

Preservación documental, curación digital y colaboración. La experiencia de la Fundación Histórica Neogranadina¹

Maria José Afanador-Llach, Universidad de los Andes

Coloquio de Buenas Prácticas en Archivos Históricos, Universidad del Rosario, Bogotá, Agosto 2017

Esta ponencia discute las experiencias de digitalización de la Fundación Histórica Neogranadina y los proyectos de HD que estamos desarrollando para apoyar la investigación, la pedagogía y transformar la comunicación académica en la era digital. Alrededor de los principios de conocimiento abierto, la colaboración, la experimentación, la creatividad y el software libre, la charla reflexiona sobre las siguientes preguntas: ¿Cómo preservar y visibilizar el patrimonio documental manuscrito de la era colonial de forma sustentable y escalable? ¿Cómo enfrentar problemas tales como la falta de catálogos, o su desactualización? ¿Cuál debe ser el rol de los historiadores en la preservación digital del patrimonio documental?

Quiero empezar con un contexto para reflexionar sobre cómo están impactando las tecnologías de la información el trabajo de los historiadores y más ampliamente de las humanidades y las ciencias sociales, y su conexión con las ciencias de la información. Vayamos un poco más atrás en el tiempo.

La transformación del humanismo temprano hacia las disciplinas que conforman hoy las humanidades y las ciencias sociales estuvieron moldeadas por prácticas editoriales que involucraban la recuperación de un conjunto de trabajos de la antigüedad clásica. Más tarde el desarrollo de las lenguas vernáculas ayudó a expandir las posibilidades de la expresión humanística. El humanismo se nutrió de muchas fuentes, pero la transcripción, traducción, edición y anotación de textos fueron su legado. Más tarde la imprenta permitió la estandarización y diseminación del corpus cultural humanístico mientras que permitía avanzar en el desarrollo y refinamiento de las prácticas editoriales. Algunos estudiosos del tema argumentan que la migración de materiales culturales hacia los medios digitales es un proceso que se puede equiparar al proceso en el cual floreció la cultura impresa del Renacimiento y post-Renacimiento.

Hoy los medios impresos han dejado de ser el ámbito dominante en el cual se construye y disemina la información y el conocimiento. La producción cultural de la humanidad está migrando constantemente hacia formatos digitales. Si bien el internet--la infraestructura

¹ Neogranadina es una entidad sin ánimo de lucro creada en Colombia en 2016. Su consejo directivo está conformado por sus fundadores Juan Cobo Betancourt (profesor asistente del Departamento de Historia de la Universidad de California, Santa Bárbara), Santiago Muñoz (profesor asistente del Departamento de Historia de la Universidad de los Andes), y por Natalie Cobo (candidata a doctorado de la Universidad de Oxford) y Maria José Afanador Llach (profesora asistente de Humanidades digitales de la Universidad de los Andes).

tecnológica para transferir datos en una red de computadores distribuidos--se inició hace casi cincuenta años, la World Wide Web--documentos hiperconectados e hipertextuales visibles en navegadores--solo ha estado activa desde los años noventa. Aun así, el volumen de datos y materiales culturales producidos, compartidos y almacenados en el marco de este desarrollo tecnológico es el más alto en la historia de la humanidad.

El volumen de datos y materiales culturales producidos, compartidos y almacenados en la era digital es el más alto en la historia de la humanidad. Para las disciplinas de las humanidades esta revolución digital ha significado la apertura de nuevos espacios para la experimentación, la creatividad, la colaboración, el desarrollo de nuevas metodologías y la generación de nuevos modelos de producción y transmisión de conocimiento. En este proceso, las fronteras entre las humanidades y las ciencias de la información se están redefiniendo.

Acá es importante señalar que un aspecto esencial del conocimiento humanístico ha sido la creación de colecciones, la catalogación y la curación. En pocas palabras la organización del conocimiento. Sin embargo, la acumulación y cuidado de la información, se terminó desconectando del trabajo humanístico. Esto ocurrió con el auge de los medios impresos, la segmentación de las disciplinas de las humanidades y las ciencias sociales y con el surgimiento de instituciones modernas de la memoria (con aproximaciones sistemáticas a la creación de colecciones y a la conservación).

La proliferación de información histórica y los materiales culturales, llevaron además a una separación entre los guardianes de la memoria: archivistas, curadores de museos, catalogadores y bibliotecarios, de los académicos profesionales en campos tales como la historia. En la era digital, sin embargo, nos encontramos ante un escenario de reconexión entre la historia--o más ampliamente las humanidades--y el campo de las ciencias de la información. Esta reconexión asume muchas formas diferentes pero en esencia implica el trabajo interdisciplinario y experimental donde la investigación basada en datos asume un lugar central dado los altos volúmenes de información y datos de los que disponemos hoy en día.

