permanece en constante cambio.	“El océano” “Inmensidad” “Regeneration”
Fármacos aprobados para el tratamiento del VIH	Todo puede suceder en un humano VIH positivo; miedo, recuerdos, felicidad y normalidad. Los fármacos ayudan...pero no siempre son la solución absoluta.	“Vivir con sida” “Living with AIDS”
Hormonas - Neurotransmisores	Amor = hormonas + neurotransmisores + personas especiales.	“Química del amor”
Compuestos químicos asociados a la depresión	Depresión = hormonas + neurotransmisores - propósito.	“Darkness” “Un empujón a la serotonina”
Compuestos químicos alimenticios	Los grandes placeres de la vida suelen ir acompañados de sabores, colores, texturas y aromas.	“Bellyful” “Viaje de sabores”
Base de datos ZINC vs. Compuestos similares a fármacos	Sabemos mucho acerca de la naturaleza y cómo se constituye, sin embargo, el conocimiento es limitado. Nuestro saber se asemeja a una pequeña marca en un lienzo mucho más extenso, el cual aún nos resulta desconocido en su totalidad.	“Our knowledge” “Wisdom”

Para proporcionar una comprensión más profunda sobre el potencial de crear expresiones artísticas a través de la representación visual del espacio químico,, la Figura 11 exhibe un ejemplo de representación artística del espacio químico a partir de un conjunto de datos de productos naturales aleatorios, que han sido decodificados por sus descriptores de efectos secundarios (por ejemplo, mutagénesis, tumorigénesis y efectos reproductivos negativos, etc.). La paleta de colores, que va desde tonos rojizos hasta tonos azulados, simboliza la probabilidad de que cada producto natural provoque efectos secundarios. La obra en este "lienzo" se compone mediante una técnica de puntos, la cual refleja una posible variedad de texturas alcanzables mediante este enfoque artístico. Como en la Figura 11, sabemos intrínsecamente que la "naturaleza" no siempre es sana y que en nosotros existe un delicado equilibrio que es muy fácil romper.

