

## Ruta metodológica para la construcción del atlas histórico de la ciencia colombiana

Daniela González Correa  
daniela.9.5875@gmail.com  
Estefany Suaza Martínez  
estefany121212@gmail.com  
Universidad de Antioquia

### **Resumen**

Enfrentarse a los interrogantes relacionados con el proceso de institucionalización de la ciencia en Colombia se presenta como un reto intelectual que demanda una búsqueda rigurosa de material empírico, en este caso documental, admitido como evidencia de los procesos y momentos condicionantes de lo que denominamos ciencia. Aunque los estudios sociales de la ciencia en Colombia han aumentado en los últimos años es importante resaltar que por mucho tiempo fueron casi nulos, suscitando un vacío en el análisis de la producción científica local. Al igual que en este vacío sólo recientemente se han abordado una serie de pautas reguladoras para la evaluación y clasificación del conocimiento, donde se espera que este tenga una función social que contribuya a la explicación y transformación del mundo, y que sea construido en constante colaboración con otros (autores, instituciones).

El Atlas Histórico de la Ciencia Colombiana es un proyecto concebido en la Universidad de Antioquia por un grupo interdisciplinario de profesores, académicos, profesionales y estudiantes que tiene como fin presentar un estudio exploratorio que dé cuenta de la producción científica colombiana, publicada entre el siglo XVIII y finales del siglo XX en 65 bases de datos científicas. Para la elección de estas bases de datos se tuvo en cuenta los siguientes criterios: por un lado, que ellas mismas se considerarán bases de datos científicas y, por otro, que expertos del Cinvestav (Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional, México) y de la Universidad de Antioquia aprobaran que efectivamente se trataban de bases de datos científicas. Se acudió a ellas sin desconocer que la construcción del Atlas Histórico de la Ciencia Colombiana solo puede elaborarse con la información “visible”, es decir, disponible, pues sobre lo que está al margen de las publicaciones no tenemos certezas; quedamos expuestos a una “ceguera” involuntaria, que,

aunque lamentable, no nos hace responsables del escrutinio que definía lo que merecía ser “ciencia”.

Las bases fueron susceptibles de algunos filtros: El primero de ellos consistió en identificar los documentos en los que pudieran hallarse las denominaciones que tuvo el país según el momento histórico (La Gran Colombia, Nueva Granada, Confederación Granadina, Provincias Unidas de Colombia, República de Colombia); el segundo se concentró en la delimitación temporal en que fueron publicados los documentos (1800-1979), acotación planteada al considerar el periodo de tiempo en que los textos históricos obtuvieron licencia para ser publicados libremente. Adicionalmente, y sin abandonar la interdisciplinariedad del grupo de personas que conforman el proyecto, se realiza un acercamiento al objeto de estudio desde tres enfoques: el bibliotecológico, compuesto por análisis documental para la identificación de metadatos y posteriormente la estandarización de campos, en este enfoque se tuvo en cuenta la concepción actual de ciencia y los campos que constituían lo que anteriormente se reconocía como científico. Esto permite una conexión con un segundo enfoque: el historiográfico, el cual abordó discusiones sobre lo que era ciencia, permitiéndonos identificar estándares y rasgos de los documentos científicos. El tercer enfoque, sociológico, orientó el proyecto en el debate sobre institucionalización de la ciencia.

**Palabras clave:** Fuentes digitales, bases de datos, humanidades digitales, documento científico.

## **Introducción**

En Colombia, en la Universidad de Antioquia, se está gestando un proyecto que lleva por nombre Atlas Histórico de la Ciencia Colombiana, es objeto de la presente ponencia no solo resumir la ruta metodológica que hasta ahora se ha llevado a cabo para el desarrollo del proyecto, sino también, de manera mas concreta, exponer el proceso de identificación del documento científico en las bases de datos digitales.

El proyecto prevé explorar 65 bases de datos científicas e identificar en cada una de ellas los documentos científicos de un periodo de tiempo que oscila entre el siglo XVIII, XIX y finales del siglo XX, es por ello que surge la necesidad de hacer una breve contextualización no necesariamente de lo que se concebía por ciencia, sino mas bien de la forma en la que se empiezan a comunicar los productos que se construían a partir de un esfuerzo de trabajo

científico. Para empezar, en el siglo XVIII (donde ya existía un interés por generar un desarrollo intelectual en el país), los primeros acercamientos comprenden estudios sobre botánica, geografía y un particular interés por mapear las regiones, lo cual coincide con los documentos hallados en la revisión de las bases de datos, entre los cuales se encuentran contenidos sobre exploración de plantas, flora, fauna, historia de Colombia, topografía y corografía. El siglo XVIII permite señalar elementos que hablan del desarrollo científico en el país, aunque de forma algo imprecisa, pues la producción de documentos científicos era muy limitada, aun así, los documentos de la época dan cuenta de las prácticas científicas que se empezaban a llevar a cabo, como los estudios de la geografía del país donde se puede identificar los instrumentos y resultados de las investigaciones o en las expediciones botánicas donde se clasificaban la fauna y flora.

La incorporación de la ciencia en el país está marcada por varios sucesos y personajes, de ahí la importancia de recurrir a la historia para realizar una adecuada lectura de esta; José Celestino Mutis y Francisco Antonio Moreno son algunos de esos personajes y permiten ver que el interés naciente a mediados del siglo XVIII por realizar actividades científicas en el país contaba con influencia europea. Por un lado, Mutis, matemático, geógrafo y médico español, es quien propone la expedición botánica en el país donde primaban los intereses económicos sobre los científicos, la expedición dura aproximadamente 30 años y es en ella donde se forman gran cantidad de criollos en áreas como la botánica y la topografía. Durante la expedición la lámina se convirtió en el documento cumbre para plasmar los hallazgos, ella contenía el dibujo de las plantas, la descripción y clasificación, todo con mucho detalle, de igual manera la difusión de las láminas no era pasada por alto, las que no eran llevadas a España eran publicadas en el Semanario del Nuevo Reyno de Granada que era un periódico que se concentraba en la producción de textos alrededor de temas científicos (Nieto, s.f). Estos son acontecimiento de gran importancia para el análisis del documento científico y el valor de las fuentes digitales que los contienen.

