

¿Quién conforma la comunidad de la TEI en español? Análisis de los datos de una encuesta.

María Gimena del Rio Riande y Allés-Torrent, Susanna.

Cita:

María Gimena del Rio Riande y Allés-Torrent, Susanna (2023). *¿Quién conforma la comunidad de la TEI en español? Análisis de los datos de una encuesta.* *Journal of the Text Encoding Initiative*, 16.

Dirección estable: <https://www.aacademica.org/gimena.delrio.riande/235>

ARK: <https://n2t.net/ark:/13683/pdea/QyG>



Esta obra está bajo una licencia de Creative Commons.
Para ver una copia de esta licencia, visite
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es>.

Acta Académica es un proyecto académico sin fines de lucro enmarcado en la iniciativa de acceso abierto. Acta Académica fue creado para facilitar a investigadores de todo el mundo el compartir su producción académica. Para crear un perfil gratuitamente o acceder a otros trabajos visite: <https://www.aacademica.org>.



¿Quién conforma la comunidad de la TEI en español? Análisis de los datos de una encuesta

Who Defines the Community of the TEI in Spanish? Data Analysis of a Survey

Gimena del Rio Riande y Susanna Allés-Torrent



Edición electrónica

URL: <https://journals.openedition.org/jtei/4927>

DOI: 10.4000/jtei.4927

ISSN: 2162-5603

Editor

TEI Consortium

Referencia electrónica

Gimena del Rio Riande and Susanna Allés-Torrent, "¿Quién conforma la comunidad de la TEI en español? Análisis de los datos de una encuesta", *Journal of the Text Encoding Initiative* [Online], 16 | 2023, Online since 01 August 2023, connection on 18 August 2023. URL: <http://journals.openedition.org/jtei/4927> ; DOI: <https://doi.org/10.4000/jtei.4927>

For this publication a Creative Commons Attribution 4.0 International license has been granted by the author(s) who retain full copyright.

¿Quién conforma la comunidad de la TEI en español? Análisis de los datos de una encuesta

Who Defines the Community of the TEI in Spanish? Data Analysis of a Survey

Gimena del Rio Riande y Susanna Allés-Torrent

RESUMEN

En este artículo analizamos los resultados obtenidos en una encuesta sobre el uso de la Text Encoding Initiative para proyectos de Humanidades Digitales en español que se llevó a cabo a principios de 2022 en el seno del proyecto “Communicating the Text Encoding Initiative to a Multilingual User Community” financiado por la A. Mellon Foundation. Los objetivos principales del análisis de esta encuesta son entender quien forma parte de la comunidad de usuarios de TEI en el mundo hispanohablante o en iniciativas relacionadas con textos en español, qué tipo de

formación académica relacionada con la TEI se ofrece hoy en día, cómo y con qué finalidades se usa este tipo de codificación de textos y cuáles son los recursos existentes para su enseñanza y aprendizaje.

ABSTRACT

In this article we analyze the results obtained in a survey on the use of the Text Encoding Initiative for Digital Humanities projects in Spanish that was conducted in early 2022 within the project “Communicating the Text Encoding Initiative to a Multilingual User Community” funded by the A. Mellon Foundation. The main objectives of our analysis are to understand who conforms the TEI user community in the Spanish-speaking world or in Spanish text-related initiatives, what types of TEI-related academic training is offered today, how and for what purposes this type of text encoding is used, and what are the existing resources for teaching and learning.

ÍNDICE

Palabras Clave: comunidad TEI, investigación, enseñanza, codificación, edición digital

Keywords: TEI community, research, teaching, encoding, scholarly digital editing

1. Extendiendo las fronteras lingüísticas y geográficas de la TEI: el proyecto Communicating the Text Encoding Initiative to a Multilingual User Community

- ¹ Las guías directrices o Guidelines de la TEI están escritas y editadas íntegramente en inglés, con traducciones incompletas a varios idiomas. Los esfuerzos por internacionalizar la documentación del TEI se remontan al menos a 2005 con una iniciativa liderada por el siempre visionario Sebastian Rahtz, quien diseñó una primera infraestructura colaborativa para estas traducciones (Rahtz 2005, 2006).¹ El resultado de esta iniciativa fue una serie de traducciones parciales al japonés, chino, coreano, italiano, español, francés y alemán. A lo largo de los años transcurridos desde aquel

esfuerzo inicial, se han realizado actualizaciones periódicas a idiomas concretos, las más recientes alemán, japonés y español (Stadler y Scholger 2016, Allés-Torrent, del Río Riande y Calarco 2021-, Okada, Nakamura y Nagasaki s.f.).

- 2 El proyecto “Communicating the Text Encoding Initiative to a Multilingual User Community” fue financiado por la Fundación Andrew W. Mellon a través del programa Public Knowledge.² Se desarrolló a lo largo de un año a partir de marzo de 2021, y posteriormente se prorrogó hasta diciembre de 2022. Los principales objetivos de este proyecto son tres y están relacionados con el apoyo al grupo de internacionalización de la TEI, que es parte de las iniciativas del Consorcio:³ en primer lugar, el desarrollo de una nueva infraestructura, TranslateTEI,⁴ para mejorar la experiencia de usuario al colaborar con traducciones multilingües de las especificaciones de la TEI;⁵ en segundo lugar, una encuesta sobre el uso de la TEI por parte de la comunidad hispanohablante global y/o para textos académicos en español, y, en tercer lugar, el desarrollo de materiales de formación sobre la TEI –tutoriales, ejercicios, ejemplos y recursos académicos– en español.
- 3 Este trabajo recopila y analiza los datos de una encuesta bilingüe (español-inglés) que llevamos a cabo, como parte de este proyecto, durante la primera mitad de 2022, y en la que diferentes usuarios de la TEI para textos en español ofrecieron sus valiosas perspectivas sobre su adopción y uso, y sobre sus necesidades como parte de esta comunidad.

2. Nuestra encuesta: Acceso, uso y necesidades de quienes usan la TEI para textos en español

- 4 La encuesta bilingüe “Use of the Text Encoding Initiative in Spanish for Digital Humanities Projects” / “Uso de la Text Encoding Initiative para proyectos de Humanidades Digitales en español” fue desarrollada como parte de la participación del grupo “TTHub o Hub de Tecnologías del Texto”⁶ y la colaboración de HD Lab⁷ en este proyecto en los últimos meses de 2021 y se distribuyó desde el 24 de enero de 2022 a través de diferentes listas académicas de correo electrónico y redes sociales,⁸ y permaneció abierta hasta el 10 de marzo de 2022. La encuesta constó de 22 preguntas que podían responderse en aproximadamente 10 minutos. Privilegiamos el anonimato en las respuestas, aunque recogimos datos relacionados con la afiliación y la nacionalidad de los participantes, con el fin de obtener una mejor comprensión de nuestra

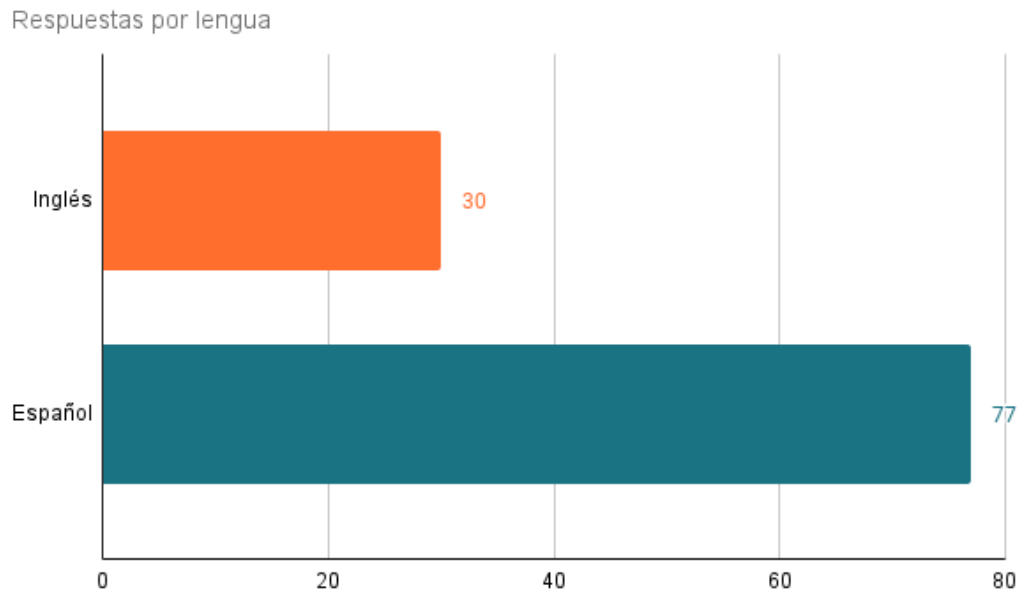
comunidad desde una perspectiva geográfica y geopolítica. El grupo al que apuntamos estaba compuesto por profesores, investigadores, estudiantes o cualquier persona que hubiera utilizado la TEI para la investigación, la docencia en español o para textos en español, o en cualquier situación de aprendizaje en esta lengua en cualquier país. Pensar que la investigación sobre textos en español solo se realiza desde países hispanohablantes es no comprender que la investigación va más allá de las fronteras geográficas y lingüísticas de un país. El campo de las humanidades digitales viene demostrando además cómo la investigación radicada en un país casi siempre cuenta con colaboraciones de otras instituciones (Isasi Velasco y del Río Riande 2022). Esta amplitud en la definición de las características del grupo nos permitió trazar conclusiones alrededor de la pregunta “¿quién conforma la comunidad de la TEI en español?”. Desde nuestro punto de vista, y desde los objetivos de la encuesta, esta no solo está compuesta por usuarios de la TEI en países hispanohablantes o que usan el español para la comunicación científica sino también por aquellos insertos en proyectos de investigación que trabajan con la codificación o edición de textos en español con TEI desde cualquier país o institución.