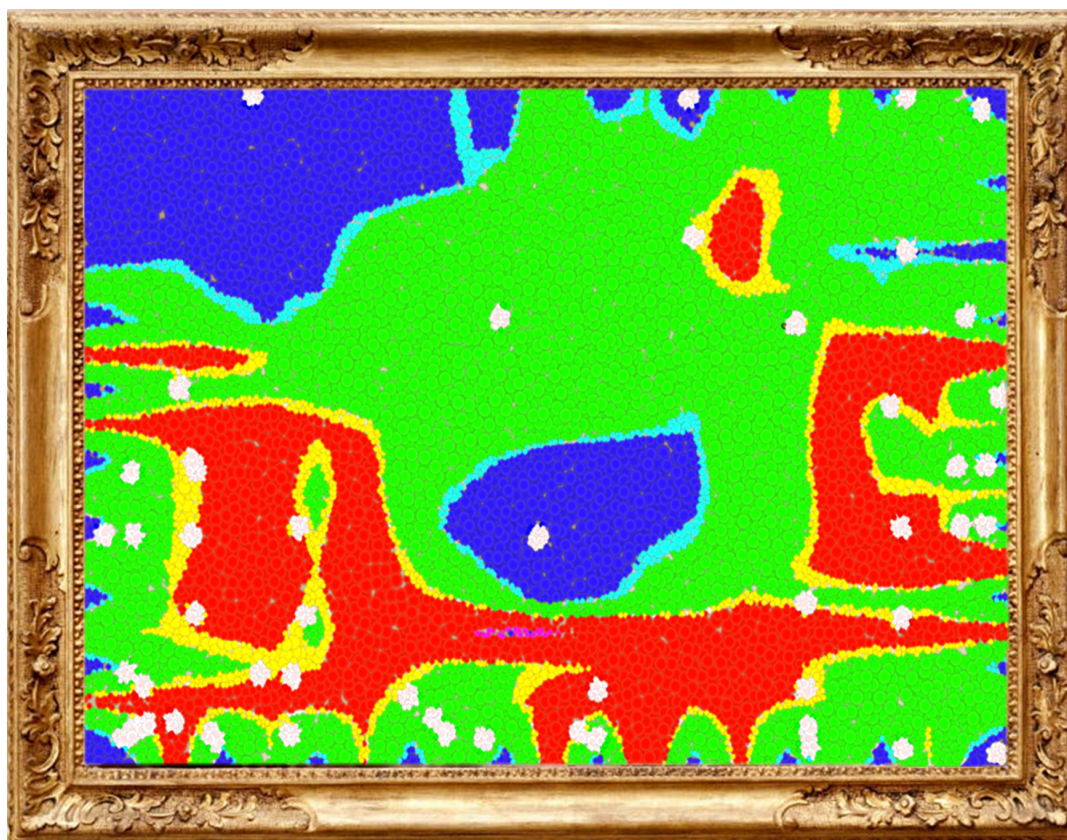
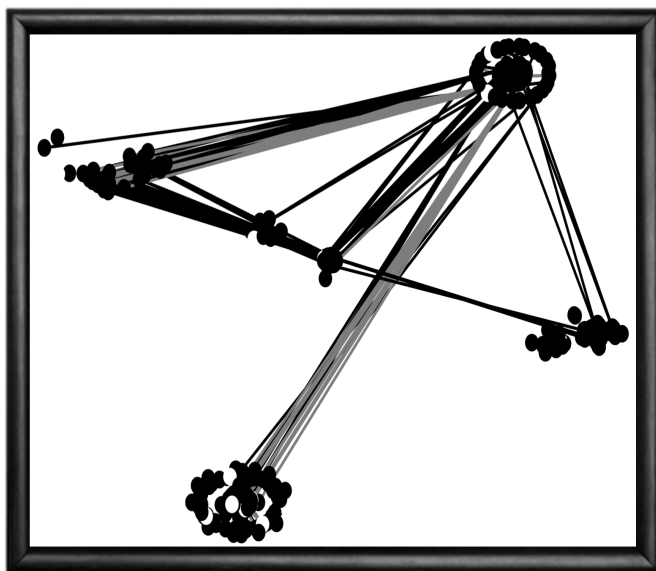


Figura 11. Ejemplo de arte del espacio químico. **Título:** "Wise nature"; **Autor:** Edgar López-López; **Técnica:** SOM - using DataWarrior software⁴⁰⁻⁴¹; **Base de datos:** Compuestos naturales aleatorios (1000 compuestos); **Descriptores:** Mutagénico predictivo, tumorigénico, eficaz para la reproducción, e irritante; **Descripción técnica:** Cada punto blanco es un producto natural, las regiones iluminadas en rojo representan el espacio químico con una alta probabilidad predictiva de contener compuestos con efectos secundarios, lo contrario para el color azul; **Interpretación artística:** La "naturaleza" no siempre es sana, siempre ha existido una dualidad entre lo que nos llena de vida y lo que acaba con ella.

La Figura 12 muestra ejemplos adicionales de obras de arte o ilustraciones del espacio químico que combinan diferentes métodos de reducción de datos y descriptores para generar una representación visual artística de las bases de datos de compuestos químicos. Aquí invito a los lectores a reflexionar y encontrar alguna otra interpretación artística que se pueda adjudicar a estas visualizaciones.



TÍTULO: Chemical umbrella

Técnica: PCA + Data fusion (chemical multiverse approach).

Descriptores: Datos de inhibición celular y enzimática. Los puntos están conectados en función de su actividad inhibidora frente a distintos tipos de citocromos (proteínas relacionadas con la protección hepática).

Compuestos hepatotóxicos.