El siglo XIX, por su parte, se caracteriza por el surgimiento de centros de enseñanza y la formación de grupos de personas a favor del cumplimiento de los intereses científicos de los ilustrados que llegaban en su mayoría de Europa. La minería, química, matemática, física y la geografía, realizándose algunas expediciones. La geografía, objeto de variadas investigaciones en aquella época agrupó a varios académicos colombianos quienes empiezan a colaborar para publicar lo que ayudó a que tiempo después se constituyera la primera gran muestra histórica científica del país (Melo, 1987). Hasta ahora la práctica científica

propiamente dicha no había alcanzado niveles altos de desarrollo lo que originó que los documentos científicos de aquella época comprendieran publicaciones sobre historia y geografía. Posterior al año 1850 la medicina dominada hasta entonces por médicos extranjeros empieza a ser protagonista de publicaciones periódicas en revistas de carácter científico como La Gaceta Médica de Colombia, la Revista Médica de Colombia y los Anales de la Academia de Medicina de Medellín, en estas se publicaban los resultados de trabajos de investigación sobre enfermedades locales.

En lo que respecta al siglo XX, la naciente industria moderna posibilitó a la ciencia un mayor vínculo con el Estado, en cuanto proporcionaba elementos necesarios para el crecimiento de la economía nacional. De este modo, los estudios en ingeniería, geología, minería, química y economía entran en auge, al principio bajo la dirección de científicos extranjeros, y luego, de colombianos que habían completado sus estudios en Europa. Este periodo se destacó durante sus primeros años por la consolidación de la geología en distintas instituciones como la Oficina de Longitudes, dirigida por Ricardo Lleras Codazzi, la Comisión Científica Nacional, el Servicio Geológico Nacional, Escuela Nacional de Minas, el Laboratorio Químico Nacional, el Instituto Geográfico Nacional, la Escuela Normal Superior; más tarde se caracterizó por la diversificación de profesiones en el terreno educativo, incluidas las ciencias sociales; y finalmente, en los años 50s del siglo, los procesos de institucionalización científica fueron orientados al acelerado proceso de transformación económicas, social y cultural (Melo, 1987).

Con relación a lo anterior, los documentos producidos durante los dos primeros siglos expuestos (XVIII y XIX), eran almacenados en archivos extranjeros, fenómeno históricamente explicable como lo relata Augusto Gómez (1986), pues ello expresa el caos administrativo del país en aquellos años y el poco interés en la producción científica. Por otro lado, poseían la característica de ser, en su mayoría, información restringida, pues se tenía conciencia de que los documentos eran fuentes de información para el poder. Aun así, a lo largo de la historia el hombre ha tenido la necesidad de transmitir la información compilada en los documentos, por lo que ha recurrido a diferentes soportes como el papel.

En concreto, es determinante para el proceso de investigación científica dar a conocer sus hallazgos, resultados y avances, puesto que uno de los postulados fundamentales de la ciencia es el de ampliar el conocimiento que se tiene respecto a un problema. Por esta razón la publicación y difusión son quizá el elemento más contundente de la producción científica,

debido a que es común la creencia en el ámbito investigativo, de que lo que no es visible entonces no existe.

De acuerdo con esa noción de visibilidad, está también la necesidad de preservación de aquellos escritos que contienen evidencias de valor histórico, artístico, científico, literario y cultural en tanto representan el testimonio sobre sucesos, hechos o acontecimientos que han sustentado la vida y quehacer del hombre. En esta medida, la conservación de bibliografía de los documentos científicos hace parte de la memoria colectiva (Gutiérrez, Serra, Hinojosa, & Luis, 2014), no sólo de los investigadores que se entregaron a la ciencia en diferentes épocas, sino también de quiénes se interesan aún por los resultados.

La consulta de documentos científicos antiguos tiende a ser más ardua que la de los recientes; en primer lugar, porque muchos de ellos son impresos, por lo cual se ven expuestos al deterioro propio de los años (que los hace ilegibles o frágiles al tacto), o condiciones de ambientes de las bibliotecas, casas museo u otros; en segundo lugar, porque a veces el documento tiene un solo ejemplar, ocasionando que su préstamo sea restringido y/o muy estricto. Ante esta dificultad es imprescindible contar con otras herramientas que permitan el acceso al conocimiento y con ello, al desarrollo, divulgación, preservación y recuperación de material científico.

Actualmente, las alternativas diferentes al material impreso se han ido ampliando en forma de dispositivos electrónicos, internet, repositorios, bases de datos bibliográficas (académicas, científicas, literarias, etc); medios digitales que en su mayoría facilitan la consulta de diferentes temas y escritos, entre ellos, lo concerniente a la ciencia

Ahora bien, los medios digitales cambiaron, la manera en la que estas características generales de producción y circulación del conocimiento se implementan. El tamaño, la velocidad y la interconectividad de las producciones académicas se han incrementado tremendamente. Los medios digitales proveen plataformas para producir y almacenar mucha más información en menos espacio y tiempo; lo cual a su vez permite acceder a esa información de forma más rápida que lo que es posible en la dinámica del libro o la revista impresos en papel. Esto se hace a través de herramientas como bases de datos y bibliografías electrónicas (...) (Vargas, 2014).

Así mismo, con la celeridad de las herramientas digitales surgen alternativas que facilitan la conservación y consulta de los documentos científicos históricos, una vez han sido digitalizados y su información ha sido indexada, son almacenados, para el caso concreto, por bases de datos que compran el derecho de estos documentos. En la actualidad los artículos científicos en soporte digital tienen problemas de accesibilidad y la facilidad de uso del contenido de los documentos digitales recae en la forma en la que son publicados. Las bases de datos son herramientas que nos permiten consultar, seleccionar, analizar o relacionar información de diversos contenidos relacionados con distintas actividades, ya sean científicas, de investigación, académicas o cotidianas se encuentran organizados para cumplir un propósito específico que obedezca a las necesidades de información de una comunidad determinada de usuarios (Gil, 1994). En el caso de una base de datos de bibliográfica, específicamente, son dispositivos de almacenamiento y procesamiento de fácil, eficiente y oportuno acceso al conocimiento; que hoy por hoy benefician a investigadores, estudiantes, profesores, y todo tipo de usuarios, de todo el mundo, en la consulta de sus temas de estudio. Para el caso de la ponencia se hará evidente la información encontrada en la base de datos Archive, además de lo que significó la identificación de los documentos científicos en ella.