- 5 El software utilizado para crear y distribuir la encuesta fue Qualtrics,⁹ y la licencia de uso fue proporcionada por la Universidad de Miami. 135 participantes iniciaron la encuesta, aunque sólo 107 respondieron a todas las secciones y preguntas. 77 de estas 107 respuestas se escribieron en español y 30 en inglés (figura 1).
- 6 La encuesta se dividió en cuatro bloques principales: Información general sobre los participantes, Formación académica en TEI, Uso de la TEI, y Enseñanza y aprendizaje de la TEI. A continuación, ofrecemos un resumen de las respuestas que consideramos más relevantes, junto con unas breves conclusiones y una propuesta de trabajo a futuro.

2.1 Información general sobre los participantes

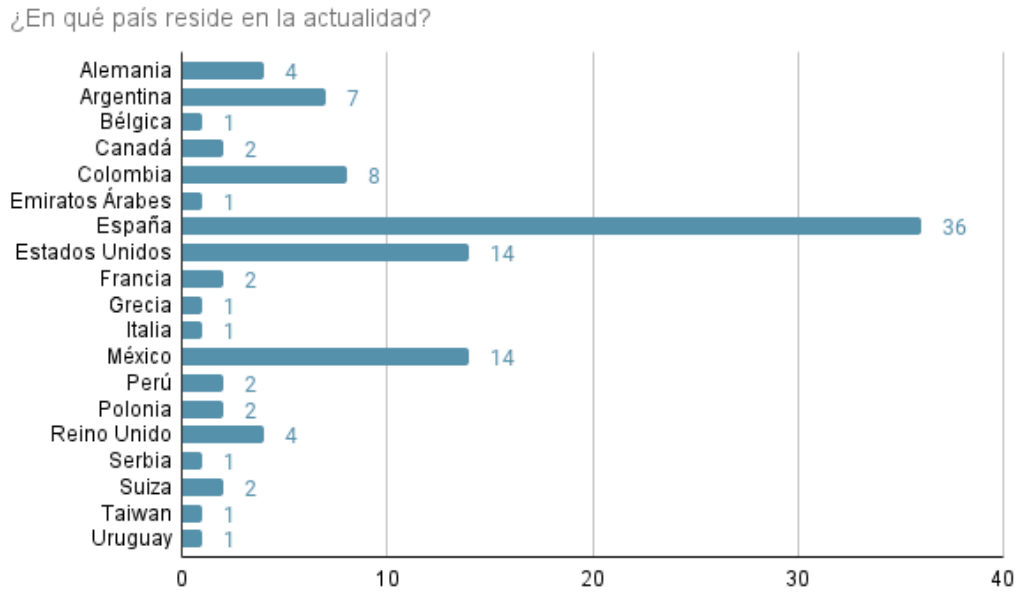
- 7 Esta sección nos ayudó a averiguar quién es la población que trabaja actualmente en proyectos TEI con textos en español y dónde. Además, nos proporcionó información sobre intercambios globales y perfiles académicos. Como ya hemos mencionado, ofrecimos la encuesta en español e inglés, sin embargo, como era de esperar, la mayoría de las respuestas se completaron en español:

Figura 1. Lengua en la que se respondió la encuesta.



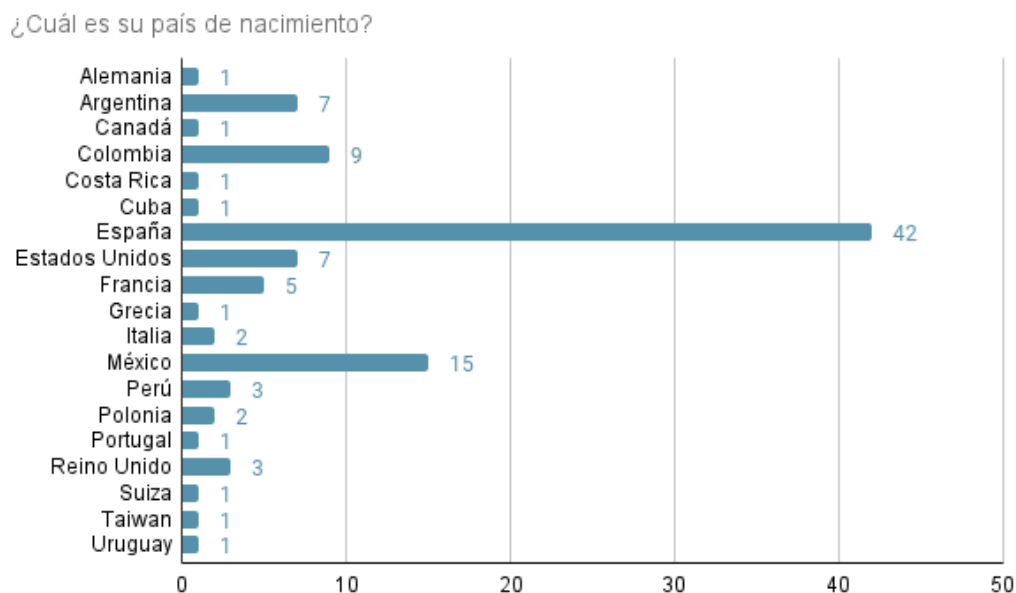
- 8 Según las respuestas, la mayoría de las personas que respondieron a la encuesta viven actualmente en España. Se trata de un primer gran grupo de participantes en un país que fue pionero en el uso de la TEI (Allés-Torrent y del Río Riande 2020), pero también podemos distinguir otros dos grupos: uno formado por participantes que viven en México y EE.UU. (14 participantes en cada caso) y un tercer grupo más pequeño constituido por quienes están en Argentina y Colombia.

Figura 2. País de residencia de los encuestados.



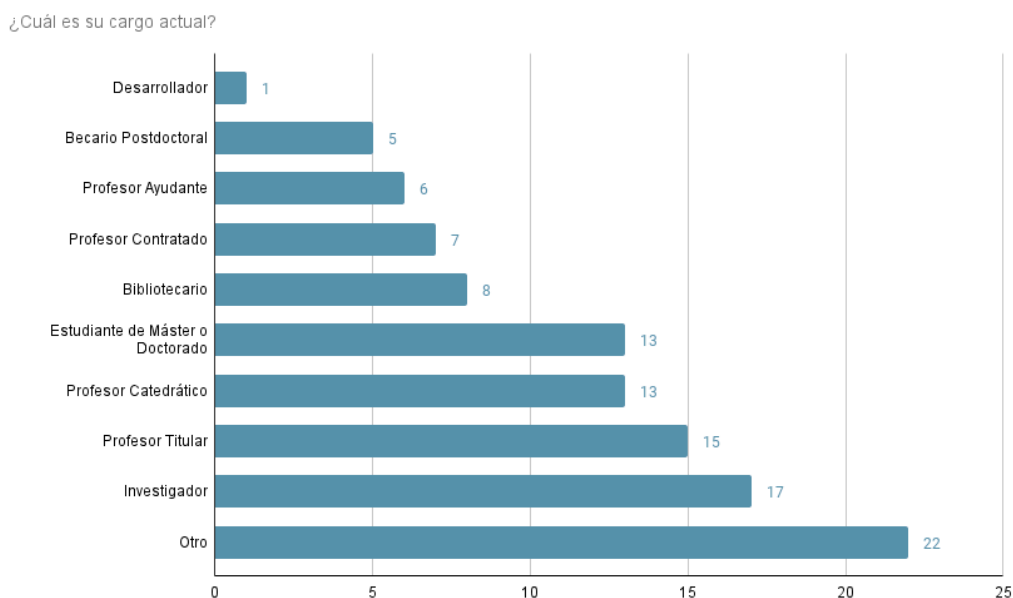
- 9 La movilidad que hace a la comunidad global interesada en el uso de la TEI en español que antes mencionamos puede verse en los datos de nuestra encuesta: muchos participantes nacieron en un país distinto de aquel en el que viven actualmente. La mayoría de estos participantes son españoles (42 frente a 36):

Figura 3. País de origen de los participantes.



