Seminario di informatica umanistica

Marta Materni (chercheur post-doc, Marie Curie Fellow, Université Grenoble Alpes)

Padova, 22-24 aprile 2020

**La Galassia Busa-Bush: analisi e rappresentazione del testo in epoca digitale**

Benché la presenza del digitale nel nostro quotidiano costituisca ormai una sorta di ovvietà, la coscienza critica rispetto a questa presenza è ancora lontana dall’essere solida. Per limitarsi all’ambito accademico, nonostante i decenni di pratica ed esperimenti e nonostante un *Manifesto delle Digital Humanities* formulato nel 2010, la discussione identitaria – cosa sono le Digital Humanities? come si inquadrano le Digital Humanities nella cornice accademica? – è ancora oggi un argomento scottante. Il recente volume di sintesi di Pierre Mounier, *Les Humanités Numériques: une histoire critique* (2018), riassume perfettamente già nel titolo il problema.

Rispetto a questo mare magnum, il seminario si concentrerà su un aspetto specifico: l’interazione possibile tra lo strumento/metodologia digitale e il Testo, in una prospettiva filologica. Da una parte, nel solco di Roberto Busa, padre nobile della linguistica computazionale, si offrirà un panorama degli strumenti informatici di analisi e rappresentazione del testo; un’attenzione particolare verrà riservata alla codifica testuale in XML-TEI, che attualmente costituisce lo standard nell’ambito delle edizioni digitali. Dall’altra, nel solco di Vannevar Bush, inventore del Memex, si introdurrà la questione della nascita di una nuova forma testuale, l’ipertesto, per la quale sarà forse necessario fondare una nuova filologia, la filologia del testo digitale.

**22 aprile.** Sala “Carlo Diano” (ex Presidenza), presso il Liviano

*Una Archéologie du savoir o epistemologia della scienza informatica applicata al testo*

9.00-11.00 Digital Humanities/Humanités Numériques/Informatica Umanistica: un problema di definizione identitaria. La galassia Busa-Bush: *Index Tomisticum* e Ipertesto.

11.30-13.30 Nozioni di base del pensiero informatico: decontestualizzazione e formalizzazione. Lessico tecnico indispensabile.

14.30-16.30 Filologia digitalizzata/Filologia computazionale/Filologia del digitale: nuovi strumenti per vecchie metodologie e vecchie metodologie per nuovi contesti testuali.

**23 aprile.** Aula 5 del Capitaniato (Palazzo del Capitanio)

*La codifica del testo come forma di rappresentazione della conoscenza*

9.00-11.00 Introduzione ai linguaggi di marcatura del testo (XML, HTML): sistema WYSIWYG *vs.* sistema WYSIWYM. Grammatica e sintassi dei linguaggi di marcatura.

11.30-13.30 Introduzione alla TEI (*Text Encoding Initiative*): il modello testuale proposto dalla TEI; le Guidelines TEI; il problema strumentale.

14.30-16.30 Sessione pratica. Primi passi con un file XML-TEI.

**24 aprile.** Sala “Carlo Diano” (ex Presidenza), presso il Liviano

*E dopo la formalizzazione dei dati? Visualizzazione e analisi*

9.00-11.00 Sessione pratica. Costruzione parallela di un file XML-TEI e di un file HTML. Introduzione ai principi dei linguaggi XSLT e CSS.

11.30-13.30 Dalla rappresentazione all’analisi (1). L’analisi linguistica: lemmatizzazione, etichettatura morfosintattica e concordanze.

14.30-16.30 Dalla rappresentazione all’analisi (2). L’analisi testuale: la sfida della collazione automatica.