TÍTULO: Brake cancer

Técnica: Diagramas de constelación.

Descriptores: Datos de inhibición de células anticancerígenas.

Fármacos contra el cáncer.

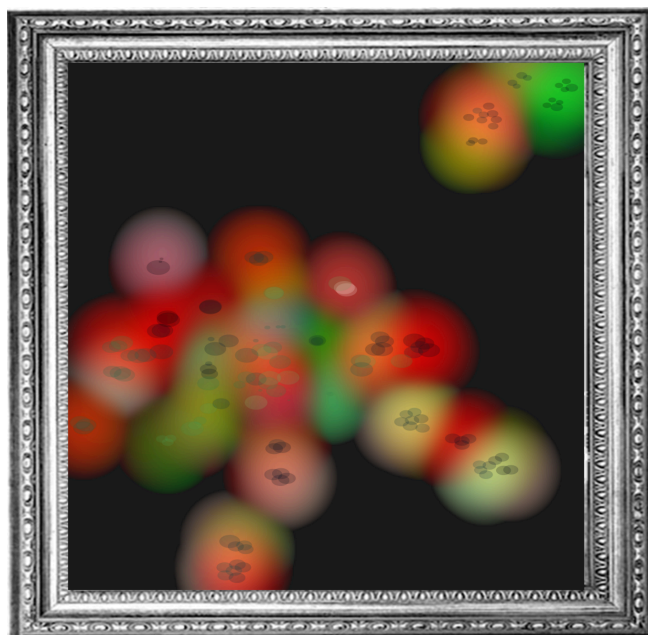


Figura 12. Ejemplos de obras del espacio químico. Las obras de arte se generaron con datos del dominio público ⁴²⁻⁴⁴.

6.2. Inteligencia artificial y arte digital

IA se emplea para crear representaciones, o expresiones artísticas ^{11,45}. Aunque no constituye el enfoque central de este escrito, la Figura 13 ejemplifica imágenes generadas a través de recursos de libre acceso, utilizando palabras clave asociadas al “espacio químico”. La figura ilustra un caso de

multiverso químico generado por un servidor web de IA, cuyo entrenamiento se basa en la alimentación de múltiples palabras u oraciones. Aunque las imágenes son atractivas, hay una diferencia notable con las obras de arte del espacio químico presentadas en secciones previas (Figuras 7-10) y esta radica en que las imágenes de la Figura 13 se originan de un proceso de entrenamiento mediante palabras clave, mientras que las primeras obras de la sección de resultados (Sección 5) emanan directamente de estructuras químicas codificadas con descriptores moleculares. Otro aspecto importante es que en las primeras representaciones hay una mayor comprensión, intervención humana y abstracción de los conceptos e ideas, algo cuestionable en las representaciones guiadas por IA.