Como lo señala Bethany Nowviskie, la investigación digital requiere de sistemas de producción que requieren colaboración más cercana que nunca entre académicos individuales y técnicos, estudiantes e investigadores postdoctorales, creadores de contenido, diseñadores, archivistas y profesionales de patrimonio cultural que trabajan de forma colaborativa para construir, difundir y preservar nuevos conocimientos y nuevas interpretaciones (Nowviskie, 2011: 169).

Fundación Histórica Neogranadina

La Fundación Histórica Neogranadina es una entidad, sin ánimo de lucro, constituida acá en Colombia. Liderada por historiadores colonialistas de los siglos XVI, XVII y XVIII, Neogranadina tiene como objetivo utilizar nuevas tecnologías para proteger, rescatar y promover el patrimonio histórico documental de Colombia. Nos dedicamos también a

desarrollar proyectos de historia digital utilizando los últimos avances de las humanidades digitales.

Actualmente, Neogranadina se concentra en la digitalización del patrimonio documental manuscrito y bibliográfico de archivos privados e instituciones no estatales en Colombia. A futuro Neogranadina está desarrollando plataformas en Internet para poner este patrimonio documental a la disposición de investigadores, estudiantes, docentes, y el público interesado, alrededor del mundo.

Nuestra aproximación a la digitalización documental buscar innovar a partir de un modelo de alta calidad pero de bajo costo. Desarrollamos proyectos de digitalización a costos muy bajos. Una gran parte del costo de digitalizar es la máquina en si misma y los escáneres comerciales de libros pueden costar miles de dólares. En Neogranadina construimos escáneres de bajo costo lo que ayudan a hacer el acceso a esta tecnología más fácil y a poder escalar proyectos de digitalización en varias regiones del país. De hecho estamos a punto de iniciar un proyecto con el Centro de Memoria Histórica para alquilar nuestros escáneres para la digitalización de periódicos. No vendemos escáneres, los ponemos en uso por medio de convenios o contratos de alquiler en lugares con necesidades de preservación.

El problema que buscamos resolver es el de los costos muchas veces prohibitivos de hacer proyectos de digitalización. Muchas veces estos son ejecutados por empresas privadas que se especializan en la digitalización de este tipo de material con ánimo de lucro, o las instituciones mismas tienen que invertir en escáneres comerciales, y contratar personas para que los operen.

Así, o las personas o instituciones que quieren consultar la documentación tienen que pagar costos de suscripción o de acceso a las empresas privadas que llevaron a cabo la digitalización, o los archivos terminan gastando cantidades de dinero en proyectos de digitalización muy costosos. Más aún, estos escáneres comerciales en su mayoría no están diseñados específicamente para digitalizar documentación histórica, y pueden deteriorar los documentos. Esto nos parece ineficiente e innecesario, y por eso hemos desarrollado una alternativa.

Actualmente tenemos proyectos de digitalización de manuscritos coloniales y libros de la temprana edad moderna en el Archivo Histórico Regional de Boyacá en Tunja. Tenemos también un proyecto de digitalización en el Archivo Central del Cauca que es parte del Centro de Estudios "José María Arboleda Llorente" y en la Biblioteca del Colegio de Misiones en Popayán. Este es una muestra de las imágenes digitalizadas.

Pregunta clave: ¿Cómo se hace un archivo digital?

Problemas con archivos digitales. ¿Cómo nos imaginamos una plataforma digital innovadora? Acá voy a mostrar dos referentes.

Referente de BIA

A diferencia de otros archivos digitales BIA fue diseñado para innovar y experimentar con las fronteras académicas tradicionales. De esta forma, BIA, más que ser una herramienta de

búsqueda, es un portal al mundo de la Europa moderna, en particular a través de colecciones epistolares. La plataforma permite búsquedas avanzadas y los usuarios pueden navegar volúmenes enteros o documentos individuales en búsquedas de persona, lugar, etiquetas de categorías, localización del archivo, lugar de creación y palabras clave. La plataforma también opera como una comunidad académica en la cual se pueden hacer preguntas y pedir ayuda. Permite además anotar y transcribir documentos en línea. En pocas palabras, está diseñado para la investigación académica.