- 10 En esta sección, la pregunta sobre los cargos académicos de los participantes nos llevó a algunas cifras interesantes con respecto a quiénes hoy usan la TEI. Investigadores, Profesores Asociados y Profesores Titulares parecen ser los cargos académicos de la mayoría, pero no son muchos más que los Lectores o Profesores Adjuntos. Por otro lado, Magister o Graduado son cargos más populares entre el grupo de estudiantes que los postdoctorales. También encontramos bibliotecarios y sólo un desarrollador. Ha de destacarse que, en la categoría *Otros*, los participantes respondieron con categorías que podrían haber coincidido con alguna de las propuestas: Estudiante de doctorado (2), postdoctorado (1), Técnico de apoyo a la investigación (4), Investigador (2), Archivista (2), Especialista en sistemas de metadatos (1), Estudiante de grado (2), Editor (2), otras categorías de Profesorado (3), Gestor de proyectos (1).

Figura 4. Cargos académicos de los participantes.



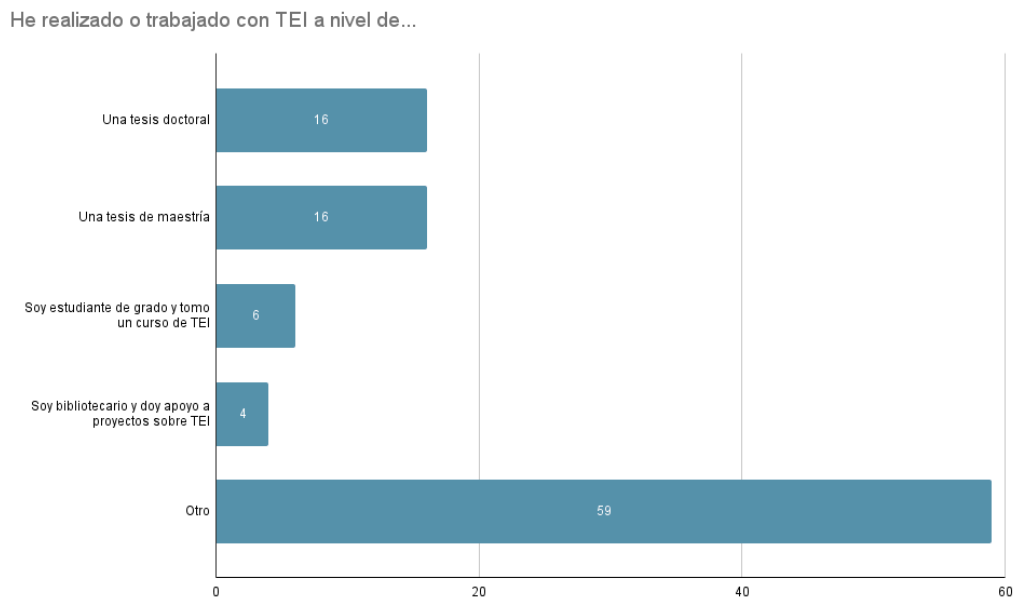
- 11 En síntesis, aunque el carácter académico de quienes usan o están interesados en el uso de la TEI para textos en español persiste y está representado principalmente por cargos académicos altos, puede verse una extensión hacia otros niveles en la carrera universitaria. Sería deseable una mayor permeabilidad en las instituciones culturales (bibliotecas, archivos) y en los equipos técnicos.

2.2 La formación académica en TEI

- 12 El conjunto de preguntas propuestas en la segunda sección, Formación académica en TEI, nos ayudó a comprender los antecedentes de los participantes. Las respuestas nos mostraron que algunos participantes habían estudiado o utilizado la TEI para su máster (16 respuestas) o tesis doctoral (16). El porcentaje menor fue el relacionado con el uso de la TEI en estudios de grado (6) o como parte del apoyo bibliotecario (4). Estos resultados no son sorprendentes ya que, por un lado, la mayoría de los programas que incluyen formación en TEI son parte de cursos de postgrado y, por el otro, roles como el de bibliotecario en humanidades digitales son casi inexistentes en España y Latinoamérica. La categoría *Otros* recibió la puntuación más alta con 59 respuestas; algunos participantes simplemente repitieron su trabajo: aficionado (1), archivista (1), revistas científicas (1), gestor de proyectos (1), investigador (4), instructor (3). Otros especificaron su nivel actual de

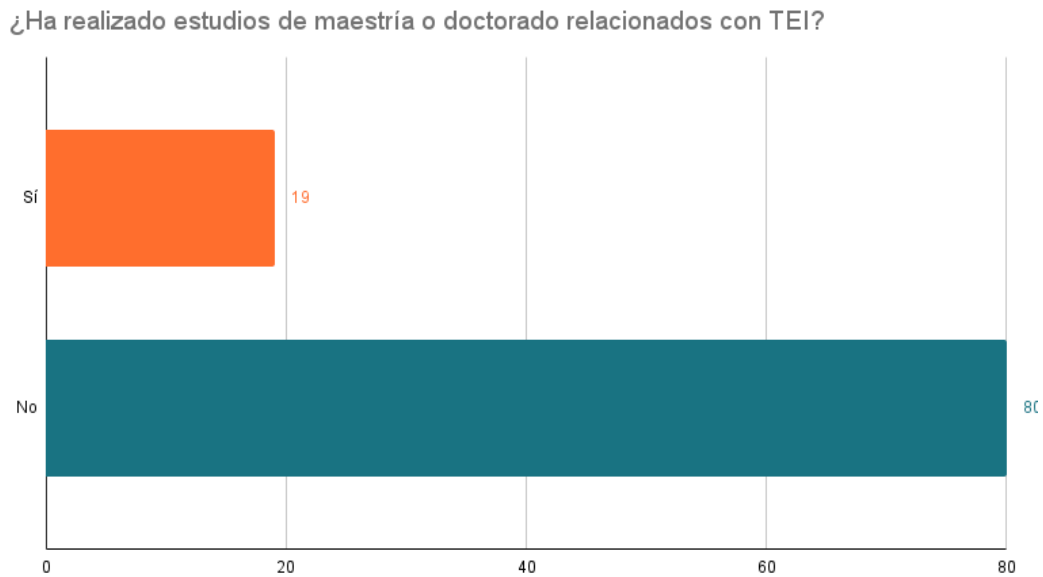
estudios: Licenciatura (4), Máster (3), Posgrado (1). Unos pocos participantes indicaron también que habían realizado algunos cursos a nivel universitario, incluidas las Escuelas de Verano (8). Pero el dato más llamativo es el número de participantes que eligieron *Proyecto de investigación* (21 respuestas), es decir, personas que utilizan TEI porque forman parte de un proyecto en curso.

Figura 5. Formación en TEI.



- 13 También fue muy interesante descubrir que la mayoría de los participantes no habían realizado ningún curso sobre TEI como parte de su máster o de sus cursos de posgrado (19 sí versus 80 no). Esto significa que, aunque la oferta en cursos de posgrado es cada vez mayor, sólo unos pocos imparten cursos sobre TEI.

Figura 6. Estudios de maestría o doctorado relacionados con la TEI.



- 14 Por último, destaca el porcentaje de participantes que habían realizado sus estudios de posgrado en España y Francia, seguido por Reino Unido. Es difícil saber si se trata de una coincidencia de los perfiles de los participantes o en realidad evidencia una mayor oferta formativa a nivel de máster y doctorado en estas áreas. Ciertamente en estos tres países hace ya casi una década que se ofrecen formaciones en humanidades digitales que incluyen codificación de textos en TEI. A modo de ejemplo, traemos aquí la formación permanente de la Universidad Nacional de Educación a Distancia a través de su Laboratorio de Innovación en Humanidades Digitales (LINHD) de la UNED (Madrid) en España, la École Nationale des Chartes o el Centre d'Études Supérieures de la Renaissance en Francia, o el Máster en Humanidades digitales de King's College.¹⁰

Figura 7. País en el que realizó su posgrado.