De esta manera, adquirir información científica tomada de las bases de datos bibliográficas, es el recurso principal disponible en el momento de identificar los antecedentes históricos en la ciencia en Colombia, y en este caso, de construir un Atlas Histórico de la Ciencia Colombiana, puesto que permite recuperar el acervo de conocimiento existente, aquel sobre el cual se tienen registros o fueron considerados tan relevantes que hoy constituyen piezas patrimoniales.

### **Marco referencial**

Si bien el Atlas Histórico de la Ciencia Colombiana cuenta con la construcción de un marco teórico que de cierto modo sustenta el proyecto, en la presente ponencia se desarrollará un marco referencial que dé cuenta de manera más precisa del objeto que le concierne. La ponencia se centra en la identificación de documentos científicos en 65 bases de datos digitales, por tal motivo se hace importante recurrir a una definición de lo que se entenderá por humanidades digitales, documentación científica y fuentes digitales.

### **Humanidades digitales:**

Tal y como lo plantea Isabel Galina (2011), las humanidades digitales (HD) buscan entender el impacto de las tecnologías de cómputo en el quehacer de los investigadores de las humanidades y tiene como uno de sus objetivos la creación de bases de datos con documentos relevantes, lo que a su vez alude a la preservación de los mismos. Es importante resaltar que las humanidades digitales no solo se concentran en las humanidades, sino que también abarcan un sinnúmero de disciplinas que generan que cada vez se consolide más, lo que ha hecho que alrededor de este tema surjan posgrados, congresos y centros especializados que contribuyen, a su vez, a su institucionalización. Tal vez la importancia de las humanidades digitales radica en que a lo largo de su historia se han convertido en un elemento central en las investigaciones científicas “Hoy día, prácticamente la totalidad de la información primaria de las ciencias se produce en formato digital” (Galina, 2011, p. 7).

Desde nuestro análisis las humanidades digitales hacen, de cierto modo, alusión al término documentación científica, usado con frecuencia por los académicos interesados en encontrar en la web documentos científicos agrupados, por lo general, en bases de datos digitales. Por su lado, desde las humanidades digitales se encuentra que mucha información física (no digital) ha sido transferida al mundo web lo que ha generado un incremento considerable de documentos científicos en plataformas digitales “...existen hoy en día un gran número de bases datos con información científica” (Galina, 2011, p.8). Los estudios sobre humanidades digitales son relativamente nuevos y la mayoría apuntan a tratar de encontrar una definición adecuada, lo cual se traduce en un gran interés por incorporar lo científico y social en el ámbito digital. Ahora, es importante en este recorrido por las humanidades digitales referirnos a una serie de definiciones que se tuvieron en cuenta en la presente ponencia, para empezar, Nuria Rodríguez (2013) en su texto “Humanidades digitales, Digital Art History y cultura artística: relaciones y desconexiones” define las humanidades digitales como “...el espacio de convergencia entre ciencias de la computación, medio digital y disciplinas humanísticas” (p. 18), en la búsqueda de nuevos paradigmas de conocimiento, lo que no se aleja de lo que se plantea en el Atlas Histórico de la Ciencia Colombiana (AHCC) referente a su interés por recurrir a las fuentes digitales como otro campo de conocimiento para la identificación de documentos científicos. Por su lado, Antonio Rojas (2012) la define de una forma más diacrónica apelando a su evolución a través del tiempo, para este autor si bien la fundación de las humanidades digitales se dio de forma específica en el humanismo a partir de la creación del Index Thomisticus en 1949, trabajo que combinaba tradiciones humanísticas con innovaciones técnicas, hoy por hoy las humanidades digitales apela a la interdisciplinariedad trastocando dicotomías entre ciencias y humanidades, entre lo innovador

y lo tradicional, definiéndola así "...como un conjunto de principios, valores y prácticas en donde convergen múltiples objetos de estudio y saberes"(Rojas, 2012, p.79) que para el caso concreto del AHCC las discusiones giran en torno a los postulados de la historia, la sociología, la bibliotecología y la filosofía. Otra definición que nos ocupa por su analogía con lo propuesto desde lo trabajado en el proyecto AHCC es la de Nicanor Ursua (2016) quien se refiere a las humanidades digitales como "una "transdisciplina" portadora de los métodos, dispositivos y perspectivas heurísticas relacionadas con procesos de digitalización en el campo de las Ciencias Humanas y Sociales" (p.34), afirmando así que las HD se caracterizan por el uso de plataformas digitales para hacer de los documentos objetos más accesibles.

Las humanidades digitales han tenido tal impacto que han llevado a que las ciencias sociales y las humanidades se cuestionen sobre el valor de sus investigaciones en la era de lo digital, generando una transformación en las metodologías y formas de difusión de lo que se produce. Así mismo, las HD presuponen nuevos modos de hacer ciencia con el uso de plataformas digitales para la colaboración, comunicación y preservación entre los autores y el material producido. El definir las humanidades digitales es una cuestión que todavía ocupa gran parte de los debates teóricos, pero tal vez, su característica más importante es la intersección del mundo digital con el cúmulo de conocimientos y prácticas de las ciencias sociales y humanas. Cabe resaltar que uno de los intereses por las humanidades digitales en esta ponencia se refiere a la visibilidad que se le da en la web a documentos científicos históricos que habían sido relegados y olvidados, al tiempo que generan un trabajo interdisciplinar creando redes y cambios en la concepción de conocimiento.

### **Documentación científica y Documento científico:**

En su trabajo "Introducción a la Documentación Científica" Tomàs Baiget (2007) define la documentación científica como la actividad que se ocupa de recoger y tratar los conocimientos, almacenar la información en documentos y facilitar su recuperación para contribuir a la difusión de los mismos, afirmando que sin documentación no hay ciencia, en tanto para que esta sea debe estar escrita, como forma de perpetuarse en el tiempo. En la actualidad existen nuevas facetas de la documentación científica como la publicación electrónica, que se traduce la mayoría de las veces en bases de datos.

