

EL GEORG ECKERT INSTITUTE Y LAS HUMANIDADES DIGITALES

The Georg Eckert Institute and the Digital Humanities

José Luis Rodríguez Gómez

Patrimonio Nacional. Real Biblioteca de Palacio

jluis.rodriguez@patrimonionacional.es

Cómo citar este artículo / Citation: Rodríguez Gómez, José Luis (2023). «El Georg Eckert Institute y las Humanidades Digitales». *Avisos. Noticias de la Real Biblioteca*, 29 (99), pp. 15-21.

Disponible en: <https://avisos.realbiblioteca.es/index.php/Avisos/article/view/815>

Resumen: El Georg Eckert Institute cuenta con una importante colección de libros de texto de todo el mundo, que constituyen una fuente de primer orden para la historia de la educación. El planteamiento multidisciplinar y la descripción de los principales proyectos desarrollados por este instituto de investigación, en el marco de las Humanidades Digitales, es el objeto de esta reseña. Se destaca la función de la biblioteca como infraestructura fundamental para la investigación académica.

Palabras clave: Georg Eckert Institute; Humanidades Digitales; Bibliotecas Digitales, Libros de texto

Abstract: The Georg Eckert Institute has an important collection of textbooks from around the world, which constitute a first-rate source for the history of education. The multidisciplinary approach and the description of the main projects developed by this research institute, within the framework of Digital Humanities, is the subject of this review. The role of the library as a fundamental infrastructure for academic research is highlighted.

Keywords: Georg Eckert Institute; Digital Humanities; Digital Libraries; Textbooks

El Georg Eckert Institute (GEI), con sede en Braunschweig (Alemania)¹, nace con su actual configuración como instituto independiente de investigación en 1975, dando continuidad y estabilidad al International Institute for the Improvement of Textbooks, que su fundador, Georg Eckert, crea en el Teacher Training College de Braunschweig. Esta institución embrionaria y la personalidad de su fundador son elementos esenciales para explicar la misión y el devenir del centro actual, perteneciente a la Asociación Leibniz².

Georg Eckert, nacido en 1912, no es ajeno al conflicto bélico que protagoniza su país y en el que participará primero como miembro del ejército de ocupación, en la Grecia ocupada y, a partir de 1944, en la resistencia. A la finalización de la guerra se implica en la construcción de un sistema de educación democrático, que combina con su función docente e investigadora en el campo de la metodología historiográfica, siempre centrado en el análisis de los libros de texto con una visión global, con el objetivo de depurar los textos de contenidos y estereotipos xenófobos para integrarlos en los principios democráticos. Fruto de este trabajo continuo, es nombrado en 1964 presidente de la German UNESCO Commission, hasta su muerte en 1974. Su trayectoria biográfica e intelectual están en la base de la labor actual del Georg Eckert Institute, en concreto, en su misión doble de investigación y revisión del corpus internacional de libros de texto.

En el núcleo de esta enorme empresa intelectual se sitúa la biblioteca, infraestructura básica para la investigación sobre el libro de texto. Desde los inicios del Georg Eckert Institute es patente la necesidad de dotar al proyecto de una colección bibliográfica, que tiene como conjunto fundacional unos 2000 libros de texto de materia histórica procedentes de la sección de educación del British Military Government. A partir de este núcleo, el crecimiento de la biblioteca es exponencial, gracias a intercambios, donaciones e importante financiación pública y privada, alcanzando el reconocimiento de colección especializada de la German Research Association. En cuanto a las materias recogidas, la historia, la geografía y los estudios sociales constituyen la base inicial, para dar cabida a partir de 2011 a la educación religiosa y a la educación en valores. Por lo que se refiere a la colección patrimonial, en los últimos años se ha hecho un esfuerzo en la adquisición de volúmenes en lengua alemana del siglo XVII. Además, entre otros cometidos institucionales, el GEI administra la base de datos de libros de texto autorizados en Alemania.