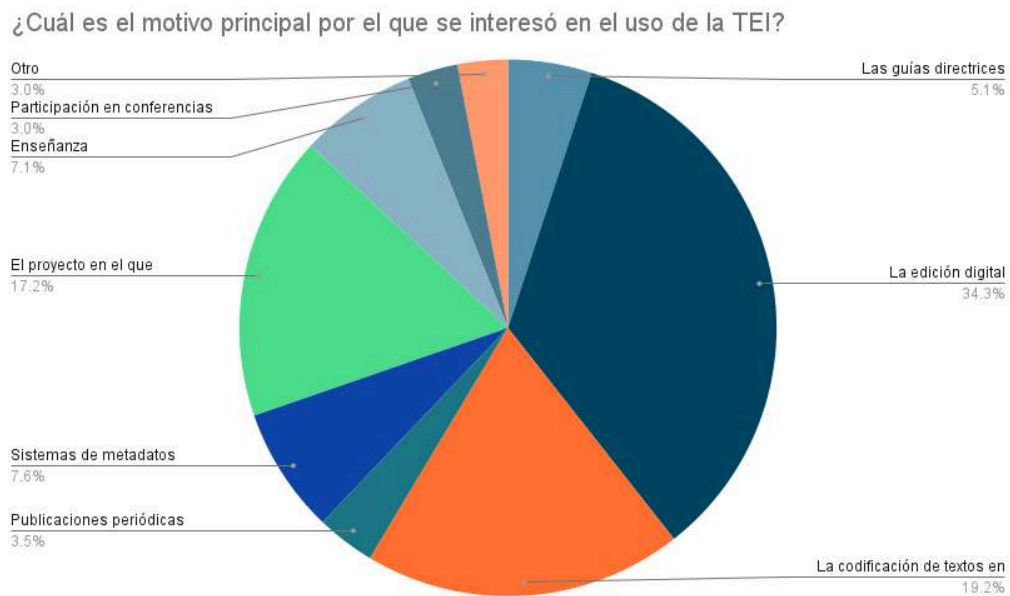


2.3 Utilización de la TEI en investigaciones actuales

- 15 En la tercera sección de la encuesta, pretendíamos conocer mejor los usos actuales de la TEI por parte de los participantes. En primer lugar, les preguntamos por la razón principal de su interés en el uso de la TEI, dándoles las siguientes opciones que debían ordenar de más a menos relevante: 1. Guías directrices de la TEI, 2. Ediciones digitales, 3. La codificación de textos en general, 4. Artículos u otras publicaciones académicas, 5. Sistemas de metadatos, 6. Participación en un proyecto en el que se adoptaba el marcado TEI, 7. Docencia, 8. Participación en conferencias, 9. Otros. Como se indica en la [figura 8](#), el 34,3% de los participantes (68 respuestas) eligió como principal motivo de su interés por la TEI la creación de ediciones digitales, seguido del interés por el mercado general de textos (19,2%, 38) y la participación en proyectos de investigación en los que se les exige trabajar con la codificación de textos (17,2%, 34). En un nivel inferior, encontramos el interés por los sistemas de metadatos (15), la docencia (14), las guías directrices de la TEI (10) o los artículos y otras comunicaciones académicas (7). Sorprendentemente, la participación en conferencias se

valora más bajo de lo que cabría esperar (6), ya que estos lugares son los más adecuados para conocer proyectos, métodos o herramientas en curso. El 3% restante eligió *Otros* y añadió términos generales como investigación, o humanidades digitales en general.

Figura 8. Uso de la TEI en investigaciones actuales de los participantes.



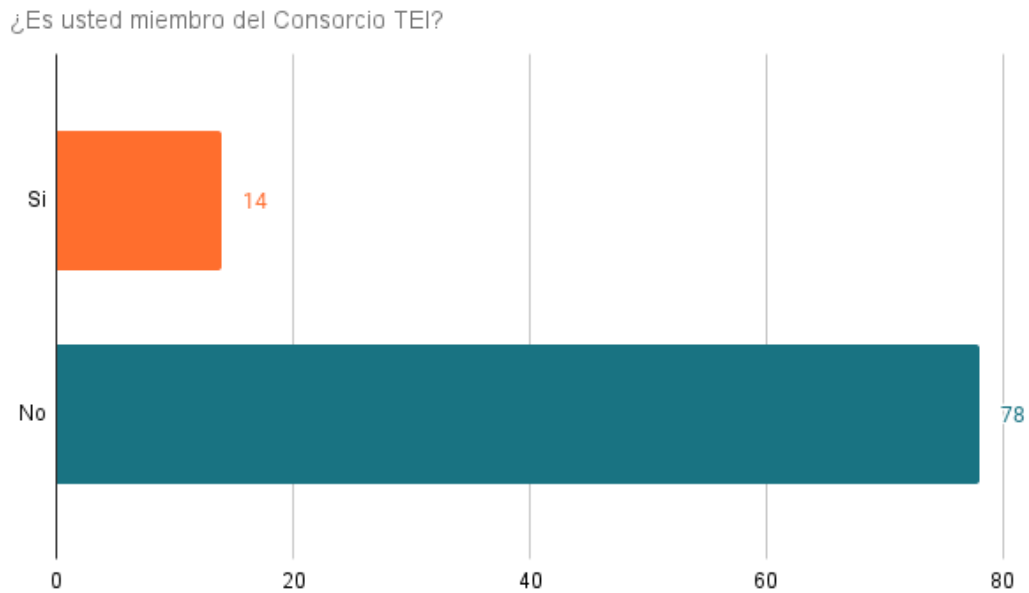
- 16 La siguiente pregunta se refería a las funciones de los participantes en los proyectos de investigación, y se dieron las siguientes opciones: 1. Director / Investigador principal, 2. Miembro del equipo de investigación, 3. Miembro del equipo técnico, 4. Consultor, 5. Colaborador externo, 6. No participo en ningún proyecto. Obtuvimos los siguientes números: la mayor parte de los participantes (30,6%, 38) confirmaron ser miembros de un equipo de investigación. Los directores o investigadores principales (24) y los miembros de equipos técnicos (17) figuraban entre las respuestas más frecuentes. Sólo 13 participantes eran colaboradores en un proyecto que utilizaba TEI, y 9 ofrecían servicios de consultoría. Resulta curioso que un número considerable de participantes (18,5%, 23) admitieran que no participaban en ningún proyecto, lo que debe tenerse en cuenta como una característica específica de esta comunidad y merece una mayor exploración sobre las causas que dificultan la participación.

Figura 9. Rol de los participantes en un proyecto de investigación.

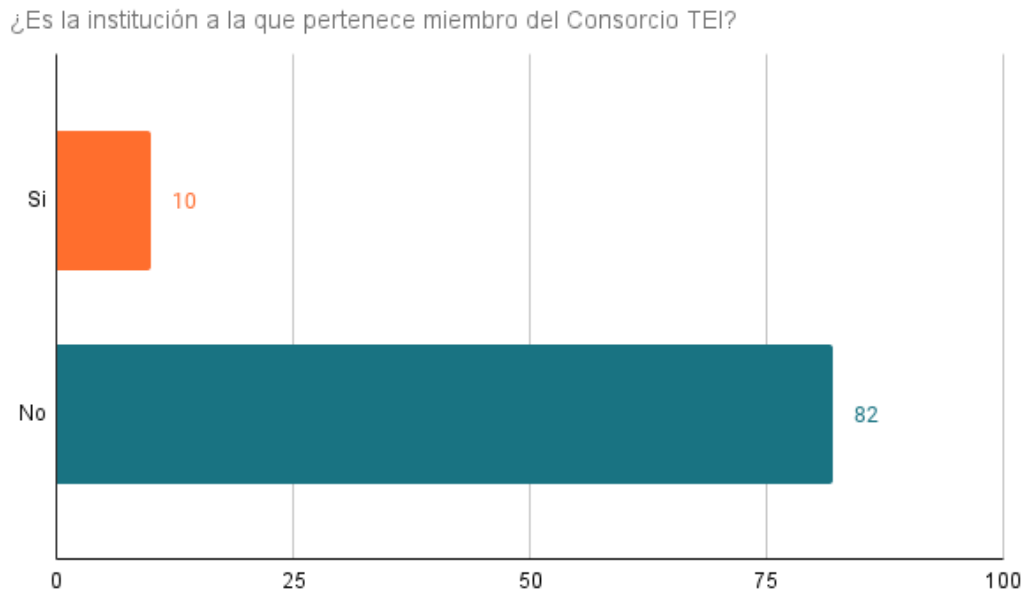


- 17 En cuanto a la participación en la comunidad TEI internacional, el número de miembros del Consorcio TEI sigue siendo extremadamente bajo: sólo el 15% de los participantes son miembros individuales. La gran mayoría, el 85%, no son miembros de la TEI. Estas cifras son muy similares a las de la encuesta de 2017 llevada a cabo por Gimena del Rio Riande en el marco de su plenaria en el congreso internacional de la TEI en la University of Victoria, Canadá (del Rio Riande 2017). Cabe destacar además que el 13% de los participantes ni siquiera contestó.

Figura 10. Membresía individual del consorcio de la Text Encoding Initiative.

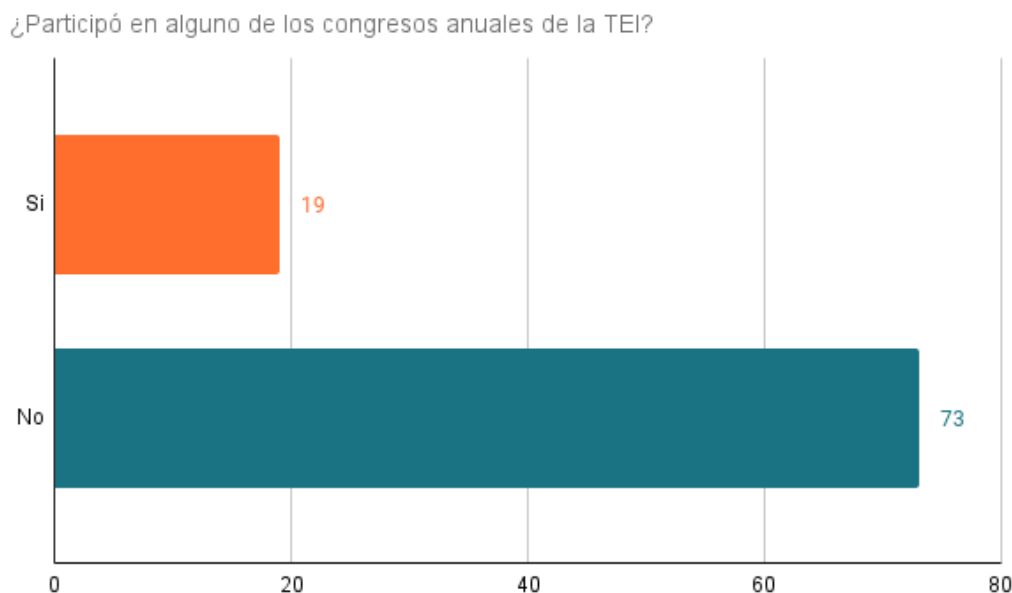


- 18 Finalmente, otro dato de interés es que de 82 respuestas sólo 9 mencionaron una afiliación institucional al consorcio de la TEI. 13 participantes no contestaron. Las 9 respuestas relacionadas con la afiliación institucional corresponden a proyectos fuera de América Latina.