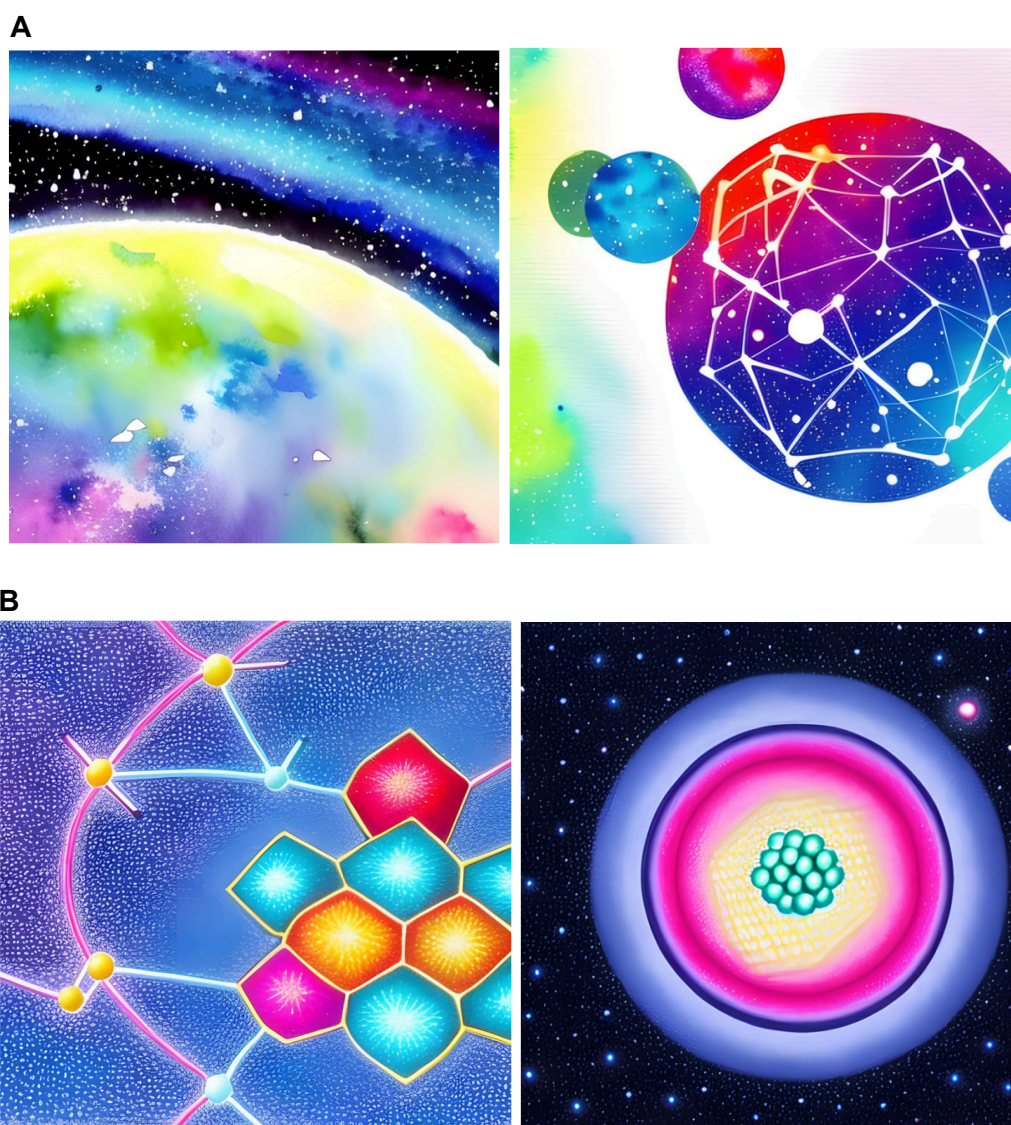


Figura 13. Ejemplo(s) de arte impulsado por inteligencia artificial con la aplicación gratuita Canva (<https://www.canva.com/>) utilizando la palabra clave espacio químico y A) acuarela y B) lápiz de color.

7. Conclusiones

La ciencia y el arte están íntimamente relacionados desde hace mucho tiempo. Un ejemplo típico se resume en la frase: “El descubrimiento de fármacos es tanto un arte como una ciencia”. Es cierto que la química desempeña un papel fundamental en el ámbito artístico, incluyendo la restauración y preservación de piezas artísticas. Pero hay una tendencia emergente a aplicar. Sin embargo, está emergiendo una tendencia que busca aplicar los conceptos químicos para generar expresiones artísticas u obras de arte. En este trabajo se buscó un enfoque que busca fusionar el arte con la quimioinformática a través de las representaciones visuales del espacio químico. Se presentaron algunos ejemplos de obras de arte que tomaron como inspiración el espacio químico, que pueden considerarse “cuadros digitales”. El autor de los gráficos de dimensiones reducidas puede utilizar los gráficos con un doble propósito general: comunicar datos y generar información química (como se hace normalmente con las visualizaciones del espacio químico) y comunicar un significado emocional o personal al gráfico (impulsado por principios de química e informática).

