



Resumo TFG

Identificación de TFG (para o profesorado):

VTZNMKT

Apelidos e Nome do/a Titor/a

González Rufino Encarnación

Correo electrónico do/a Titor/a

nrufino@uvigo.es

Apelidos e Nome do CoTitor/a (se procede)

Javier Perez Guerra

Apelidos e Nome do Alumno/a

Gómez Vilar Santiago

DNI do Alumno/a

Título do TFG

DAT: Detección y anotación computacional de 'temas' sistémico-funcionales en corpus de lengua inglesa

Resumo

Como contribución a un proyecto sobre conexión entre género textual y realización temática en inglés escrito desde la lingüística empírica, este estudio describe el procedimiento computacional empleado en la detección automática y anotación XML de 'temas' en *AmE06* (Baker 2009), un corpus multi-género, analizado sintácticamente, de un millón de palabras de inglés americano contemporáneo (2006). 'Tema' es un concepto fundamental de la Lingüística Sistémico-Funcional (Halliday y Matthiessen 2004) que identifica, como factor de diseño oracional, el segmento inicial de cada oración.

El proceso computacional mencionado hace referencia al análisis, diseño e implementación de un servicio que, partiendo de una entrada de texto plano en inglés y mediante un conjunto de algoritmos a desarrollar, identifique los 'Temas' en dicho texto, los procese y genere una salida estructurada empleando el lenguaje de marcado XML.

Las tecnologías a emplear son Java para el desarrollo y Maven como herramienta de construcción del proyecto y gestor de dependencias. Debido a la alta incertidumbre en torno al proyecto, se debe analizar la posibilidad del empleo de distintas herramientas o librerías como JAXB, Project Lombok, JUnit, etc.

Respecto al proceso de desarrollo se empleará una adaptación de Scrum como marco de trabajo. La flexibilidad de este modelo de proceso es ideal en desarrollos con una incertidumbre en cuanto a requisitos técnicos como la ya mencionada. Esta herramienta permite dividir el desarrollo en diferentes 'Sprint' que, mediante la entrega constante de valor al cliente, acompañada de su retroalimentación, permite la adaptación ante el cambio.

Se formará un equipo de trabajo multidisciplinar con roles fuertemente marcados:

- **El alumno:** realizará análisis, diseño e implementación del proceso computacional ya citado junto con el rol de *Scrum Master*.
- **El cotutor:** encarnará el rol de *Product Owner* a la vez que ejerce de asesor y experto en lingüística.

Referencias bibliográficas:

Baker, Paul. 2009. The BE06 Corpus of British English and recent language change. *International Journal of Corpus Linguistics* 14/3: 312–337.

Halliday M.A.K. y C.M.I.M.
Matthiessen. 2004. *An introduction to
Functional Grammar*. London: Edward Arnold (3ª ed).

(A empresa debe ter convenio asinado en vigor coa Universidade de Vigo. **Deberase entregar copia do nomeamento do/a titor/a pola empresa**)

*** O/A profesor/a recibirá copia desta solicitude, e deberá dar o Visto e Prace do mesmo dende o formulario online dispoñible na web da ESEI.**

Código do TFG (para o alumno): EI 19/20-83

**Introduce un correo electrónico válido
(a continuación recibirá un código
para a firma da solicitude)**

*** Unha vez enviada a solicitude recibirá por correo electrónico copia da mesma**