

Indústries de la Llengua (21540)

Titulació/estudi: Grau en Llengües Aplicades

Curs: tercer o quart

Trimestre: tercer

Nombre de crèdits ECTS: 4

Hores de dedicació de l'estudiant: 100

Tipus d'assignatura: Optativa

Professor GG: Maarten Janssen

Llengua de docència: català, castellà i anglès

1. Presentació de l'assignatura

Aquest curs dona una introducció pràctica a diverses tècniques i aplicacions de la lingüística computacional que tenen rellevància directa en l'àmbit professional de lingüistes i traductors.

2. Competències que s'han d'assolir

- Anàlisi de situacions i resolució de problemes (G.2): utilització de eines informàtiques per resoldre problemes lingüístics.
- Gestió de la informació (G.5): manipulació i tractament de informació amb eines informàtiques.
- Coneixements d'informàtica (G.6): coneixement pràctic de diverses tècniques informàtiques, fins i tot HTML, etiquetatge morfosintàctic, expressions regulars, i traducció automàtic.
- Aplicació de coneixements a la pràctica (G.17): utilització en classe del coneixement d'informàtica per situacions reals i una tasca final per resoldre un problema amb eines informàtiques.
- Capacitat per poder reflexionar sobre el funcionament de la llengua (E.9): exploració de la llengua a través de eines i solucions informàtiques, fins i tot la ambigüitat estructural i lèxic.
- Utilització d'eines informàtiques aplicades a la mediació lingüística (E.11), inclòs el domini de tècniques de traducció assistida i localització.

3. Continguts

- Conceptes fonamentals de la lingüística computacional: etiquetes morfosintàctiques, tractament estadístic de textos, arbres sintàctics
- Tractament de textos mitjançant *scripts*: generació multilingüe de documents i altres exemples basats en casos reals per minimitzar la intervenció dels serveis lingüístics o per facilitar l'aplicació de memòria de traducció.
- Gestió multilingüe d'un servidor WEB: traducció HTML, configuració del servidor, arquitectura d'enllaços relatius vs tipificació d'infixos en les extensions dels fitxers.
- Configuració amb *software* públic d'entorns ofimàtics plurilingües i plurialfabètics.
- Traducció automàtica.
- Programació orientada a l'enginyeria lingüística: REGEX, Perl

4. Avaluació i recuperació

Avaluació		Recuperació			
Activitat d'avaluació	Ponderació sobre la nota final	Recuperable/ No recuperable	Ponderació sobre la nota final	Forma de recuperació	Requisits i observacions
Assistència i participació en les aules	10%	No recuperable			
Treball en grup*	20%	Recuperable	20%	Treball individual 1	
Treball individual	70%	Recuperable	70%	Treball individual 2	

* Durant els seminaris, n'hi ha quatre:

- i. Etiquetatge morfosintàctic
- ii. Expressions regulars
- iii. Montatge d'un servidor multilingüe
- iv. Traducció automàtica

5. Metodologia: activitats formatives

El curs s'organitza en una classe de grup gran i una classe de seminari cada setmana: a les classes de grup gran es proporcionen les bases teòriques i als seminaris els estudiants tenen l'oportunitat de treballar en les eines i mètodes exposats a les classes teòriques.

Classes magistrals: 15%

Seminaris: 10%

Tutories presencials: 5%

Treball individual: 70%

6. Bibliografia bàsica de l'assignatura

Hammond, Michael. (2003). *Programming for linguists: Perl for language researchers*. Blackwell.

L'Homme, M.C. 2008. *Initiation à la traductique*, Brossard (Québec)

Oliver, Antoni; Moré, Joaquim; Climent, Salvador (coord.) (2007). *Traducció i tecnologies*. Barcelona: Editorial UOC.

Manning, Christopher; Schütze, Hinrich. (1998). *Foundations of Statistical Natural Language Processing*. Cambridge: MIT Press.

Weisser, Martin. (2010). *Essential Programming for Linguists*. Edinburgh: Edinburgh University Press.