También llegué a la conclusión que las creaciones artísticas basadas en el espacio químico tienen el poder de ser auténticos embajadores de la ciencia en general y de la química en su peculiaridad, llevando consigo el mensaje a un público amplio y diverso. El concepto de espacio químico visto desde un enfoque artístico puede desempeñar un papel significativo en el gran público a ver las moléculas más allá de las entidades individuales y las representaciones convencionales. En consecuencia, las manifestaciones artísticas impulsadas por la informática química adquieren relevancia como una vía de divulgación científica. Tal enfoque se alinea con los resúmenes gráficos utilizados frecuentemente en las revistas revisadas por pares. El "arte químico" podría ser útil para representar datos complejos, pero desde una perspectiva artística y atractiva. La persona que genere la representación del espacio químico podría considerarse un "artista del espacio químico".

Preveo varios desarrollos y áreas de oportunidad adicionales para el arte impulsado por las representaciones visuales del espacio químico. La Tabla 3 resume los proyectos de arte químico en curso, abarcando desde la generación de herramientas “de fácil uso”, la instauración de la primera galería de arte químico, hasta la aplicación de este enfoque artístico para introducir a la nueva generación de expertos en quimioinformática en el concepto de espacio químico. Simultáneamente, los

métodos de inteligencia artificial continuarán expandiendo y explorando el espacio químico, brindando nuevos tipos de moléculas y descriptores que podrían utilizarse para aumentar las posibilidades de representar el espacio químico desde la perspectiva de un artista.

Tabla 3. Innovaciones representativas de la combinación del arte con la quimioinformática a través de visualizaciones artísticas del espacio químico.

Inovación	Potencial resultado o aplicación
Continuar desarrollando una compilación digital centrada en la representación artística del espacio químico.	<i>Museo de Arte del Espacio Químico</i>
Crear procesos automatizados mediante el empleo de software de código abierto u herramientas informáticas, con el propósito de ampliar la accesibilidad de esta forma artística a individuos con diversas trayectorias académicas o artísticas.	<i>Generadores ChemArt</i>
Establecer un repositorio permanente, gratuito y de libre acceso de obras de arte. Esta iniciativa promoverá la divulgación científica y artística de manera abierta. La comunidad científica y artística podría apoyar el repositorio.	<i>Galería ChemART</i>
Implementar un programa educativo o cultural sostenido como exposición abierta y permanente.	<i>Programa “Arte conducido por la visualización del Espacio Químico”</i>