Por otro lado, Juan Ros García (s.f) define la documentación científica como aquella parte de la disciplina documental que tiene por objeto de estudio el proceso de transmisión de las fuentes para la obtención de nuevo conocimiento científico. El término engloba el carácter científico de la disciplina documental y argumenta porque esta no es por sí sola, teniendo

presente que la documentación por si misma solo tiene un nivel operativo. La documentación científica desde sus múltiples definiciones hace alusión a la recolección y almacenamiento de información científica para después ser recuperada y consultada. Ros, en su trabajo sobre la documentación aclara que "...la documentación científica como parte de la disciplina documental en general, es especializada, maneja documentos científicos y trata de obtener nuevos documentos que transporten mensajes científicos" (Ros, s.f, p. 338).

La forma en que es posible acceder, producir y circular información sobre un tema de interés, una disciplina o un acontecimiento, se ha visto a lo largo de la historia, sujeta a cambios tan significativos e irreversibles como los de la ciencia; pues lo que antes podía responderse con la lectura de un papel escrito del puño y letra de un autor, luego pasó a ser conocido, sucesivamente, por intermedio de la máquina de escribir y luego la computadora; de igual manera sucedería con los papiros, libros, cassettes, cds, memorias; o con las bibliotecas, centros de documentación, repositorios, internet, entre otros.

En efecto, encontramos hoy otras formas de acceso al conocimiento e información, y una cantidad enorme de nuevas tecnologías de comunicación (Ruiz, 2016) que han ido modificando los hábitos de búsqueda y recuperación. Los productos y técnicas tradicionales como los repertorios impresos pasan segundo plano cuando existen alternativas fáciles, rápidas y cómodas como las que ofrece los recursos digitales en la localización de material para una investigación (Merlo, 1997)

Es así como el uso del documento digital y la documentación digital como actividad, se encuentran en continuo crecimiento, en tanto cumplen con varios propósitos orientados a la organización, preservación, recuperación, agrupación, clasificación, circulación y acceso libre al conocimiento (Florian et al., 2007). La proliferación de fuentes digitales de conocimiento se da en el marco de los riesgos asociados a la pérdida de información que se deposita en otros mecanismos de difusión como textos, grabaciones, imágenes, etc. Mientras que el recurso digital hace uso de tecnologías más duraderas, mayor alcance, ingreso desde casi cualquier lugar, menores costos a futuro, etc.

El mayor valor de la documentación digital lo otorgan las diferentes medidas de preservación y conservación para cada uno de los tipos documentales (filmes, fotos, libros, artículos, etc) que están considerados dentro del patrimonio histórico de la humanidad. Se aclara entonces que la conservación pretende salvaguardar y asegurar el acceso a la información tanto de las generaciones presentes como las futuras, y esto incluye labores curativas y de restauración (Romero, 2016).

Sobre la preservación digital, la Unesco propone acciones para incrementar la conciencia y protección del patrimonio documental y lograr la accesibilidad universal y permanente (Hallo & De la Fuente, 2010). Como Organización de las Naciones Unidas para la Educación, la Ciencia y la Tecnología, declara desde 1995, y reitera en el 2002, la incorporación del programa Memoria del Mundo:

(...) la Constitución de la UNESCO establece que la Organización debe ayudar a la conservación, al progreso y a la difusión del saber, velando por la conservación y la protección del patrimonio universal de libros, obras de arte y monumentos de interés histórico o científico, que su Programa Información para Todos ofrece una plataforma para el debate y la acción sobre políticas de información y sobre la salvaguardia de los conocimientos conservados en forma documental, y que su programa “Memoria del Mundo” tiene por objeto garantizar la preservación del patrimonio documental del mundo y un acceso universal al mismo, reconociendo que esos recursos de información y expresión creativa se elaboran, distribuyen, utilizan y conservan cada vez más en forma electrónica, y que ello da lugar a un nuevo tipo de legado: el patrimonio digital, Consciente de que el acceso a dicho patrimonio brindará mayores oportunidades de creación, comunicación e intercambio de conocimientos entre todos los pueblos (UNESCO, 2003, p. 1).

En esta misma publicación, la organización declara

El patrimonio digital como recurso único fruto del saber o la expresión de los humanos, que comprende recursos de carácter cultural, educativo, científico, administrativo, técnicos, jurídicos, médicos y de otras clases, que se generan directamente en forma digital o se convierten a éste a partir de un material analógico ya existente. Estos productos digitales existen solo en formatos electrónicos y pueden ser textos, bases de datos, imágenes, videos, grabaciones, material gráfico, páginas web, ect, que requieren de procedimientos específicos para su conservación, producción, mantenimiento y gestión” (UNESCO, 2003, p. 2)<sup>1</sup>.

A propósito de los fines de conservación y protección, hoy es factible el continuo crecimiento e innovación de la documentación digital, ya que constituye herramienta de apoyo a todas las

---

<sup>1</sup> Aunque en Colombia no existe una ley que hable directamente de la preservación y conservación del patrimonio digital como la propone la Unesco, si existe una normatividad, como la Ley de Cultura de 1997, que hacer referencia a todos los bienes y manifestaciones materiales e inmateriales producto de la expresión de la nacionalidad colombiana (Universidad de Antioquia, s.f). Asunto que no contradice el postulado de la Unesco en que el patrimonio digital brindará mayores oportunidades de “preservar y poner a disposición de cualquier persona el patrimonio de todas las regiones, naciones y comunidades a fin de propiciar, con el tiempo, una representación de todos los pueblos, naciones, culturas e idiomas” (UNESCO, 2003, p. 2).

actividades científicas, educativas, culturales y económicas, y permitiendo el acceso libre a un sinfín de temas relacionados con ellas.

### **Bases de datos bibliográficas:**

En lo que se refiere a la importancia de la información y el cómo se crean estrategias para su recuperación, difusión y circulación, se presenta la utilización de nuevas tecnologías para el manejo de un cúmulo de datos incalculable, de tal forma que éstas se encuentran habilitadas para consultar, descargar y analizar las revelaciones de distintas actividades del saber humano. De ahí que, actualmente las bases de datos bibliográficas tengan una importante trascendencia como artefacto de almacenamiento y procesamiento de la información (Gil, 1994), dado que, además de preservar fuentes originales, primarias y secundarias, también son útiles para que los usuarios puedan informarse, intercambiar y desarrollar un efectivo conocimiento u obtenerlo de manera oportuna, precisa, relevante y bien gestionada (García, s.f). Adicionalmente, las bases de datos bibliográficas están compuestas por grandes colecciones de documentos académicos y científicos del pasado y del presente, disponibles de forma electrónica, en distintos formatos, a toda hora y desde cualquier lugar del mundo. Igualmente, las bases de datos bibliográficas tienen la labor fundamental de “identificar los textos, certificando su existencia, mediante la aportación comprobada de los datos referenciales que se refieren a él; de otro, informar sobre los temas, dando referencias de los documentos relativos a una cuestión” (Mikelarena, 2002, p. 86).