Figura 11. Membresía de la institución al consorcio de la TEI.

- 19 La participación en los congresos de la TEI parece tener más alcance que la afiliación, aunque las cifras son muy bajas. El 21% ha participado al menos una vez en un congreso de la TEI, mientras que el 70% nunca ha presentado una ponencia en estos eventos. El 13% no contestó. Estas cifras están estrechamente relacionadas con las respuestas sobre los intereses en la TEI mencionados al describir la segunda sección de la encuesta. En general, podríamos pensar que la comunidad hispanohablante permanece al margen de la comunidad TEI global, especialmente en términos de participación como miembros o en los eventos oficiales del consorcio. Esta situación se debe probablemente no sólo a la brecha lingüística, sino también a la falta de financiación para la investigación en humanidades digitales en los países hispanohablantes y a problemas económicos de muchos países en América Latina.

Figura 12. Participación en congresos de la TEI.



- 20 A continuación, preguntamos a los participantes por las tecnologías utilizadas en sus proyectos, proponiéndoles respuestas múltiples que incluían: 1. Personalizaciones de los módulos TEI, 2. Utilización de esquemas (ODD, RelaxNG, etc.), 3. Bases de datos XML, 4. Bases de datos relacionales (MySQL, PostgreSQL, etc.), 5. Transformaciones XSLT, 6. Transformaciones XQuery, 7. Vocabularios controlados o tesauros, 8. Tecnologías sobre Sistemas de Información Geográfica (SIG), 9. Tecnologías de la información y la comunicación, 9. Tecnologías sobre Procesamiento del Lenguaje Natural (PLN), 10. Tecnologías sobre web semántica, 11. Visualización de datos, 12. Otros. El objetivo detrás de algunas de las opciones que dimos era confirmar si la comunidad de uso de la TEI en español era de hecho consciente de las mejores prácticas ya en uso a nivel internacional.
- 21 Las respuestas revelaron que la mayoría de los proyectos recurren a personalizaciones del esquema TEI (40) y utilizan esquemas y ODD (39), que son algunos de los requisitos para un flujo de trabajo de marcado documentado, reproducible y sostenible.

- 22 En cuanto a los sistemas de bases de datos para organizar y recuperar datos, parece haber un mayor uso de bases de datos XML (23) por delante de las bases de datos relacionales más tradicionales (18). Esto está en consonancia con la evolución reciente y la llegada de productos de bases de datos XML de código abierto como eXist.¹¹
- 23 En lo que respecta a la transformación y renderizado de XML, el lenguaje más utilizado parece seguir siendo el ya veterano XSLT (29),¹² mientras que las transformaciones XQuery (11)¹³ se utilizan menos. Esto no parece coincidir del todo con la pregunta anterior sobre el uso de bases de datos XML, ya que la mayoría de las bases de datos XML nativas utilizan XQuery como principal herramienta de recuperación de datos, en lugar de XSLT. Entre los participantes existe una curiosa mezcla entre las formas *antiguas* y las *nuevas*.
- 24 Otras prácticas que los participantes adoptan cuando trabajan con TEI son la visualización de datos (31), y los vocabularios controlados o tesauros (22), así como las tecnologías relacionadas con la web semántica (16), los sistemas SIG (11), y el Procesamiento del Lenguaje Natural (12%). En *Otros*, algunos participantes explicaron que utilizaban scripts de interoperabilidad entre distintos esquemas (DCAT, DDI-CDI), anotación lingüística de corpus, XSLT - LaTeX - PDF y Cocoon. Lamentablemente, el 9% eligió otros sin especificar más.

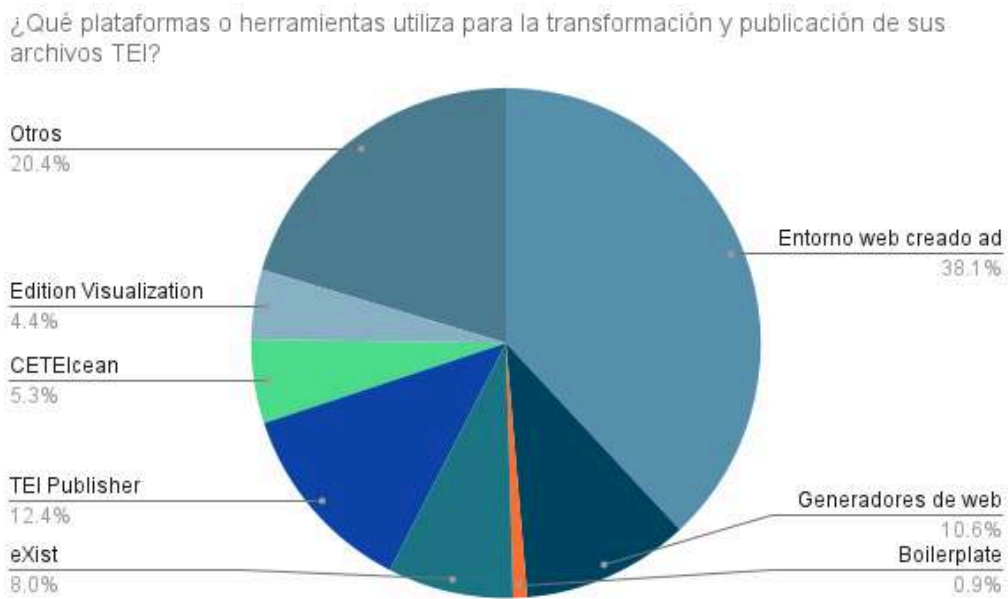
Figura 13. Tecnologías utilizadas en proyectos de edición digital con TEI.



- 25 La siguiente pregunta pretendía saber si tanto los proyectos individuales como los colectivos compartían sus archivos XML-TEI. El motivo era medir el compromiso de los participantes con la compartición del código fuente, que creemos que ayuda mucho al avance y la adopción de prácticas abiertas para la investigación científica (Kramer y Bosman 2018). Aunque los resultados son mejores de lo que esperábamos, la mayoría de los participantes (45) admitieron no compartir sus archivos, mientras que 34 respondieron afirmativamente, y para 26 participantes esta pregunta no era aplicable. El hecho de no compartir podría deberse a muchos factores, entre los que se cuentan la inseguridad en sus decisiones de marcado, la falta de tiempo para producir una versión final y optimizada, el desconocimiento de las mejores prácticas (por ejemplo, la elección de repositorios públicos y académicos), la preocupación por las malas prácticas en materia de derechos de autor, o incluso las restricciones de derechos de autor por parte de sus instituciones.
- 26 La encuesta también indagó sobre el uso de plataformas o herramientas por parte de los participantes para publicar sus archivos TEI. El objetivo era controlar si había alguna plataforma que se destacara por su uso. Por ello, propusimos las siguientes: 1. Infraestructura web creada *ad hoc* (por ejemplo, XML, XSLT, PHP, Python, etc.), 2. Generadores web estáticos (Jekyll, Gatsby, etc.), 3. Boilerplate, 4. eXist; 5. TEI Publisher, 6. CETEIcean, 7. Edition Visualization Technology, y añadimos una opción 8. Otros.
- 27 No es sorprendente que la puntuación más alta correspondiera a las infraestructuras creadas *ad hoc* (43). Así pues, la gran mayoría siente la necesidad de construir su propia infraestructura en términos de diseño web, funcionalidades, motor de búsqueda, presentación de las características del texto y publicación. La relevancia de estos datos radica en la escasez de plataformas que respondan a las necesidades de los profesionales. Sin embargo, entre las plataformas más utilizadas, los participantes eligen TEI Publisher (14),¹⁴ una herramienta diseñada en 2004 para crear repositorios basados en XML a través de una base de datos XML nativa con eXist que emplea además una biblioteca de motor de búsqueda de texto llamada Lucene.¹⁵ De hecho, los proyectos que se apoyan en la solución de base de datos eXist (9) se encuentran entre las primeras posiciones. A continuación, los generadores web estáticos aparecen como una opción viable y en alza, como Jekyll o Gatsby (11).¹⁶ El hecho de que los generadores web estáticos formen parte de las opciones mejor rankeadas responde probablemente a dos hechos: primero, el auge de la computación mínima que permite la creación de infraestructuras sin necesidad de servidores comerciales o

institucionales (por ejemplo, los sitios estáticos pueden vivir en servicios gratuitos como GitHub y GitHub Pages)¹⁷ y segundo, la falta de acceso a infraestructuras digitales especialmente en América Latina.¹⁸ Además, entre las otras plataformas para publicar archivos TEI algunos participantes respondieron estar usando CETELcean (6),¹⁹ Edition Visualization Technology (5),²⁰ mientras que sólo uno marcó la hoy bastante anticuada Boilerplate (1).²¹ Entre los que respondieron con la opción *Otros*, los participantes añadieron que reconocen su falta de conocimientos acerca de las tecnologías de transformación y publicación de archivos TEI (4), mientras que otros mencionaron estar utilizando otras opciones como Kiln,²² TEITOK,²³ R scripts,²⁴ o tener en mente para el futuro el uso de Textual Communities²⁵ o TEI Publisher.