8. Referencias

- (1) Galván-Madrid, J. L. (2011). La química y el arte: ¿Cómo mantener el vínculo? *Educación química*, 22(3), 207–211.
https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0187-893X2011000300005
- (2) Bello, D. G. *La química de lo bello*, 2nd ed.; Ediciones Paidós, 2023.
- (3) Orna, M. V. Chemistry, Color, and Art. *J. Chem. Educ.* 2001, 78 (10), 1305, DOI: 10.1021/ed078p1305.
- (4) Kafetzopoulos, C.; Spyrellis, N.; Lymperopoulou-Karaliota, A. The Chemistry of Art and the Art of Chemistry. *J. Chem. Educ.* 2006, 83 (10), 1484, DOI: 10.1021/ed083p1484.
- (5) Ball, P. (2003). *La invención del color* (1a ed.). FCE, TURNER.
<http://www.librosmaravillosos.com/lainvenciondelcolor/index.html>
- (6) Mendoza-Cuevas A; Nodarse-Martín L.; Calas M.;Alonso H. (2008). Análisis físico-químico de materiales artísticos de la pintura “Colonización de México”. *CONSERVA N°12*, 14
- (7) Lipscher, J. (2018). Chemistry and Art. *Chimia*, 72(1–2), 27.
<https://doi.org/10.2533/chimia.2018.27>
- (8) Otero, J., & Vanessa Gutiérrez Cano. (2015). *Espectroscopía Raman: Fundamento y aplicaciones*. <https://doi.org/10.13140/rg.2.1.5015.5362>
- (9) Epstein, Z., Levine, S., Rand, D. G., & Rahwan, I. (2020). Who gets credit for AI-generated art? *IScience*, 23(9), 101515. <https://doi.org/10.1016/j.isci.2020.101515>
- (10) Rodríguez-Mondragón, S., Soto-Walls, L. J., Clavé-Almeida, M. M., & Herrera-Alcántara, O. (2019). Métodos computacionales para la creación artística. *Research in Computing Science*, ISSN 1870-4069 (28), 375-389
- (11) DALL·E 2 <https://openai.com/dall-e-2/> (accessed Jun 20, 2023).
- (12) Mahjour, B., Bench, J., Zhang, R., Frazier, J., & Cernak, T. (2022). Molecular sonification for molecule to music information transfer. En *ChemRxiv*.
<https://doi.org/10.26434/chemrxiv-2022-g7xkl>
- (13) Cuevas, G. (2005) ‘Química computacional’, *Revista Ciencias Matematicas*, 56(2), pp. 33–42.
- (14) Damm-Ganamet, K. L., Rives, M.-L., Wickenden, A. D., McAllister, H. M., & Mirzadegan, T.

- (2020). A computational approach yields selective inhibitors of human excitatory amino acid transporter 2 (EAAT2). *The Journal of Biological Chemistry*, 295(13), 4359–4366. <https://doi.org/10.1074/jbc.ac119.011190>
- (15) Varnek, A., & Baskin, I. I. (2011). Chemoinformatics as a theoretical chemistry discipline. *Molecular Informatics*, 30(1), 20–32. <https://doi.org/10.1002/minf.201000100>
- (16) López-López, E., Bajorath, J., & Medina-Franco, J. L. (2021). Informatics for chemistry, biology, and biomedical sciences. *Journal of Chemical Information and Modeling*, 61(1), 26–35. <https://doi.org/10.1021/acs.jcim.0c01301>
- (17) Saldívar-González, F. I., Aldas-Bulos, V. D., Medina-Franco, J. L., & Plisson, F. (2022). Natural product drug discovery in the artificial intelligence era. *Chemical Science (Royal Society of Chemistry: 2010)*, 13(6), 1526–1546. <https://doi.org/10.1039/d1sc04471k>
- (18) Saldívar-González, F. I., Fernández- de Gortari, E., & Medina-Franco, J. L. (2023). Inteligencia artificial en el diseño de fármacos: hacia la inteligencia aumentada. *Educación Química*, 34(2), 17–25. <https://doi.org/10.22201/fq.18708404e.2023.2.83233>
- (19) Greener, J. G.; Kandathil, S. M.; Moffat, L.; Jones, D. T. A Guide to Machine Learning for Biologists. *Nat. Rev. Mol. Cell Biol.* **2021**, 23 (1), 40–55, DOI: 10.1038/s41580-021-00407-0.
- (20) van der Maaten, L.; Hinton, G. Visualizing Data using t-SNE <https://www.jmlr.org/papers/volume9/vandermaaten08a/vandermaaten08a.pdf?fbcl> (accessed Jun 1, 2023).
- (21) Probst, D.; Reymond, J.-L. Visualization of Very Large High-Dimensional Data Sets as Minimum Spanning Trees. *J. Cheminf.* **2020**, 12 (1), 1-13, DOI: 10.1186/s13321-020-0416-x .
- (22) Wigh, D. S., Goodman, J. M., & Lapkin, A. A. (2022). A review of molecular representation in the age of machine learning. *Wiley Interdisciplinary Reviews. Computational Molecular Science*, 12(5). <https://doi.org/10.1002/wcms.1603>
- (23) Weininger, D. SMILES, a Chemical Language and Information System. 1. Introduction to Methodology and Encoding Rules. *J. Chem. Inf. Comput. Sci.* **1988**, 28 (1), 31-36, DOI: 10.1021/ci00057a005.
- (24) David, L., Thakkar, A., Mercado, R. *et al.* Molecular representations in AI-driven drug discovery:

- a review and practical guide. *J Cheminform* 12, 56 (2020).
<https://doi.org/10.1186/s13321-020-00460-5>
- (25) Virshup, A. M., Contreras-García, J., Wipf, P., Yang, W., & Beratan, D. N. (2013). Stochastic voyages into uncharted chemical space produce a representative library of all possible drug-like compounds. *Journal of the American Chemical Society*, 135(19), 7296–7303.
<https://doi.org/10.1021/ja401184g>
- (26) Saldívar-González, F. I.; Medina-Franco, J. L. Approaches for Enhancing the Analysis of Chemical Space for Drug Discovery. *Expert Opin. Drug Discov.* **2022**, 17 (7), 789-798, DOI: 10.1080/17460441.2022.2084608.
- (27) Miranda-Salas, J.; Peña-Varas, C.; Valenzuela Martínez, I.; Olmedo, D. A.; Zamora, W. J.; Chávez-Fumagalli, M. A.; Azevedo, D. Q.; Castilho, R. O.; Maltarollo, V. G.; Ramírez, D.; Medina-Franco, J. L. Trends and Challenges in Chemoinformatics Research in Latin America. *Artif. Intell. Life Sci.* **2023**, 3 (1), 100077, DOI: 10.1016/j.aillsci.2023.100077.
- (28) Medina-Franco, J. L., Chávez-Hernández, A. L., López-López, E., & Saldívar-González, F. I. (2022). Chemical multiverse: An expanded view of chemical space. *Molecular Informatics*, 41(11). <https://doi.org/10.1002/minf.202200116>
- (29) Osolodkin, D. I.; Radchenko, E. V.; Orlov, A. A.; Voronkov, A. E.; Palyulin, V. A.; Zefirov, N. S. Progress in Visual Representations of Chemical Space. *Expert Opin. Drug Discov.* 2015, 10 (9), 959-973, DOI: 10.1517/17460441.2015.1060216.
- (30) Medina-Franco, J.; Martinez-Mayorga, K.; Giulianotti, M.; Houghten, R.; Pinilla, C. Visualization of the Chemical Space in Drug Discovery. *Curr. Comput. Aided Drug Des.* 2008, 4 (4), 322–333, DOI: 10.2174/157340908786786010.
- (31) FooDB <https://foodb.ca/> (accessed Apr 20, 2023).
- (32) RDKit <https://www.rdkit.org> (accessed 08 January 2022).
- (33) MolVS <https://molvs.readthedocs.io/en/latest/> (accessed Jan 08, 2022).
- (34) Durant, J. L.; Leland, B. A.; Henry, D. R.; Nourse, J. G. Reoptimization of MDL Keys for Use in Drug Discovery. *J. Chem. Inf. Comput. Sci.* **2002**, 42 (6), 1273–1280, DOI: 10.1021/ci010132r.
- (35) Rogers, D.; Hahn, M. Extended-Connectivity Fingerprints. *J. Chem. Inf. Model.* **2010**, 50 (5),

742-754, DOI: 10.1021/ci100050t.