### **Metodología**

La construcción del Atlas Histórico de la Ciencia Colombiana tiene implícito la responsabilidad de identificar y revelar el acervo de producción científica publicada entre el siglo XVIII, XIX y finales del siglo XX, de ahí que, bajo la orientación del Cinvestav (Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional, México), un equipo de profesores, estudiantes y profesionales de diferentes disciplinas de la Universidad de Antioquia y la Universidad Pontificia Bolivariana realizáramos una serie de exploraciones y acercamientos a bases de datos científicas, teorías y métodos de los cuales fuera posible obtener información para la elaboración y solidez del proyecto.

El proyecto del Atlas Histórico se ha ido construyendo a partir de varios momentos, que se describirán más adelante. Para empezar, es importante hablar de la ruta metodológica abordada previo al inicio del Atlas, que se basó en una serie de reuniones realizadas con los

encargados de elaborar el Atlas de la Ciencia Mexicana donde se tomaron decisiones basadas no solo en la experiencia que se había tenido en México con la realización de su Atlas, sino también en las discusiones teóricas que giraron alrededor de la sociología de la ciencia, historia de Colombia y geografías del conocimiento. Posterior a ello se empezó a formular lo que se quería alcanzar con la realización de un atlas que abordará la documentación científica histórica de Colombia y sobre Colombia, y en síntesis el objetivo planteado estaba ligado a la identificación de información histórica mediante la exploración de fuentes digitales.

Ahora, el primer momento de la investigación se caracterizó por ser una etapa de exploración virtual y referencial de bases de datos bibliográficas autodenominadas científicas, debido a que éstas proporcionan información sobre las producciones científica a nivel mundial, y en el caso específico de Colombia, cuáles de ellas tienen contenido relevante para el desarrollo de un Atlas Histórico de la ciencia sobre el país. En concreto, la consulta se realizó con un volumen inicial de 124 bases de datos internacionales (Scopus, Web of Science, Microsoft Academic, Scielo, Semantic Scholar, OpenEdition.org, Isidore, Jstor, Archive, Directory of Open Access Journals, Google Scholar, CINAHL: Cumulative Index to Nursing and Allied Health, National Diet Library Collection etc), de las cuales fueron elegidas solo 65 para explorar con mayor profundidad, puesto que la elección se dio gracias al primer filtro que consideramos esencial para el desarrollo del AHCC.

El criterio de selección inicial consistió en identificar aquellas bases de datos que comprendieran documentos en los que se hallaran las distintas denominaciones que tuvo el país según el momento histórico (Nueva Granada, Confederación Granadina, Provincias Unidas de Colombia, República de Colombia), además de la delimitación temporal en que fueron publicados los documentos (1800-1979), acotación planteada al considerar el periodo de tiempo en que los textos históricos obtuvieron licencia para ser publicados libremente. Fueron descartadas las bases de datos que solicitaran inscripción generando costo, requerían password, no tienen contenido científico y no correspondieran con el rango de fechas establecido.

Respecto a la segunda etapa —que se fue desarrollando paralela a la primera etapa— del proyecto AHCC, se llevó a cabo la construcción de una matriz analítica que comprendiera categorías sociológicas, históricas y bibliométricas con el fin de tener un control bibliográfico de la información que se fuera recolectando. La matriz fue elaborada de acuerdo al formato Dublin Core, que es un modelo de metadatos, y se contó con la guía constante del Cinvestav. Los campos aprobados en el formato son: colaborador, cobertura, creador, fecha, descripción, formato, identificador, idioma, editor, relación, derechos, fuentes, asunto, título, fuente y 132

etiquetas más; esta etapa del proyecto fue desarrollada en gran medida por los bibliotecólogos adscritos al proyecto. El objetivo de la estructuración de metadatos en el formato Dublin Core para el AHCC está pensado en la carencia de metadatos estructurados disponibles para los documentos históricos queriendo presentar lo que se construya al final como un modelo que permita la descripción detallada de los documentos.

En la actualidad el proyecto Atlas Histórico de la Ciencia Colombiana está en el proceso de recuperación de textos y metadatos en las bases de datos digitales, de ahí el interés de esta ponencia por indagar sobre la importancia de las fuentes digitales en la recuperación de documentos científicos históricos, que a su vez comprende el contexto futuro de la exploración de la totalidad de las bases de datos. Para el caso de la ponencia, la metodología que se planteó

## **Resultados**

A partir de la exploración realizada en las bases de datos bibliográficas, por medio de las opciones de “búsqueda simple” y “búsqueda avanzada”, en las cuales la palabra clave que se empleó hacía referencia a cada una de las denominaciones históricas de Colombia, fue posible obtener los siguientes hallazgos:

- De las 124 bases de datos bibliográficas examinadas sólo 65 arrojaron resultados en el momento de ingresar las palabras “Colombia”, “Confederación Granadina”, “Nueva Granada” y “Provincias Unidas de Colombia”; las restantes 59 bases de datos fueron descartados por no contener información sobre nuestro país o en caso de contenerla, se encontraba por fuera del rango de fechas establecidos para la elaboración del AHCC; es decir, muchas de esas publicaciones correspondían a periodos de 1980 hasta el año 2016; contrario a las pretensiones del proyecto que consideró como delimitación temporal 1979 y precedentes.
- Entre las 65 bases de datos que sirven al propósito del AHCC, se encontraron 390.000 registros con la palabra Colombia, 493 para Confederación Granadina, 30.355 de Nueva Granada y 1346 por Provincias Unidas de Colombia. Es de aclarar que el número de resultados obtenidos por cada uno de estos nombres corresponde a la cantidad de veces que la palabra clave ha sido mencionada en el material disponible por la base de datos, sin que pueda determinarse si el contador de palabras puede discriminar el hecho de que la denominación esté implícita en el título, subtítulo,

resumen o contenido del expediente; y en algunos casos, tampoco puede comprobarse si hablan propiamente sobre el país, a menos que se descargue cada uno de los tipos documentales para su posterior revisión.

