

## **ESCUELA DE VERANO UNLP 2016**

### **1. Denominación del Curso:**

**“INTRODUCCIÓN AL MARCADO SEMÁNTICO DE TEXTOS (TEI-XML) PARA SU EDICIÓN DIGITAL”**

### **2. Docentes a cargo:**

- Docente Coordinador por la UNLP: **Dra. Natalia Corbellini**, Profesora Adjunta Ordinaria cátedra de Literatura española II, FAHCE - UNLP.
- Docente invitado de otra universidad: **Dra. Gimena del Rio Riande**, docente de grado y posgrado en la Facultad de Filosofía y Letras UBA, investigadora asistente del CONICET en el Seminario de Edición y Crítica Textual (SECRIT-IIBICRIT).

### **3. Fundamentación:**

El impacto que la tecnología y la implementación de la web 2.0 en las ciencias humanas no ha alcanzado aún su potencialidad plena debido en parte a la falta de formación en herramientas y metodologías digitales de los investigadores. Por otro lado, existen centros científicos y grupos significativos de investigadores que han desarrollado metodologías y consensuado estándares para la investigación, el intercambio de información, y para potenciar la transmisión de conocimientos también en las ciencias humanas, desarrollos que conocemos hoy como Humanidades Digitales.

En ese contexto, el Text Encoding Initiative (TEI) es el producto de una comunidad de prácticas activa desde los años 80, que ha desarrollado un estándar de marcación semántica, actualmente basado en XML, usada por gran cantidad de bibliotecas, colecciones y corpora lingüísticos. La diferencia con los otros marcados textuales (como HTML u OD) es que no se ocupa de cómo debe aparecer el texto, sino que establece el significado de cada elemento y atributo de un texto. Las ventajas van tanto desde permitir una mayor fluidez en la edición porque permite crear fácilmente múltiples presentaciones y visualizaciones, pero también transforma la investigación y el trabajo con el texto y corpus, ya que la tecnología permite interrogar cantidades enormes de datos o reorganizar información de manera dinámica. Esto no es sólo una capacidad del TEI/XML, pero su potencial de estructurar información textual

permite búsquedas más potentes, creación dinámica de resultados, y ofrece una alternativa sólida para la preservación de datos en el largo plazo. En palabras de Shillingsburg “un mundo ‘democratizado’ de la investigación académica con una metodología modularizada, accesible por marcos técnicos fuera del entorno inmediato de un proyecto determinado”.

Las ventajas de las Humanidades Digitales tienen como correlato la necesidad de una formación específica para comprender los alcances y potencialidades en cada campo del saber específico y para cada proyecto particular. Por eso se propone un curso intensivo introductorio que permita a los tesisistas, investigadores y profesionales entrar en contacto con una herramienta específica y una metodología de trabajo.

#### **4. Objetivos:**

Que los participantes:

- Adquieran conocimientos sobre las posibilidades tecnológicas de tratamiento de textos mediante diferentes lenguajes de marcado para la estructuración, presentación y descripción.
- Conozcan y sepan implementar técnicas de marcación digital de textos, en particular los estándares TEI/XML para textos de humanidades y estudios sociales.
- Conozcan y puedan incorporar como herramientas de análisis las diferentes TEI *Guidelines* (diccionarios, teatro, corpus lingüísticos, transcripción de fuentes primarias, etc.)
- Se familiaricen con diferentes análisis posibles sobre textos marcados, ya sea textos de la tradición literaria, o de diferentes corpora lingüísticos, así como de textos de estudios sociales plausibles de ser analizados.
- Se familiaricen con la transformación del lenguaje marcado a diferentes formatos para su exportación, análisis o publicación como recurso compartido.

#### **5. Perfil del estudiante:**

Graduados o profesionales que necesiten contar con herramientas para analizar grandes volúmenes de textos. Graduados o profesionales interesados en la edición digital de textos académicos que necesiten estructurarlos semánticamente. Graduados o profesionales que trabajen en la edición crítica de textos literarios, filosóficos, históricos, testimonios de textos antiguos, o ediciones filológicas tradicionales. Graduados interesados en profundizar sobre la actualidad en el campo edición digital científica.

No son necesarios conocimientos específicos previos, aunque se recomienda familiarizarse con los lenguajes XML antes de comenzar el curso para el mejor aprovechamiento de éste.

## **6. Contenidos:**

Unidad 1: Presentación. Reseña histórica de los diferentes sistemas de marcado. Usos y limitaciones. Web 2.0.

Unidad 2: Desarrollo y actualidad de los estándares de marcación presentacionales para textos digitales. Diferentes formatos de publicación y circulación.

Unidad 3: El proyecto de Text Encoding Initiative. Las soluciones específicas que brinda para el área de humanidades. Diferentes guías de procesamiento y núcleos de marcado. La importancia de los encabezados y descriptores de los textos.

Unidad 4: La edición por marcación semántica de textos literarios. Las opciones de marcación diferentes géneros; la segmentación de textos; las variaciones filológicas e históricas. Ejemplos de proyectos en marcha de y proyectos ya publicados.

Unidad 5: Las potencialidades de análisis. Modos de lectura. Ejemplos de casos

**7. Modalidad:** El curso se dictará a distancia a través de la plataforma de la UNLP.

## **8. Metodología:**

El curso se organizará en 5 etapas, correspondiente cada una a las unidades desarrolladas en los contenidos. Cada etapa contará con una presentación por parte de las docentes del tema y del material a utilizar.