Como resultado de este proceso de adquisiciones, la biblioteca es ahora un espacio excepcional para el estudio de los libros de texto en el mundo, ya que cubre cerca de dos centenares de países, junto con una muestra más que representativa de la literatura internacional especializada sobre el libro de texto. Recorrer los estantes de la biblioteca, clasificada por países, brinda una primera aproximación visual al análisis comparativo de la diversidad editorial, de la evolución diacrónica de formatos y contenidos, y de aspectos diversos de la historia material y editorial del fondo.

¹ Leibniz Institute for Educational Media | Georg Eckert Institute, 38100 Braunschweig, Germany: <https://www.gei.de>

² Un informe sobre la historia, actualidad y propuestas de futuro del GEI, puede leerse en Fuchs y Sammler 2016.

En paralelo al desarrollo tecnológico experimentado por las bibliotecas en las dos últimas décadas, el esfuerzo por hacer accesible toda esta colección a través de internet centra los planes de este centro. El instrumento principal lo constituye el International Textbook Catalogue (ITC), que centraliza el acceso a la información bibliográfica no solo del fondo del GEI, sino que incorpora otros catálogos internacionales como EDISCO³ y MANES⁴. El primero, bajo la responsabilidad de Paolo Bianchini (coordinador científico) y Maria Cassella (coordinación biblioteconómica), es una base de datos electrónica que describe alrededor de 25.000 títulos de libros de texto publicados en Italia entre los años 1800 y 1900. Por su parte, MANES, dirigido por Gabriela Ossenbach (UNED), se centra mayoritariamente en el ámbito hispánico, ya que actúa como catálogo colectivo de un conjunto importante de bibliotecas y centros de investigación sobre historia del libro de texto de España, América Latina, Portugal y, al margen de espacio cultural, Bélgica. A diferencia de EDISCO, MANES permite la localización de los ejemplares en más de 150 bibliotecas y colecciones privadas de la mencionada área geográfica.

Progresivamente, la descripción bibliográfica se ha ido enriqueciendo con el acceso electrónico al libro. En este sentido, la biblioteca digital accesible en línea cuenta ya con un número importante de ejemplares. En su “Collection of digital historical school books with full text search”⁵ se registran más de 4000 volúmenes digitalizados y con reconocimiento óptico de caracteres, de modo que sean susceptibles de ser sometidos a métodos de análisis digital para ulteriores investigaciones. La selección de ejemplares es predominantemente germana, fruto en su mayoría del proyecto “Digitisation of German Textbooks for Different Subjects up to 1918”, que se desarrolla entre los años 2009 y 2016. Para la selección del corpus se han tenido en cuenta materiales que, por una parte, han sido objeto creciente de interés por parte de la investigación, y, por otra, son de difícil acceso. Los ejemplares digitalizados pertenecen a la colección del Georg Eckert Institute, y los recursos digitales generados se mantienen en su infraestructura digital, si bien en algunos casos se da acceso a recursos digitales externos.

El núcleo cuantitativamente más importante es el de los libros de texto de historia correspondientes al periodo imperial alemán (1871-1918) con más de 1000 ítems, seguido por los libros de lectura (*Readers*) del mismo periodo con 800 ejemplares. La representación de otros países en esta biblioteca digital es todavía escasa; tan solo unas decenas de textos correspondientes a Gran Bretaña, Suiza, Francia, México, Canadá o Estados Unidos.

Desde el punto de vista técnico y de asunción de estándares, los metadatos están disponibles para su exportación en los formatos más habituales: Marc XML, Dublin Core, METS, o RIS, para los gestores de citas bibliográficas. Además, el sistema cumple con los estándares promovidos por el International Image Interoperability Framework (IIIF⁶), y ofrece la URL del *manifest*, es decir, un archivo en formato JSON que sigue la norma IIIF Presentations, por medio del cual el recurso electrónico puede ser consumido por aplicaciones, mayoritariamente de código abierto, desarrolladas en el contexto IIIF, tales

³ EDISCO: <https://www.edisco.unito.it/>

⁴ Centro de Investigación en Manuales Escolares (MANES): <https://www.centroman.es/>

⁵ Collection of digital historical school books with full text search: <https://gei-digital.gei.de/viewer/index/>

⁶ International Image Interoperability Framework (IIIF): <https://iiif.io/>

como visualizadores (IIIF Presentation) o sistemas de anotación (IIIF Annotation). Finalmente, está disponible también el contenido del libro en formato texto plano obtenido y depurado a partir de las tecnologías del OCR.