Después de ejecutar el filtro y continuar la realización del Atlas con 65 bases de datos, se procede a dar un contexto de cada una de ellas con el fin de determinar el tipo de información que contenían, antes de iniciar con la inspección de los documentos que comprende cada base de datos. Posterior a la contextualización se obtuvieron resultados que permitieron, de cierto modo, discernir las temáticas más recurrentes; para empezar, gran cantidad de las bases de datos archivan documentos que responden a múltiples disciplinas, como: Bielefeld Academic Search Engine (BASE), Scopus, Scielo, Google Scholar, GENESIS, Directory of Open Access Journals, Microsoft Academic, Zasshi Kiji Sakuin, WorldCat, Ulrich's Periodicals Directory, SafetyLit, Science.gov, ScienceOpen, Socol@r: Socolar, SpringerLink, Book Review Index Online, Web of Science, Archive, Luis Ángel Arango Hemeroteca, Luis Angela Arango Libros Digitales y LaReferencia, otras bases de datos se dedican a la preservación de documentos más específicos, por ejemplo algunas sólo contienen información sobre medicina, enfermería y temas afines a la salud: CINAHL, Europe PMC, POPLINE, PubPsych, PubMed, FSTA – Food Science and Technology Abstracts y Public Health Reports, las ciencias sociales y las humanidades acogen otro número de bases de datos: OpenEdition.org, AgeLine, Isidore, CiNii, PsycINFO, Educational Resource Information Center (ERIC), Questia: Online Research Library, Anthropological Index Online, así como la informática, la tecnología y las comunicaciones: The Collection of Computer Science Bibliographies, Semantic Scholar, Association for Computing Machinery Digital Library, IEEE Xplore, Aerospace & High Technology Database, hay otro pequeño grupo de bases de datos que están conformadas por documentos que aluden a disciplinas más específicas como MathSciNet que trata de física, matemáticas, probabilidad y áreas afines y Chemical Abstracts Service que contiene documentos sobre química o para finalizar, Astrophysics Data System que es una base de datos financiada por la NASA y contiene documentos sobre astrofísica.

Es importante resaltar que en términos metodológicos surgió la necesidad de definir no sólo la categoría *base de datos bibliográfica*, sino también *metabuscadores*, *agregadores de bases de datos bibliográficas*, *índices bibliográficos* y *repositorios* teniendo en cuenta que la mayoría de éstos sufren variaciones no sólo en términos de su contenido sino también de su estructura. Así, por ejemplo, se dijo que los *metabuscadores* se encargan de agregar los

resultados de varios buscadores y no tienen bases de datos propias y los *agregadores de bases de datos bibliográficas* hacen mención a las plataformas que integran varias bases de datos, mientras que los *índices bibliográficos* permiten el acceso a varias bases de datos y su contenido es multidisciplinar y por su lado, los *repositorios*, son medios que almacenan, preservan y facilitan el acceso a documentos digitales, para el caso concreto del Atlas contamos con repositorios institucionales. Por lo anterior es que existe la posibilidad que durante la exploración de cada una de las bases de datos nos encontremos con fuentes de información repetidas, hipótesis que fue corroborada en el desarrollo de la presente ponencia donde se realizó la búsqueda de documentos científicos por la base de datos Archive, pero de esto se hablará más adelante.

Durante el desarrollo de esta ponencia se hizo indispensable la revisión de una base de datos, todo con el fin de exponer lo que se tiene planeado hacer con la totalidad de las bases de datos bibliográficas. Archive fue la base seleccionada por la cantidad de documentos científicos que tratan sobre Colombia, las variadas disciplinas y su compromiso con la preservación de los documentos históricos.

## **Archive**

Archive se define a sí misma como una biblioteca sin fines de lucro que contiene millones de textos, videos, música, imágenes y películas. Sus colecciones principales están conformadas por audios, música y videos, seguidas por audiolibros e imágenes, que por intereses investigativos no hacen parte de la búsqueda que se realizó en esta base de datos; las bibliotecas americanas y canadienses, por su parte, suman casi 3 millones de artículos, casi el doble del total de la colección de revistas. En total la base de datos contiene más de 18 millones de textos dentro de los cuales se encuentran libros, artículos y revistas, para el caso concreto se encuentran 4,617 textos que traen implícito ya sea en el título, resumen, colección, nombre del creador, en su bibliografía y/o referencias las denominaciones sobre Colombia, lo que equivale al 0,025% del total de textos de la base de datos.

Ahora, como ya se mencionó, el interés recae en la identificación de documentos científicos que sean sobre Colombia o que hayan sido realizados desde Colombia, lo que implica que al identificar los documentos haya que dirigirse al resumen del texto para saber si con seguridad el documento está cumpliendo con alguno de los dos requisitos, posterior a ello hay que identificar lo que sería el factor más importante, que a su vez es el objeto de esta ponencia, el documento científico, para saber si el resultado encontrado cumple con los parámetros de un

documento científico sin revisar aun cada uno de ellos se recurre a identificar su temática, metodología y quién lo publica (por ejemplo si está en una revista científica). Para el caso de los documentos más antiguos que carecen de algunos de estos parámetros se hace necesario dirigirse al resumen o parte del texto con el fin de identificar en este lo que se busca desde el proyecto.

Posterior a la identificación del documento se extraen de él una serie de variables que con seguridad en un futuro facilitarán el manejo de la información en otras fases del proyecto, las variables corresponden a las etiquetas normalizadas en Dublin Core: el título del documento, el autor, el año de publicación, el lugar de publicación, los temas que trata, el tipo de fuente, el idioma, las opciones de descarga (los formatos), la entidad o entidades que lo publican, el editor, las filiaciones y los metadatos, cabe anotar que no todos los documentos contienen la totalidad de las variables, pero no por ello se descartan, pues se tiene presente que al ser documentos históricos no contemplaban todos estos campos. Luego de efectuar cada uno de estos filtros en los documentos, se procede a descargarlos y guardarlos con un código, asignado en la base de datos de excel que es donde en un primer momento se guarda toda la información de los documentos. Los formatos más encontrados en Archive para la descarga de los documentos son PDF y HTML afirmando lo planteado por Térmens et al. (2008) de que dos de los formatos más utilizados para la publicación de artículos científicos en soporte digital son: HTML y PDF.

