Nell'àmbito delle attività del progetto ADATest, diretto da Antonio Montinaro (ricercatore in Linguistica italiana dell'Università del Salento) e promosso dal programma di ricerca denominato FutureInResearch, la dott.ssa Elena Artale (OVI-CNR Firenze) giovedì 26 maggio, ore 10.00, presso l'aula Ferrari di palazzo Codacci Pisanelli, terrà un seminario dal titolo *Il Corpus OVI dell'italiano antico: occorrenze, forme e lemmi*.

Progetto ADATest

Il progetto ADATest (Archivio Digitale degli Antichi Testi di Puglia) è finalizzato alla ricomposizione virtuale dei testi in italoromanzo prodotti in Puglia nel corso dei secoli XIII-XVI o ad essa riconducibili. I testi, di natura composita (letteraria, paraletteraria e documentaria), sono tramandati da manoscritti e stampe in caratteri latini, greci ed ebraici attualmente dispersi in biblioteche e archivi italiani e stranieri.

Tramite la Banca dati è possibile consultare le schede filologiche dedicate alle opere e ai loro testimoni, effettuando articolate ricerche. È in fase di allestimento l'archivio che consentirà anche la visione e l'interrogazione dei testi. Nella sezione Studi si riuniscono pubblicazioni attinenti all'oggetto della ricerca.

ADATest, sviluppo di ADAMaP 1.0, è promosso dal programma di ricerca denominato FutureInResearch.

www.adamap.it

Seminario di Elena Artale - *Il Corpus OVI dell'italiano antico: occorrenze, forme e lemmi* (giovedì 26 maggio 2016, ore 10.00, palazzo Codacci Pisanelli, aula Ferrari)

Il *corpus OVI* (*Opera del Vocabolario Italiano*) è la maggiore banca dati attualmente disponibile delle varietà italo-romanze medievali. Il suo nucleo principale è costituito dal *corpus TLIO*, su cui si basa la redazione del *Tesoro della lingua Italiana delle Origini* (il *TLIO*, appunto), e che conta attualmente 2114 testi e 445.776 forme diverse, per un totale di 21.920.159 occorrenze.

Il *corpus TLIO* è lemmatizzato esaustivamente sulle forme ma non sulle occorrenze (quelle lemmatizzate sono ad oggi 3.670.753), con una lemmatizzazione 'critica', che nasce dal vaglio dei singoli contesti.

Il seminario focalizzerà l'attenzione su questo tipo di lemmatizzazione, sulle sue problematiche e sull'utilità che essa riveste sia per la redazione sia per la consultazione del corpus.

CV Elena Artale

Dopo aver conseguito la laurea in Lettere Moderne all'Università di Pisa (tesi in letteratura italiana «Dal “fiorin d’oro d'amor” al “maladetto fiore”», voto 110/110 e lode), ha frequentato il corso di dottorato all'Università di Milano (corso in Storia della Lingua e della Letteratura Italiana, XI ciclo, tesi «Sulla lingua dei "libri di ricordi" del XV secolo»).Dal 1995 al 2003 ha svolto attività di ricerca presso l'Istituto CNR "Opera del Vocabolario Italiano" (OVI); dal 2001 al 2004 è stata professore a contratto presso l'Università di Pisa con l'incarico di insegnamento del corso ufficiale di "Laboratorio di scrittura italiana"; dal 2003 al 2005 è stata assegnista di ricerca per lavorare al corpus "Archivio Datini"; nel 2004-2005 ha collaborato con la casa editrice Le Monnier con l'incarico di datazione delle voci del Dizionario della lingua Italiana G. Devoto-G.C. Oli.Dal 2005 è ricercatrice all'OVI (di ruolo dal 2010). All'OVI è responsabile della lemmatizzazione del corpus del "Tesoro della Lingua Italiana delle Origini", "corpus TLIO", e codirige con Pär Larson il corpus testuale degli antichi testi italiani "corpus OVI dell'Italiano antico" [ISSN 2280-6857]. È anche redattrice di voci del TLIO, con particolare attenzione al lessico degli ambiti botanico e medico, e cura la "Bibliografia dei volgarizzamenti del corpus TLIO" (online nel sito del vocabolario, all'indirizzo http://tlio.ovi.cnr.it/BibVolg/).

Aree di interesse:

I suoi campi d'indagine riguardano: 1) Lessicografia storica italiana, in particolare del periodo medievale. 2) Storia della lingua italiana, con particolare attenzione agli scritti dei mercanti nei secc. XIV/XV.3) Letteratura italiana medievale e filologia, con particolare attenzione ai volgarizzamenti e ai testi di ambito medico medievali.4) Gestione e lemmatizzazione di banche dati testuali.