Esta biblioteca digital (GEI.digital), que forma parte del módulo Education Media Research (EDUMERES), comparte espacio y se complementa con una serie de recursos de especial interés para la investigación. Este portal, integrado en el sitio web del GEI, ofrece conjuntos de datos estructurados, entre los que se cuenta la plataforma Curricula Workstation, que da acceso abierto a una amplia colección alemana e internacional de curricula, es decir, planes de estudio oficiales que determinan el contenido de los libros de texto. El buscador, que permite la localización inmediata y el acceso a texto completo en PDF a los documentos de este corpus perteneciente a la biblioteca de investigación del Leibniz Institute for Educational Media, facilita su función de recuperación de la información con la aplicación de filtros por países, materias, nivel educativo y fechas, además de extender la búsqueda por palabras al texto completo.

Integrado también en EDUMERES, el repositorio digital edu.docs, gestionado por el sistema de código abierto Dspace⁷, da acceso a texto completo a las publicaciones académicas sobre libros de texto, publicadas por instituciones externas y por el propio Leibniz Institute, tales como los *Eckert Dossiers*, la *Eckert Beiträge* o el *Journal of Educational Media, Memory and Society* (JEMMS). Por otra parte, EDUMERES aloja la base de datos de libros de texto aprobados en las escuelas alemanas desde el curso 2009-2010 de las materias de ciencias sociales (GEI-DZS), en concreto, geografía, historia y estudios sociales y políticos. En el mismo módulo se presentan edu.news, que recoge las noticias de congresos, eventos y cualquier otro acontecimiento relacionado con la investigación en el campo de la educación, y edu.reviews, que da cuenta de 380 reseñas de libros de texto alemanes publicadas entre 2008 y 2018, fecha en que se concluye esta iniciativa.

Para el ámbito español e iberoamericano es de especial importancia el proyecto GLOTREC (Global Textbook Research Center), que tiene como objetivo ampliar globalmente la cobertura geográfica con la integración en el International Textbook Catalogue (ITC) y en GEI.digital de libros de texto de todo el mundo. Hasta el momento incorpora ya la ya mencionada base de datos MANES, que, a su vez, actúa como catálogo colectivo con alrededor de 37.000 registros pertenecientes a más de 150 bibliotecas públicas y privadas y centros de investigación, entre los que se cuentan algunos de los participantes activos en GLOTREC, tales como el Centro de Investigación y de Estudios Avanzados (CINVESTAV), de México, el Grupo de Investigación Histórica en Educación e Identidad Nacional (GIHEIN), de la Universidad del Atlántico (Barranquilla, Colombia), o el Instituto de Historia de Cuba (IHC). Pertenecientes a este mismo ámbito geográfico, pero con catálogos propios integrados directamente en GLOTREC, cabe mencionar el Museo de la Educación Gabriela Mistral (Chile), la Biblioteca Nacional de Maestros (Argentina) y la Universidad Estadual de Ponta Grossa (Brasil), colaboradores imprescindibles en esta propuesta global de acceso e investigación sobre los libros de texto.

⁷ Dspace. Sistema de gestión de repositorios digitales: <https://dspace.lyrasis.org/>

Sin embargo, el proyecto de digitalización de GEI no se detiene simplemente en la difusión *online* de las fuentes de estudio y materiales de referencia, que de por sí constituye un apoyo fundamental para la investigación y la base imprescindible para futuras actuaciones, sino que se propone el desarrollo en paralelo de herramientas y métodos digitales para la exploración de los textos y el apoyo a propuestas de proyectos concretos de investigación, protagonizando siempre la biblioteca el papel de agente nuclear en todo este proceso colaborativo e interdisciplinar.