Las dos primeras unidades, con temas más historiográficos y descriptivos, contarán con una exposición extensa con ejemplos por parte de las docentes, y con textos teóricos breves sobre diferentes consideraciones sobre los estándares de marcación. Estos contenidos los alumnos deberán reponerlos luego en un cuestionario sencillo con el formato verdadero/falso para cada unidad. Se propone así un formato que en poco tiempo permita a los alumnos contar con un acercamiento al tema y estado de la cuestión de la edición digital científica, y que a su vez tome distancia crítica de la metodologías propuestas. Se prevé un tiempo de 5 horas para cada unidad para la visualización y lectura del material, así como la resolución del cuestionario.

La unidad 3 luego de la presentación se trabajará a partir de ejemplos de textos crudos y marcados para ir ejemplificando el aumento de la complejidad de la marcación hasta llegar a los estándares actuales de TEI/XML. Se propondrán ejercicios de escritura y procesamiento de textos breves pero de distintas áreas temáticas para que entren en contacto con los diversos núcleos y guías de marcación semántica.

La unidad 4 profundizará en la marcación de diferentes formatos (verso, prosa, paratextos, variantes históricas,...). Se propondrán ejercicios de procesamiento de textos y visualización de

los trabajos realizados por las docentes en los sendos proyectos colectivos de los que participan (Diálogo Medieval y ArteLope).

Se prevé un tiempo de 6 horas para cada unidad para la visualización, lectura del material, y práctica con el proceso de los textos. La evaluación se realizará mediante los ejercicios de procesamiento de textos.

La unidad 5 contará con exposiciones de las docentes y presentaciones de resultados de proyectos de Humanidades Digitales que tienen como base la marcación TEI/XML. Se propondrá a los alumnos que elaboren una hipótesis de propuesta de análisis factibles sobre su propio material o corpus de trabajo con el formato o la utilización de la tecnología TEI/XML. Se prevé para este último núcleo 6 horas de trabajo. La propuesta final de análisis podrá ser grupal en el caso de estudiantes que provengan del mismo campo de estudios.

Se establecerán horarios y modos de tutorías por parte de las docentes para la elaboración de los trabajos con las 2 horas restantes de la planificación.

#### **9. Forma de Evaluación y fecha límite de presentación:**

La evaluación se hará por cada unidad dado que los contenidos se presentarán de modo progresivo. Como evaluación final, se pedirá la propuesta esbozada en la Etapa 5 desarrollada como proyecto: objetivo general, objetivo específico, resumen de la propuesta, hipótesis, materiales de los que se parte, metodología y resultados esperados. Se evaluará que el estudiante haya incorporado las herramientas de las Humanidades digitales a su metodología de trabajo. La fecha límite para su presentación será el 31 de marzo.

#### **10. Bibliografía:**

BORGMAN, Christine, "The Digital Future is Now: A Call to Action for the Humanities", en *Digital Humanities Quarterly*, vol. 3, 4, 2009.

Burnard, L. 2014. *What is the Text Encoding Initiative? How to add intelligent markup to digital resources*. Marseille: OpenEdition Press. Extraído de <http://books.openedition.org/oep/426> ISBN: 9782821834606.

BUZZETTI, Dino y MCGANN, Jerome, «Editing in a Digital Horizon», en *Electronic textual editing*, ed. Lou Burnard, Katherine O'Brien O'Keefe y John Unsworth, New York, Modern Language Association of America, 2006.

Corbellini, Natalia, "Edición digital de *Lo fingido verdadero*: herramientas para leer traducciones y tradiciones" *Teatro de palabras. Revista sobre teatro áureo*, nro. 7 junio de 2013. Québec: L'Université du Québec à Trois-Rivières, 2013, pp. 193-205, ISSN 1911-0804

González-Blanco García, E., Martínez Cantón, Martos, del Río Riande G.. Una propuesta de integración del sistema de formularios de bases de datos MYSQL con etiquetado TEI: ReMetCa...”, en *Humanidades Digitales: desafíos, logros y perspectivas de futuro*, Janus, Anexo 1, 2014, pp. 209-219.

MUÑOZ PONS, Carlos, «El proceso de edición digital en Artelope y CTCE», en *Teatro de palabras. Revista sobre teatro áureo*, nro. 7 junio de 2013. Québec: L'Université du Québec à Trois-Rivières.

PIEZ, Wendell, “Towards Hermeneutic Markup: An architectural outline”, ponencia en congreso *Digital Humanities 2010*, King’s College London. Otros ejemplos

RENEAR, Allen, “Text Encoding”, en *A companion to Digital Humanities*, ed. Susan Schreibman, Ray Siemens y John Unsworth, Oxford, Blackwell, 2004, pp. 222-224.

SHILLINGSBURG, Peter L., *From Gutenberg to Google: electronic representations of literary texts*, Cambridge, Cambridge University Press, 2006.

Spence, Paul, “Teatro clásico y humanidades digitales: el cruce entre método, proceso y nuevas tecnologías”, *Teatro de palabras. Revista sobre teatro áureo*, nro. 7 junio de 2013. Québec: L'Université du Québec à Trois-Rivières, Número 7, 2013.

SVENSSON, Patrik, «Envisioning the Digital Humanities», en *Digital Humanities Quarterly*, vol. 6, 1, 2012.

Text Encoding Initiative: <http://www.tei-c.org/About/>

#### **11. Dirección de e-mail del profesor Coordinador:**

Dra. Natalia Corbellini: [ncorbellini@fahce.unlp.edu.ar](mailto:ncorbellini@fahce.unlp.edu.ar)