El dato más antiguo encontrado en la base de datos es del año 1673 y es un libro sobre la historia del Nuevo Reyno de Granada, escrito por Juan Flórez de Ocariz y publicado en Madrid el libro lleva como título “Libro primero de las genealogías del nuevo Reyno de Granada” está escrito en español y fue descargado en PDF, por el contrario, el archivo más reciente comprende un estudio sobre el descubrimiento de antiguas plantaciones de guayusa en el país, el documento se titula “Discovery of an ancent guayusa plantation in Colombia”, de 1979 está escrito en inglés y es un artículo de revista, lo que permite observar cómo las formas de difusión del conocimiento han ido cambiando.

En esta base de datos se obtuvieron varios hallazgos, uno de ellos concierne a la temática de publicación según el siglo, es decir, para el siglo XVIII se encontró que la mayoría de los documentos científicos respondían a los siguientes temas: historia, geografía, agricultura y topografía, todos correspondían a libros y eran publicados en Europa, el siglo XIX inicia con una tendencia en tanto a la producción de documentos muy parecida a lo que había ocurrido durante el siglo XVIII, pero es a mediados de siglo cuando empieza a notarse un cambio, la mayoría de documentos están escritos en inglés, las investigaciones etnograficass toman

importancia pero la historia sigue siendo la temática principal con la variante del idioma de publicación y el tipo de fuente, dejan de encontrarse tantos libros y empiezan a tomar importancia los artículos de revista. El siglo XX inicia con un panorama muy parecido a su siglo antecesor, las principales temáticas se encuentran en la historia, con la diferencia de que son memorias históricas lo cual permite ver que las investigaciones a inicio de este siglo (desde lo evidenciado en Archive) no fueron muy dinámicas, lo que si es cierto es que el idioma ingles es cada vez más recurrente en los documentos y que los libros van pasando a un segundo plano; a mediados de siglo otras temáticas investigativas empiezan a cobrar protagonismo, ahora son las ciencias sociales y las humanidades las que abarcan gran parte de los documentos: sociología, antropología y psicología, por su parte la filosofía, las ciencias políticas, los estudios de suelos, y la botánica tienen gran parte de estudios, ya no son solo artículos ahora se evidencia que las tesis aumentan, a diferencia de otros siglos Europa deja de ser el lugar predilecto para publicar y es Estados Unidos quien se queda con el puesto publicando mayormente en ciudades como Washington, Massachusetts y Florida, otro cambio notorio se refiere al tipo de autor, ahora son los autores corporativos los que más registran en los documentos.

Fueron varios los elementos que se identificaron en la búsqueda por Archive referentes a las fuentes digitales, para empezar se puede plantear que la facilidad de uso del contenido de los documentos científicos digitales depende en gran medida de la forma en la que se publiquen, con ello refiriéndonos a la cantidad de información que brinden sus metadatos, teniendo presente que su contenido responde a una serie de decisiones de sus autores, por ejemplo, según el formato de archivo concreto en que hayan sido subidos a la base de datos tienen un potencial de accesibilidad que puede ser explotado en mayor o menor grado. Por otro lado, nos encontramos con que la totalidad de los documentos de Archive tenían acceso abierto lo que demuestra la importancia y beneficios de lo digital en el acceso al conocimiento.

### **A modo de conclusión**

Las primeras etapas de la construcción del Atlas histórico de la Ciencia Colombiana se han ido desarrollando transversales a una serie de obstáculos, para empezar, algunas de las diferentes denominaciones que ha tenido el país a lo largo de la historia llevan implícito el término “Colombia” (Estados Unidos de Colombia, La Gran Colombia, República de Colombia) lo que generó en un inicio que el total de resultados variara considerablemente. Por otro lado, al realizar los filtros de la información que se quería encontrar aparecía en los

resultados muchas veces el mismo documento, ya fuera porque su editor cambiaba o porque lo habían subido a la plataforma distintas personas, de igual manera en la búsqueda de los documentos nos encontramos con que Colombia no siempre era el país, en ocasiones los resultados arrojados por la base también contemplaban nombres de personas, haciendas y/o barrios con el término “Colombia”. Estos obstáculos generan que la identificación del documento científico en las bases de datos tome más tiempo y por ende sea un proceso más complejo.

Como se dijo a lo largo de la ponencia México ha sido guía y ficha clave en la construcción del AHCC, a diferencia de ellos quienes identificaron un hilo temático (el descubrimiento del eritronio conocido hoy como vanadio, en 1801) que les sirviera de inicio para considerar cuáles habían sido los primeros acercamientos a la ciencia en su país, por ejemplo, ellos consideran una carta el primer documento científico hasta ahora hallado. Aun así, la búsqueda de documentos científicos por las diferentes bases de datos nos ha dado indicios de cuáles serían los temas predominantes en Colombia para cada momento histórico, lo que puede significar un acercamiento a algún tema. Ahora, es importante nombrar que la identificación de documentos científicos en las diferentes bases de datos también recoge el interés por saber de qué manera se plasma en los textos los objetos de estudio, los procedimientos, métodos y resultados según el momento histórico y cómo a su vez se hacia la difusión de los mismos. En la búsqueda del documento científico por Archive se hizo evidente la escasez de trabajos en bases de datos digitales pues la mayoría de los archivos se han hecho en bibliotecas físicas, tal vez porque se ha desconocido un poco la importancia y la visibilidad que generan las plataformas digitales.

No obstante, a través de Archive, hasta ahora la base de datos explorada para la construcción del Atlas, fue posible identificar una dinámica de internacionalización de la producción científica colombiana, puesto que, los documentos publicados en ella pueden ser consultados desde cualquier lugar del planeta, otorgando considerablemente, mayor visibilidad y acceso al conocimiento, desvaneciendo las fronteras espacio temporales entre países. Aunado a ello está el hecho de que los documentos científicos encontrados en las bases de datos, además de preservar los frutos del saber humano, se han convertido en una herramienta de reconocimiento para los investigadores, científicos, escritores, centros y grupos de investigación, que de esta forma hacen valer y legitimar sus aportes como construcciones propias del campo científico.