Figura 14. Plataformas y herramientas para transformar y/o publicar TEI.

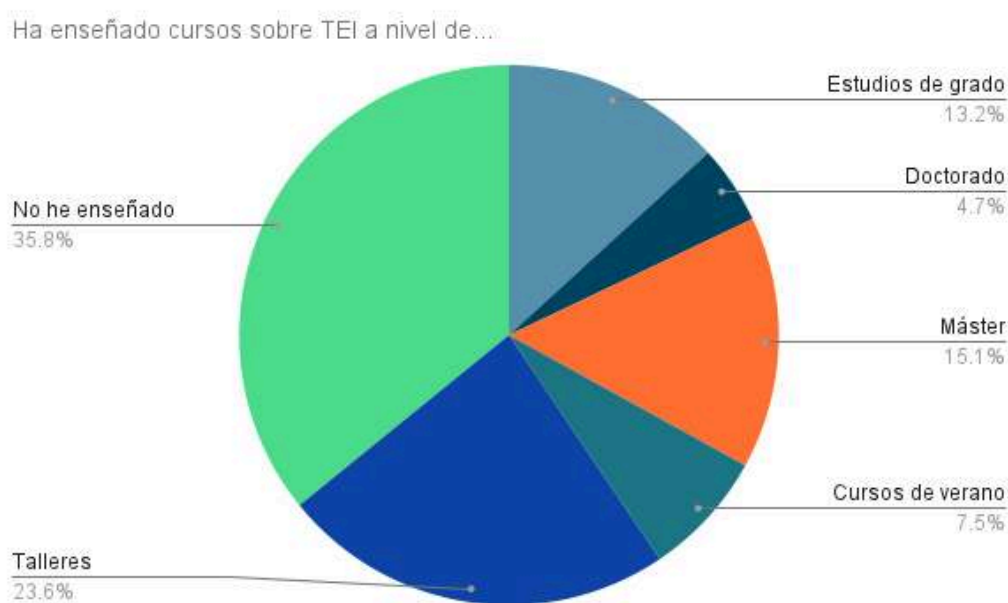


2.4 Enseñanza/aprendizaje de la TEI

- 28 El último bloque de nuestra encuesta pretendía recoger algunas ideas sobre el estudio y la formación en prácticas de codificación de textos, concretamente la TEI, y comprender mejor si los profesionales estaban satisfechos con los recursos existentes para enseñar y aprender a trabajar con este lenguaje de marcado.

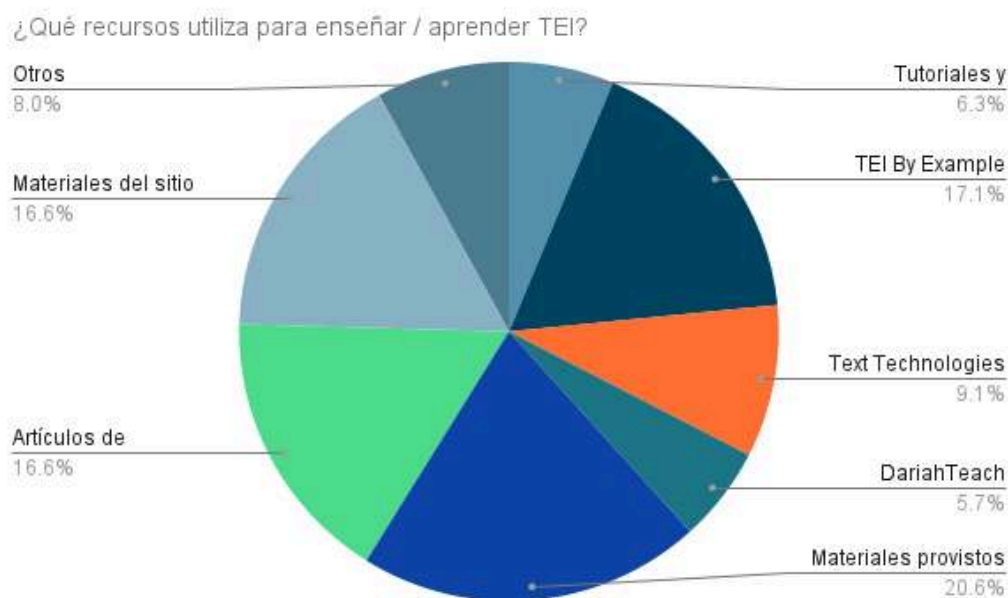
- 29 Empezamos preguntando si los participantes habían participado en la enseñanza de la TEI y a qué nivel, ofreciendo las siguientes opciones: 1. Licenciatura o estudios de grado, 2. Posgrado, 3. Máster, 4. Escuelas de verano, 5. Talleres, o bien 6. *No he enseñado cursos sobre TEI*.
- 30 La mayoría de los participantes (38) escogió esta última opción y declaró no haber impartido ningún curso. Sin embargo, entre los que sí lo habían hecho, el 23,6% (25) impartía talleres sobre el uso de la TEI, el 15,1% (16) cursos regulares de nivel MA y el 13,2% (14) cursos de licenciatura. Menor presencia parece tener la docencia TEI en Escuelas de Verano (7,5%, 8) y cursos de postgrado o doctorado (4,7%, 5). Como era de esperar, podríamos concluir diciendo que la codificación de textos aún no forma parte del plan de estudios establecido ni en los programas académicos de grado ni de posgrado.

Figura 15. Nivel de enseñanza de formación en codificación y publicación con TEI.



- 31 En relación con los recursos que los participantes utilizan para enseñar o aprender la TEI, dimos las siguientes opciones: 1. Tutoriales y recursos en línea; 2. TEI By Example; 3. TTHub; 4. DARIAHTeach; 5. Materiales proporcionados por instructores; 6. Artículos académicos; 7. Recursos del sitio web TEI-C, y 8. Otros. Las cifras fueron las siguientes:

Figura 16. Recursos para la enseñanza y el aprendizaje de la TEI.



- 32 El escenario más común es aquel en el que los instructores proporcionan sus propios materiales (20,6%, 36 respuestas), creados (suponemos) ad hoc para contextos específicos. Un dato curioso que surge es que el recurso más utilizado es TEI By Example (17,1%, 30 respuestas),²⁶ un recurso sólo disponible en inglés. Esto demuestra que el mundo hispanohablante sigue confiando en los recursos en inglés, seguido también por DARIAH Teach (5,7%, 10 respuestas).²⁷ El único recurso en español con cifras de uso no despreciables es nuestro TTHub (9%, 16 respuestas), que proporciona materiales educativos sobre codificación, edición y publicación de textos con TEI escritos en español y concebidos para un público hispanohablante.
- 33 Además, 29 participantes (16,6%) declararon utilizar los recursos enumerados en el sitio web de TEI-C; mientras que el mismo número recurre a artículos académicos que tratan sobre la TEI. Entre las respuestas dadas en la categoría *Otros*, algunos de los participantes insistieron en el hecho de que utilizan su propio material, o que no imparten clases, mientras que dos señalaron los cursos sobre el uso de la TEI de la UNED y uno de la Universidad de Birmingham que está relacionado con un proyecto de edición en curso.

2.5 Retos a la hora de usar la TEI

- 34 Para la pregunta sobre los retos a la hora de trabajar con la TEI dimos una serie de opciones que los participantes debían clasificar (1. Escasez de infraestructuras, 2. El trabajo aislado, 3. Ausencia de una disciplina como la filología o la crítica textual en el currículo académico, 4. Falta de financiación, 5. Falta de conocimiento sobre la TEI en su comunidad académica, 6. Imposibilidad de contar con un grupo de investigación híbrido (humanistas, programadores, bibliotecarios, etc.), 7. Falta de traducción de las Directrices de la TEI, 8. Falta de disponibilidad de formación técnica y de programación, y 9. Otros.

Figura 17. Obstáculos para trabajar con la TEI.



- 35 La mayor preocupación entre los participantes parece ser el hecho de que en su comunidad académica todavía no se entienda lo que es la TEI (15,8%, 33). Probablemente relacionado con esta preocupación está el malestar por el trabajo individual (13,9%, 29), sin apoyo de la comunidad, de otros académicos o de la institución. El siguiente reto indicado por los participantes corresponde a la falta de disponibilidad de formación técnica y de programación (14,8%, 31), seguido de la falta

de financiación (12,4%, 26) y la imposibilidad de contar con un grupo de investigación híbrido (humanistas, programadores, bibliotecarios, etc.) (12%, 25), así como la falta de infraestructuras (9,6%, 20).