Miremos un ejemplo de cómo está ocurriendo esta reconexión en el caso del trabajo histórico. Este proyecto se llama History Lab, un laboratorio colaborativo entre varias instituciones en Estados Unidos incluyendo la Universidad de Columbia, la Universidad de Brown, la MacArthur Foundation, entre otros. En el laboratorio participan investigadores, diseñadores, científicos e ingenieros con el fin de construir herramientas para preservar la historia. Con la pregunta sobre ¿cómo hacer historia en la era de la computación y la llamada "big data"? El "big data" se refiere a conjuntos de datos extremadamente grandes que se pueden analizar con computadores para revelar patrones, tendencias y asociaciones, relacionadas al comportamiento e interacciones humanas. El proyecto trabaja con cantidades muy grandes de información, en particular metadatos sobre archivos que se mantuvieron clasificados por largos periodos de tiempo en Estados Unidos.

El proyecto utiliza la llamada ciencia de datos, un campo interdisciplinario que involucra métodos científicos, procesos y sistemas para extraer conocimiento o un mejor entendimiento de datos en sus diferentes formas, ya sea estructurados o no estructurados. De esta forma, el proyecto reúne en una sola bases de datos documentos desclasificados para utilizar lenguajes de procesamiento y "machine learning", algo así como aprendizaje automático, para explorarlos.

En pocas palabras este laboratorio busca transformar millones de documentos digitalizados y sus metadatos en datos estructurados y a partir de ahí crear nuevas herramientas digitales para visualización. Por ejemplo, al contar personas y lugares, se puede revisar la frecuencia relativa en que estos han sido discutidos en documentos. También se pueden encontrar temas clave en cada colección, o los documentos más representativos de cada tema. He aquí una de las posibilidades que pueden emerger del trabajo colaborativo entre las ciencias de la información, la computación y las humanidades.

Neogranadina está conceptualizando su plataforma de archivo digital como un repositorio de conocimiento nuevo. En el proceso nos hemos enfrentado a diversos problemas y estamos acudiendo a los principios de la colaboración y la experimentación para buscar resolverlos. El problema de la falta de catálogos.

El proyecto de Catalogación colaborativa buscar catalogar miles de documentos del fondo Notaría Segunda que no cuentan con un catálogo. Se hizo un llamado a voluntarios, investigadores, paleógrafos, archivistas y estudiantes de historia para participar en el proceso de descripción documental. La inscripción se lleva a cabo en nuestra página Web. Tenemos

además otros proyectos: Edición Digital de la Recopilación de las Leyes de los Reinos de Indias; enfoque en metadatos que permitan crear bases de datos relacionales para investigación con herramientas digitales; Cartografías del Nuevo Reino de Granada.

Retos a futuro

Neogranadina quiere crecer y fortalecer varios de sus procesos e insertarse en una práctica de preservación documental e investigación en historia. De esta forma creemos que podremos expandir los espacios de colaboración que contribuyan a generar nuevo conocimiento y formas de divulgación. Y acá es donde creo que es importante seguir pensando la relación entre la historia y las ciencias de la información.

Conclusiones

La preservación del patrimonio documental ha sido por mucho tiempo el objeto de estudio de las ciencias de la información. Bibliotecas, archivos, museos, universidades e instituciones dedicadas a la conservación en el mundo han avanzado con proyectos y estrategias para mejorar las prácticas para preservar el patrimonio cultural en formatos digitales.

Sin embargo, la emergencia de la comunidad de práctica de las humanidades digitales expande las fronteras tradicionales entre disciplinas creando espacios comunes de práctica, reflexión teórica e investigación. En este contexto, la preservación del patrimonio cultural digital debe entenderse desde una perspectiva expansiva y crítica. En ella, los humanistas deben cumplir un rol activo en las prácticas de preservación y curación digital al igual que en la redefinición de los instrumentos y metodologías de investigación.

Las prácticas de curación digital incluyen la selección, preservación, mantenimiento, colección y archivo de recursos digitales que buscan añadir valor a los datos de investigación digital. Dichos datos constituyen la materia prima para proyectos de investigación, pedagogía y divulgación en las humanidades digitales. Por esta razón, el manejo de la información en la era digital requiere de la generación de metadatos de calidad, abiertos, que se conecten entre sí y que puedan estructurarse teniendo en cuenta su potencial de reutilizarse para análisis computacionales y visualizaciones en sistemas de información geográfico, entre otros. La gestión de la información en formatos digitales se puede pensar como una metodología de investigación en si misma que atañe a todas las disciplinas de las humanidades.