- (36) *Foodinformatics: Applications of Chemical Information to Food Chemistry*; Martinez-Mayorga, K., Medina-Franco, J. L., Eds.; Springer International Publishing: Cham, Switzerland, 2014.
- (37) Peña-Castillo, A.; Méndez-Lucio, O.; Owen, J. R.; Martínez-Mayorga, K.; Medina-Franco, J. L. Chemoinformatics in Food Science. In *Applied Chemoinformatics*; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2018; pp 501–525.
- (38) Kirchmair, J. Molecular Informatics in Natural Products Research. *Mol. Inf.* **2020**, *39* (11), 2000206, DOI: 10.1002/minf.202000206.
- (39) Medina-Franco, J. L.; Saldívar-González, F. I. Cheminformatics to Characterize Pharmacologically Active Natural Products. *Biomolecules* **2020**, *10* (11), 1566, DOI: 10.3390/biom10111566.
- (40) Sander, T.; Freyss, J.; von Korff, M.; Rufener, C. DataWarrior: An Open-Source Program for Chemistry Aware Data Visualization and Analysis. *J. Chem. Inf. Model.* **2015**, *55* (2), 460-473, DOI: 10.1021/ci500588j.
- (41) López-López, E.; Naveja, J. J.; Medina-Franco, J. L. DataWarrior: An Evaluation of the Open-Source Drug Discovery Tool. *Expert Opin. Drug Discov.* **2019**, *14* (4), 335-341, DOI: 10.1080/17460441.2019.1581170.
- (42) López-López, E.; Medina-Franco, J. L. Towards Decoding Hepatotoxicity of Approved Drugs through Navigation of Multiverse and Consensus Chemical Spaces. *Biomolecules* **2023**, *13* (1), 176, DOI: 10.3390/biom13010176.
- (43) Medina-Franco J. L.; Naveja J. J.; López-López E. Reaching for the Bright StARs in Chemical Space. *Drug Discov. Today* **2019**, *24* (11), 2162-2169, DOI: 10.1016/j.drudis.2019.09.013.
- (44) López-López, E.; Cerda-García-Rojas, C. M.; Medina-Franco, J. L. Tubulin Inhibitors: A Chemoinformatic Analysis Using Cell-Based Data. *Molecules* **2021**, *26* (9), 2483, DOI: 10.3390/molecules26092483.
- (45) Ramesh, A.; Dhariwal, P.; Nichol, A.; Chu, C.; Chen, M. Hierarchical Text-Conditional Image Generation with CLIP Latents, 2022, DOI: arXiv:2204.06125.

ANEXO

Los resultados de este trabajo están disponibles públicamente en:

Gaytán-Hernández D, Chávez-Hernández AL, López-López E, Miranda-Salas J, Saldívar-González FI, Medina-Franco JL. Art Driven by Visual Representations of Chemical Space. ChemRxiv. Cambridge: Cambridge Open Engage; **2023**; DOI: [10.26434/chemrxiv-2023-3tbcl](https://doi.org/10.26434/chemrxiv-2023-3tbcl)

The screenshot shows the ChemRxiv website interface. At the top, the ChemRxiv logo is on the left, and navigation links for 'How To Submit', 'Browse', 'About', and 'News' are in the center. A search bar and a 'Log In' button are on the right. Below the navigation is a banner for 'Chemical Education' with a search bar and a 'Search within Chem...' button. The main content area features the article title 'Art Driven by Visual Representations of Chemical Space' and the date '27 June 2023, Version 1'. A 'Working Paper' badge is present. The authors listed are Daniela Gaytán-Hernández, Ana L. Chávez-Hernández, Edgar López-López, Jazmín Miranda-Salas, Fernanda I. Saldívar-González, and José L. Medina-Franco. A 'Show author details' link is available. A 'Comment' button is at the bottom left. On the right side, there is a 'Download' button and a 'Version History' section showing 'Jun 27, 2023 Version 1'. Below that, a 'Metrics' section displays: 566 Views, 255 Content Downloads, and 0 Citations. A 'License' section is partially visible at the bottom.

A continuación se adjunta la liga para acceder al artículo en formato PDF:

<https://chemrxiv.org/engage/api-gateway/chemrxiv/assets/orp/resource/item/649990731dccb92a5e975f79/original/art-driven-by-visual-representations-of-chemical-space.pdf>