- 36 Entre las opciones que ofrecimos queríamos que los participantes consideraran la falta de una disciplina como la filología o la crítica textual en el currículo académico (11.5%, 24). Tal y como nosotros lo vemos, esta cuestión tiene tres vertientes: en primer lugar, la filología y la crítica textual han impuesto, desde finales del siglo XIX, metodologías relativamente sencillas y claras para editar y presentar textos, pero este no es todavía el caso de la contraparte digital. En consecuencia, carecemos de un *cómo* que se inscriba en un marco teórico (Allés-Torrent 2020). Por último, esta tradición e interés por la filología, si bien tiene una larga historia en España, ha sido escasa o nula en América Latina, donde la mayoría de las universidades ni siquiera incluyen estas disciplinas en sus planes de estudio.
- 37 Más adelante preguntamos: “¿Qué tipo de recursos cree que necesitamos para mejorar la enseñanza y el aprendizaje de la TEI?” Algunos de los participantes insistieron en la necesidad de recursos y materiales de formación a todos los niveles y en español, incluyendo además temas específicos (por ejemplo, transformaciones con XSLT), así como otros tipos de recursos, como referencias bibliográficas. La necesidad de formación formal e informal dentro y fuera del mundo académico surge también como otra preocupación legítima. Entre las demandas de formación aflora el desconocimiento de las etapas finales del proceso editorial, es decir, la transformación de los archivos codificados en XML-TEI y su publicación en línea. Esto parece indicar que la comunidad está deseosa de avanzar en sus conocimientos técnicos. El deseo de contar con editores XML libres y fiables también es evidente, de hecho, uno de los editores XML más populares en uso es propietario y los profesionales, especialmente de América Latina, no pueden permitirse su compra o simplemente no quieren confiar en el software comercial. Además, y como preocupación genuina, encontramos demandas claras sobre la necesidad de financiación para realizar un proyecto de edición digital con el fin de remunerar la mano de obra y costear el soporte técnico.
- 38 Algunos de los participantes insistieron en esta sección en la necesidad de divulgación a distintos niveles para avanzar y, tal vez, resolver la ya mencionada falta de comprensión por parte de la comunidad académica, mejorar el diálogo dentro de los equipos e incluso promover la interoperabilidad real entre otros lenguajes de marcado y tecnologías.

3. Conclusiones

- 39 La TEI es una de las iniciativas globales más sólidas y extendidas para la codificación de textos de cualquier tipo de fuente escrita en Humanidades y Ciencias Sociales. Los escenarios en los que se aplica la TEI van desde la edición digital de textos hasta la creación de corpus lingüísticos o la creación de bases de datos nativas en XML. La TEI está gestionada por un consorcio formado por personas e instituciones, principalmente de Norteamérica y Europa. Este consorcio se encarga de elaborar y actualizar las guías directrices para la codificación de textos en formato digital, un conjunto de buenas prácticas y recomendaciones que especifican instrucciones generales para marcar textos de modo que puedan ser procesados por ordenador. El consorcio de la TEI está insuflado de un espíritu de apertura y colaboración que hace que los propios usuarios puedan proponer mejoras y modificaciones a través de su repositorio en GitHub. Este espíritu de colaboración y difusión se manifiesta también a través de un diálogo continuo mediante una lista de discusión en línea y la organización anual de un congreso internacional. Por último, el consorcio es un gran paraguas que da lugar a diferentes grupos de trabajo y de interés para determinados temas, como la creación de diccionarios, el trabajo con lingüística de corpus, o la descripción de manuscritos, entre otros. Sin embargo, lo anterior se aplica principalmente al mundo anglófono, donde, como decíamos al principio, nació la TEI, pero esto no quiere decir que no contemos con proyectos e iniciativas en países hispanohablantes. Con este estudio hemos tratado de profundizar en una comunidad particular de practicantes, no sólo definida por la lengua de los textos con los que trabaja, el español, sino también por sus peculiaridades y necesidades culturales y sociales específicas. En este intento internacional por parte de la comunidad de la TEI de respetar la diversidad e incluir otros idiomas además del inglés, está más claro que nunca que tanto el conocimiento lingüístico como el cultural son necesarios si se quiere entender a las personas y a sus comunidades (Modern Language Association 2007). La recopilación de los datos más destacados de nuestra encuesta ha buscado ofrecer un sucinto estado de la cuestión sobre el uso de la TEI en la actualidad para textos en español. Las respuestas de los participantes nos han ayudado a constatar algunas de nuestras percepciones. Desde el punto de vista geográfico, ahora tenemos claro que la mayoría de los estudiosos y profesionales de la TEI tienen su sede en España, seguidos de México, Estados Unidos, Colombia y Argentina. Las diferencias en el lugar de origen y residencia actual y la institución reflejan algunas tendencias mundiales globales; por

ejemplo, los académicos dejan España y buscan insertarse en otras universidades; mientras que los Estados Unidos reciben académicos del extranjero. La encuesta también ha ofrecido una primera exploración de quiénes son esos profesionales que usan la TEI: en su mayoría académicos de diferente rango, pero también técnicos de apoyo a la investigación, bibliotecarios y archivistas y, en menor medida, desarrolladores de software.

- 40 Un segundo gran bloque de preguntas se refería a la formación académica en la TEI, y puso de manifiesto que la mayoría de los cursos se realizan durante el periodo de máster y/o doctorado, pero en contextos informales, ya que, en la mayoría de los casos, la TEI no forma parte de los programas oficiales. Otro dato interesante es que los profesionales se inician en su uso lo hacen porque forman parte de un proyecto de investigación.
- 41 El tercer bloque trató sobre el interés en el uso de la TEI. Los alicientes más populares son la creación de ediciones digitales, un interés general por los lenguajes de marcado y sus posibilidades digitales y, de nuevo, la TEI como habilidad técnica a contratar en un determinado proyecto. Asimismo, pudimos descubrir que, en términos de compromiso, la comunidad hispanohablante no invierte en la afiliación individual o institucional al consorcio. Queda por dilucidar si esto se debe a la falta de financiación o a la falta de interés. Un caso similar es la baja participación en el congreso anual de la TEI, donde casi el 80% de los participantes nunca ha presentado una ponencia. Sin embargo, surgieron algunas buenas noticias desde el punto de vista de las prácticas y plataformas digitales, destacando especialmente el número de proyectos que comparten sus archivos TEI en abierto.
- 42 El tercer y último bloque tenía como objetivo explorar cómo se aprende y se enseña la TEI. En cuanto a la enseñanza, la mayoría de los profesionales no enseñan, y si lo hacen esto se da en contextos informales, como talleres, cursos breves en programas de máster o escuelas de verano. Ha quedado claro que no existen textos o recursos canónicos para enseñar o aprender, y la mayoría de los participantes reconoce utilizar recursos anglófonos como TEI By Example. En cualquier caso, nuestro TTHub parece abrirse camino para convertirse en un recurso esencial dentro de este panorama, ya que es el único con materiales originales en español y con ejemplos de la literatura o la historia hispanoamericana. Por último, creemos haber evidenciado algunos de los retos que preocupan a nuestra comunidad, a saber, el desconocimiento de la TEI entre la comunidad académica y la falta de disponibilidad para la programación y la alfabetización digital en general.

Además, diríamos que han surgido de forma audible otras dos preocupaciones, en primer lugar, la necesidad de materiales de formación para todo el proceso de labor editorial, con especial insistencia en la transformación del XML-TEI (XSLT, XQuery) y publicación de los archivos TEI. Esto significa que la comunidad ya ha superado las fases iniciales de familiarización con la TEI, y que ahora urgen temas más avanzados. En segundo lugar, han aparecido varias voces que defienden la necesidad de contar con editores XML gratuitos. El hecho de que el software más popular sea propietario desanima a algunos usuarios. Ni que decir tiene que ya se han dado algunos pasos para adaptar software de código abierto para trabajar con archivos TEI, como Visual Studio Code, que se ha beneficiado del desarrollo del plugin Scholarly XML, que permite una codificación básica pero rigurosa en XML-TEI.²⁸

- 43 En resumen, la encuesta nos permitió comprender mejor no sólo a la comunidad hispanohablante que utiliza XML-TEI, sino también pensar en estrategias que puedan contribuir con prácticas más inclusivas para los académicos de países menos representados y en contextos económicos y sociales menos favorables. Por último, pero no menos importante, creemos que la encuesta será útil para diseñar acciones que puedan apoyar una gama más amplia de modos de interacción y colaboración –por ejemplo, entre la comunidad latina en los Estados Unidos de Norteamérica y la hispanoamericana– dentro de la comunidad TEI global.