Así, en este contexto, vigente durante el bienio 2021-2022, el proyecto “Data curation for research into historical educational media”, incluye el desarrollo de los *softwares* OCR4all y LAREX, diseñados para la mejora del proceso de reconocimiento de caracteres y la estructuración de los textos sometidos a digitalización, con interfaces de usuarios apropiadas para investigadores particulares e instituciones. La colaboración entre los técnicos e investigadores de la institución y estudios externos del área de historia del libro de texto, que constituyen el verdadero *target* del proyecto, orienta la definición de un flujo de trabajo eficiente y la consecuente obtención de resultados fiables y útiles. En efecto, solo un control de calidad riguroso de los textos seleccionados y sometidos a reconocimiento óptico de caracteres puede garantizar resultados de investigación satisfactorios. Hertling (2022) ofrece detalles de todo este proceso que pueden servir de modelo para proyectos semejantes.

En la base de todo este plan integrado de digitalización están las consideraciones primordiales y programáticas de Ernesto De Luca, director del departamento Human-Centered Technologies for Educational Media, del GEI, en la línea de avanzar desde las interfaces de simple consulta y acceso hacia la generación de herramientas de análisis más complejas, imprescindibles en la práctica efectiva de la investigación historiográfica actual, en el marco de la digitalización, la web semántica o el procesamiento de lenguaje natural:

Current Digital Libraries technologies don't allow its researchers to fully exploit the potential offered by Natural Language Processing technologies, thus a new platform, based on open source solution, has been realized with the aim to offer a set of tools to the researchers, allowing them to adapt the research process to the single case without predefining a workflow (De Luca et al. 2019, p. 386).

La aproximación metodológica acorde con este planteamiento debe aportar la suficiente flexibilidad para adaptarse al modo de proceder de la investigación. Así, se considera en un primer bloque el proceso de digitalización, entendido de forma integral, que tiene como resultado no solo el acceso digital a las imágenes, sino también la incorporación de los metadatos de acuerdo con los estándares internacionales, y la aplicación de tecnologías de la web semántica en el marco de los Linked Open Data (LOD), tales como, por ejemplo, las ontologías, que permiten definir con precisión relaciones entre cualesquier entidades identificadas en la red. A partir de los textos así digitalizados se procede a la creación de corpus estructurados de textos seleccionados en virtud de características bibliográficas, de localización o de cualquier otro tipo, siempre atendiendo a hipótesis y necesidades de la investigación. Estos corpus, mediante herramientas y métodos digitales apropiados, pueden ser manipulados en diferentes sentidos. Por ejemplo, a través del uso de lenguajes de marcación como TEI (Text Encoding

Initiative)⁸, se pueden obtener ediciones digitales enriquecidas, poniendo de relieve la estructura del texto, identificando y normalizando entidades de distinta naturaleza, añadiendo anotaciones de tipo interpretativo, etc. Finalmente, las técnicas del Procesamiento del Lenguaje Natural (PLN) permiten el análisis léxico, morfológico y sintáctico del corpus, no solo aportando una solución a necesidades concretas de una investigación en particular, sino abriendo nuevas vías de estudio. La arquitectura informática que sostiene esta caja de herramientas (*toolbox*) de investigación se dibuja con precisión en De Luca (2019, p. 394-396).