## **Bibliografía**

- Baiget, T. (2007). Introducción a la Documentación Científica. Universidad de Salamanca. Recuperado de: [http://sabus.usal.es/bib\\_virtual/doc/baiguet.pdf](http://sabus.usal.es/bib_virtual/doc/baiguet.pdf)
- Becerra Ardila, D., & Restrepo Forero, O. (1993). Las Ciencias en Colombia. 1783-1990. Una perspectiva histórico-sociológica. *Revista Colombiana De Educación*, (26). <https://doi.org/10.17227/01203916.5296>
- Florian, B., Valencia, M., Rodríguez, P., Millán, M., Gaona, C., Carrillo, J., & Mauricio, C. (2007). Diseño de una plataforma experimental para la búsqueda y recuperación de documentos en una biblioteca digital. *Ingeniería y competitividad*, 105-117.
- Galina, I. (2011) ¿Qué son las humanidades digitales? UNAM. Recuperado de: <https://www.oei.es/historico/divulgacioncientifica/?Que-son-las-Humanidades-Digitales-Isabel-Galina-Russell>
- García, E. (s.f). *Bases de datos bibliográficas. Herramientas para profesionales y técnicos de la salud*. Obtenido de Universidad Virtual de la Salud Manuel Fajardo: <http://uvsfajardo.sld.cu/bases-de-datos-bibliograficas-herramientas-para-profesionales-y-tecnicos-de-la-salud>
- Gil, M. d. (1994). La base de datos. Importancia y aplicación en educación. *Perfiles Educativos*, 1-9.
- Gómez, A. (1986). LA HISTORIA COLOMBIANA Y LOS ARCHIVOS. Recuperado de: <https://revistas.unal.edu.co/index.php/maguare/article/view/75-87/16469>
- Gutiérrez, T., Serra, S., Hinojosa, M. d., & Luis, I. (2014). Importancia del desarrollo de la colección de documentos patrimoniales en una institución del sector de salud. *Revista Cubana de Higiene y Epistemología*, 52(3), 418-427. Recuperado el 15 de agosto de 2018, de [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S1561-30032014000300014](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1561-30032014000300014)

- Hallo, M., & De la Fuente, P. (2010). Preservación digital en el contexto de repositorios institucionales universitarios. *Revista Cubana de Ciencias Informáticas*, 4(3-4), 1-15. Recuperado el 15 de agosto de 2018, de <http://www.redalyc.org/articulo.oa?id=378343670001>
- Melo, J. (1987). Historia de la ciencia en Colombia. Conferencia de divulgación leída en Santa Marta, 1987. Recuperado de: <http://www.jorgeorlandomelo.com/hisciencia.htm>
- Merlo, J. A. (1997). Fuentes de información electrónicas para la investigación en biblioteconomía. *Educación y Biblioteca*, 48-56.
- Mikelarena, F. (2002). Bases de datos bibliográficas para la investigación en temas vasconavarros. *Revista internacional de estudios vascos*, 47(1), 79-91. Recuperado el 30 de Agosto de 2018, de <http://www.euskomedia.org/PDFAnlt/riev/47/47079091.pdf>
- Nieto, M. (s.f). Semanario del Nuevo Reyno de Granada. Recuperado de:  
htt  
[p://bibliotecanacional.gov.co/esco/colecciones/bibliotecadigital/publicacion?nombre=Semanaio+del+Nuevo+Reyno+de+Granada](http://bibliotecanacional.gov.co/esco/colecciones/bibliotecadigital/publicacion?nombre=Semanaio+del+Nuevo+Reyno+de+Granada)
- Rodríguez, N. (2013). Humanidades digitales, Digital Art History y cultura artística: relaciones y desconexiones. *Revista de Ciencia Arte y Tecnología*. Recuperado de: <https://artnodes.uoc.edu/articles/abstract/10.7238/a.v0i13.2017/>
- Rojas, A. (2012). Las humanidades digitales: principios, valores y prácticas. Universitat Pompeu Fabra: JANUS
- Romero, M. E. (2016). Conservación de documentos analógicos y digitales. *Anales de Documentación*, 19(1), 255. Recuperado el 2 de septiembre de 2018
- Ros, J. (s.f). El concepto de documentación en la doctrina española. Universidad de Murcia. Recuperado de: <http://www.geocities.ws/lourdesco1/id/idt1doc2.pdf>
- Ruiz, A. (2016). Fuentes digitales y fuentes impresas. Prácticas letradas y plagio en el marco universitario. *Revista Chilena de literatura* (94), 215-230. Recuperado el 2 de septiembre de 2018
- Térmens, M., Barrios, M., Díaz, M., Guasch, D., Ponsa, P., Ribera, M. (2008). Estudio de la accesibilidad de los documentos científicos en soporte digital. *REVISTA ESPAÑOLA DE DOCUMENTACIÓN CIENTÍFICA*, Vol 31, No 4, 552-572

UNESCO. (15 de octubre de 2003). *Carta sobre la preservación del patrimonio digital*.  
Obtenido de portal.unesco.org: [http://portal.unesco.org/es/ev.php-URL\\_ID=17721&URL\\_DO=DO\\_TOPIC&URL\\_SECTION=201.html](http://portal.unesco.org/es/ev.php-URL_ID=17721&URL_DO=DO_TOPIC&URL_SECTION=201.html)

Universidad de Antioquia. (s.f). *Normatividad sobre el patrimonio documental digital*.  
Obtenido de udea.edu.co: [http://www.udea.edu.co/wps/wcm/connect/udea/45ed2286-3d42-4335-a9d3-e57adc0f0cdd/patrimonio\\_digital.pdf?MOD=AJPERES](http://www.udea.edu.co/wps/wcm/connect/udea/45ed2286-3d42-4335-a9d3-e57adc0f0cdd/patrimonio_digital.pdf?MOD=AJPERES)

Ursua, N. (2016). El pensar humanístico frente a las “humanidades digitales”. España:  
Universidad del país Vasco. Recuperado de:  
<http://www.redalyc.org/html/836/83646546003/>

Vargas, E. J. (2014). El conocimiento científico y los medios digitales en la academia: el  
problema de los estándares de calidad. *Civilizar*, 14(26), 149-162. Recuperado el 1 de  
septiembre de 2018, de <http://www.scielo.org.co/pdf/ccso/v14n26/v14n26a09.pdf>