BIBLIOGRAFÍA

- Allés-Torrent, Susanna. 2020. “Crítica textual y edición digital o ¿dónde está la crítica en las ediciones digitales?” *Studia Aurea* 14. <https://doi.org/10.5565/rev/studiaaurea.395>.
- Allés-Torrent, Susanna, Gabriel Calarco y Gimena del Rio Riande. 2018-. “TTHub. Text Technologies Hub. Recursos sobre tecnologías del texto y edición digital.” <https://tthub.io/>.
- Allés-Torrent, Susanna y Gimena del Rio Riande. 2020. “The Switchover: Teaching and Learning the Text Encoding Initiative in Spanish.” *Journal of the Text Encoding Initiative* 12. <https://doi.org/10.4000/jtei.2994>.
- Isasi Velasco, Jennifer y Gimena del Rio Riande. 2022 “¿En qué lengua citamos cuando escribimos sobre Humanidades Digitales?” *Revista de Humanidades Digitales* 7: 127-143. <https://doi.org/10.5944/rhd.vol.7.2022.36280>.

- Kramer, Bianca y Jeroen Bosman. 2018. “Rainbow of open science practices.” Zenodo. <https://doi.org/10.5281/zenodo.1147025>.
- Modern Language Association. 2007. “Foreign Languages and Higher Education: New Structures for a Changed World.” *Profession* 1, 234–245. https://app.dimensions.ai/discover/publication?and_facet_journal=jour.1053132&and_facet_source_title=jour.1142962<https://doi.org/10.1632/prof.2007.2007.1.234>.
- Okada, Kazuhiro, Satoru Nakamura y Kiyonori Nagasaki. s.f. SIG East Asian/Japanese, TEI Consortium. <https://tei-c.org/Activities/SIG/EastAsian/>.
- Rahtz, Sebastian. 2005. “TEI Internationalization.” <https://tei-c.org/Vault/Talks/2005/Wuerzburg/i18n.pdf>.
- . 2006. “Towards an Internationalized and Localized TEI.” https://ora.ox.ac.uk/objects/uuid:69a8d091-6d15-4554-b513-a3e01da5a5a2/download_file?file_format=application%2Fpdf&safe_filename=Towards_an_internationalized_and_localized_TEI.pdf&type_of_work=Conference+item.
- Stadler, Peter y Martina Scholger. 2016. “Report on the TEI2German Translatathon at Vienna 2016.” <https://tei-c.org/tools/i18n/tei2german>.
- TEI Consortium. 2023. *TEI P5: Guidelines for Electronic Text Encoding and Interchange*. Versión 4.6.0. <https://tei-c.org/Vault/P5/4.6.0/release/doc/tei-p5-doc/en/html/index.html>.

NOTAS

- 1 Una semblanza de Sebastian Rahtz puede leerse en Wikipedia: fecha de consulta 16 de julio de 2023, https://en.wikipedia.org/wiki/Sebastian_Rahtz
- 2 El proyecto “Communicating the Text Encoding Initiative to a Multilingual User Community” fue financiado por la Andrew W. Mellon Foundation (Estados Unidos) con la referencia 2001-07353. Véanse los detalles de la subvención en: fecha de consulta 16 de julio de 2023, <https://mellon.org/grants/grants-database/grants/duke-university/2001-07353/>. Asimismo, los orígenes de este proyecto que empezó a gestarse años atrás pueden encontrarse en el evento de TriangleSci organizado en 2019: fecha de consulta 31 de julio de 2023, <https://trianglesci.org/2019/07/17/communicating-the-tei-to-a-multilingual-user-community/>.
- 3 Más información sobre este grupo de trabajo en: fecha de consulta 16 de julio de 2023, <https://tei-c.org/activities/workgroups/internationalization-i18n-workgroup/>.

- 4 TranslateTEI ha sido desarrollada por Hugh Cayless (Duke University) y está disponible en: fecha de consulta 16 de julio de 2023, https://translate.tei-c.org/doc_en.html
- 5 Las especificaciones de la TEI son las descripciones de los elementos de los diferentes módulos.
- 6 TTHub: Hub de Tecnologías del Texto. Recursos sobre tecnologías del texto y edición digital, fecha de consulta 16 de julio de 2023, <https://tthub.io/> (Allés-Torrent et al. 2021-).
- 7 HD Lab: Laboratorio de Humanidades Digitales del CONICET, fecha de consulta 16 de julio de 2023, <https://hdlab.space/>
- 8 Las listas de distribución donde se mandó la encuesta fueron: Red de Humanidades Digitales (<mailto:redhumanidadesdigitales@gmail.com>), Humanidades Digitales de América Latina (<mailto:latamhd@lists.humanidadesdigitales.net>), Humanidades Digitales Hispánicas (<mailto:HDH@listserv.rediris.es>), Asociación Argentina de Humanidades Digitales (<mailto:aahd-l@uleth.ca>), TEI-List (<mailto:TEI-L@listserv.brown.edu>), Lista Latinoamericana sobre Acceso Abierto y Repositorios (<mailto:llaar@googlegroups.com>), Humanist (<mailto:humanist@dhhumanist.org>), Air-L (<mailto:airl@listserv.aoir.org>), Digital Classicist (<mailto:digitalclassicist@jiscmail.ac.uk>), Digital Medievalist (<mailto:dm-l@uleth.ca>), Medieval Text - Philology, Codicology and Technology (<mailto:medtextl@lists.illinois.edu>).
- 9 Qualtrics, fecha de consulta 16 de julio de 2023, <https://www.qualtrics.com>
- 10 Más información sobre los cursos del LINHD en, fecha de consulta 16 de julio de 2023, <https://linhd.uned.es/>, para la ENC, fecha de consulta 16 de julio de 2023, <https://www.chartes.psl.eu/fr/rubrique-enseignements/masters>, el CESR, fecha de consulta 16 de julio de 2023, <https://cesr.univ-tours.fr/>, o el de King's College, fecha de consulta 16 de julio de 2023, <https://www.kcl.ac.uk/study/postgraduate-taught/courses/digital-humanities-ma>.
- 11 Exist Database, fecha de consulta 16 de julio de 2023, <http://exist-db.org/>
- 12 XSLT, fecha de consulta 16 de julio de 2023, <https://www.w3.org/TR/xslt/>.
- 13 XQuery, fecha de consulta 16 de julio de 2023, <https://www.w3.org/XML/Query/>.
- 14 TEI Publisher, fecha de consulta 16 de julio de 2023, <https://teipublisher.com/index.html>
- 15 Exist Database, fecha de consulta 16 de julio de 2023, <http://exist-db.org/>.
- 16 Para más información véase sobre Jekyll, fecha de consulta 16 de julio de 2023, <https://jekyllrb.com/> y Gatsby, fecha de consulta 16 de julio de 2023, <https://www.gatsbyjs.com/>.
- 17 GitHub Pages, fecha de consulta 16 de julio de 2023, <https://pages.github.com/>.

- 18 Para más información sobre el movimiento de la computación mínima, véase la página del grupo Minimal Computing, fecha de consulta 16 de julio de 2023, <https://go-dh.github.io/mincomp/about/>.
- 19 CETEIcean, fecha de consulta 16 de julio de 2023, <https://github.com/TEIC/CETEIcean>.
- 20 EVT, fecha de consulta 16 de julio de 2023, <http://evt.labcd.unipi.it/>.
- 21 TEI Boilerplate, fecha de consulta 16 de julio de 2023, <https://dcl.ils.indiana.edu/teibp/>.
- 22 Kiln, fecha de consulta 16 de julio de 2023, <https://kiln.readthedocs.io/en/latest/>.
- 23 TEITOK, fecha de consulta 16 de julio de 2023, <http://teitok.clul.ul.pt/>.
- 24 R Scripts, fecha de consulta 16 de julio de 2023, http://mercury.webster.edu/aleshunass/R_learning_infrastructure/R%20scripts.html.
- 25 Textual Communities, fecha de consulta 16 de julio de 2023, <https://textualcommunities.org/app/>.
- 26 TEI By Example, fecha de consulta 16 de julio de 2023, <https://teibyexample.org/>.
- 27 DARIAH Teach, fecha de consulta 16 de julio de 2023, <https://teach.dariah.eu/>.
- 28 Este trabajo se debe a Raffaele Vigiante (Maryland Institute for Technology in the Humanities) y el plugin puede descargarse en línea en: fecha de consulta 16 de julio de 2023, <https://marketplace.visualstudio.com/items?itemName=raffazizzi.sxml>.

AUTORS

GIMENA DEL RIO RIANDE

Gimena del Rio Riande es investigadora adjunta del Instituto de Investigaciones Bibliográficas y Crítica Textual (IIBICRIT) del Consejo Nacional de Investigaciones Científicas (CONICET), donde dirige el laboratorio de humanidades digitales (HD Lab). Es fundadora de la Asociación Argentina de Humanidades Digitales (AAHD) y directora de la Diplomatura en humanidades digitales de la Universidad de Ciencias Empresariales y Sociales (UCES). Desde 2018 es miembro de Board of Directors de la Text Encoding Initiative. ORCID: <https://orcid.org/0000-0002-8997-5415>.

SUSANNA ALLÉS-TORRENT

Susanna Allés-Torrent es profesora en el Department of Modern Languages and Literatures de la University of Miami (Fl, Estados Unidos), donde enseña cursos sobre literatura medieval y del Renacimiento y Humanidades digitales. ORCID: <https://orcid.org/0000-0002-3616-2285>