Esta misma línea de adhesión al paradigma de las Humanidades Digitales (HD) de las prácticas profesionales del proyecto del GEI y de su metodología y concepción es sostenida por Maret Nieländer (2023), en su intervención en la 2ª Conferencia del Proyecto GLOTREC (Global Textbook Research Center), celebrada en Braunschweig del 20 al 22 de marzo, coordinada por Kerstin Schwedes, directora de GLOTREC. Su contribución hizo un recorrido, a partir de ejemplos específicos extraídos de los proyectos en curso del GEI, por la aplicación de las tecnologías de las HD, tales como la anotación de los textos, los análisis de sentimientos o argumentos, la estilometría, etc., tanto en su aplicación a textos concretos como a grandes corpus, moviéndose entre la *close reading*, es decir el análisis pormenorizado de un texto particular, y la *distant reading*, entendida como la aplicación de métodos computacionales a corpus extensos⁹. En este mismo contexto y por lo que se refiere a herramientas de análisis y marcado, CATMA¹⁰, desarrollada en la Universidad de Hamburgo, se presentó como una de las más eficientes en el ámbito de las humanidades, por sus amplias funcionalidades y por su facilidad de uso.

Este congreso sirvió, además, para dar a conocer los resultados de investigaciones propuestas a partir de un corpus concreto de libros de texto sometido previamente al proceso de digitalización y reconocimiento óptico de caracteres, acción esta que contó con la financiación del GEI. La selección de los materiales se realizó en función de dos temas de estudio, “La difusión de la enseñanza intuitiva, 1880-1930” y “La imagen de los otros”, considerados ambos en los libros de texto publicados en Europa y Latino-América. Con este planteamiento, se ponía en práctica la interacción entre la labor propiamente historiográfica, llevada a cabo por investigadores de historia de la educación de España, Italia y países latinoamericanos, como Cuba, Argentina, México, Colombia, Chile o Brasil, y el equipo tecnológico del GEI, en un flujo de trabajo verdaderamente interdisciplinar.

En conclusión, el equipo del GEI está comprometido, en el contexto de las Humanidades Digitales, en el desarrollo de un entorno tecnológico que se propone como parte de la solución de los problemas planteados por la historiografía y las ciencias sociales. El constante diálogo interdisciplinar entre especialistas en Historia de la Educación, desarrolladores de *software* o expertos en usabilidad¹¹ garantiza la eficacia de las

⁸ Text Encoding Initiative (TEI): <https://www.tei-c.org>

⁹ En Moretti, F. (2000) se introducen estos conceptos, que son ya de uso habitual en el ámbito de las HD.

¹⁰ Computer Aided Textual Markup & Analysis (CATMA): <https://catma.de>

¹¹ En la 2ª Conferencia del Proyecto GLOTREC, el equipo de usabilidad presentó el laboratorio y se realizaron prácticas de monitorización con los asistentes.

propuestas tecnológicas. Por su parte, la biblioteca de investigación está en el centro de esta empresa innovadora, aportando las fuentes de estudio y contribuyendo al desarrollo y aplicación de las herramientas tecnológicas, abriendo nuevos caminos de análisis e interpretación.

Referencias

- De Luca, E.W., Fallucchi, F., Ligi, A. y Tarquini, M. (2019). «A Research Toolbox: A Complete Suite for Analysis in Digital Humanities». En: E. Garoufallou, F. Fallucchi y E. William de Luca (eds.), *Metadata and Semantic Research. MTSR 2019. Communications in Computer and Information Science, vol. 1057*. Cham: Springer International Publishing, pp. 385-397.
- Fuchs, E. y Sammler, S. (2016). *Textbooks between Tradition and Innovation. A Journey through the History of the Georg Eckert Institute* [en línea]. Braunschweig: Georg Eckert Institute
- Hertling, A. y Klaes, S. (2022). «Volltexte für die Forschung: OCR partizipativ, iterativ und on Demand». *O-Bib. Das Offene Bibliotheksjournal / Herausgeber VDB*, vol. 9, n. 3, pp. 1-11. DOI: <https://doi.org/10.5282/o-bib/5832>.
- Moretti, F. (2000). «Conjectures on World Literature». *New Left Review*, vol. 1.
- Nieländer, M. (2023). Personal Experiences in using DH Tools in Cultural Studies Projects. . S.I.: 2ª Conferencia GLOTREC (Global Textbook Research Center), Braunschweig, 20 al 22 de marzo de 2023.