



FIȘA DISCIPLINEI

1. Date despre program

1.1 Instituția de învățământ superior	Universitatea “Alexandru Ioan Cuza” din Iași
1.2 Facultatea	Facultatea de Informatică
1.3 Departamentul	Informatică
1.4 Domeniul de studii	Informatică
1.5 Ciclul de studii	Master
1.6 Programul de studii / Calificarea	Informatică/Master în Lingvistică Computațională

2. Date despre disciplină

2.1 Denumirea disciplinei	Traducere automată						
2.2 Titularul activităților de curs	Conf. Dr. Diana TRANDABĂȚ						
2.3 Titularul activităților de seminar	Conf. Dr. Diana TRANDABĂȚ						
2.4 An de studiu	II	2.5 Semestru	1	2.6 Tip de evaluare	E	2.7 Regimul disciplinei*	OB

* OB – Obligatoriu / OP – Opțional

3. Timpul total estimat (ore pe semestru și activități didactice)

3.1 Număr de ore pe săptămână	4	din care: 3.2 curs	2	3.3 seminar/laborator	2
3.4 Total ore din planul de învățământ	56	din care: 3.5 curs	28	3.6 seminar/laborator	28
Distribuția fondului de timp					ore
Studiu după manual, suport de curs, bibliografie și altele					28
Documentare suplimentară în bibliotecă, pe platformele electronice de specialitate și pe teren					28
Pregătire seminarii/laboratoare, teme, referate, portofolii și eseuri					48
Tutoriat					-
Examinări					4
Alte activități					-
3.7 Total ore studiu individual					104
3.8 Total ore pe semestru					164
3.9 Număr de credite					8

4. Precondiții (dacă este cazul)

4.1 De curriculum	Nu este cazul
4.2 De competențe	Nu este cazul

5. Condiții (dacă este cazul)

5.1 De desfășurare a cursului	- Studenții se vor prezenta la prelegeri, seminarii/laboratoare cu telefoanele mobile închise.
5.2 De desfășurare a seminarului/ laboratorului	- Termenul predării temelor de laborator este stabilit de titularul de laborator, de comun acord cu studenții. - Nu se acceptă cererile de amânare decât pe motive obiectiv întemeiate. - Pentru predarea cu întârziere a temelor de laborator, acestea vor fi depunctate cu 0,5 pct./zi de întârziere.



6. Competențe specifice acumulate

Competențe profesionale	C1. Înțelegerea conceptelor de bază în traducere automată; C2. Utilizarea adecvată a instrumentelor de traducere automată existente; C3. Abilitatea de a dezvolta aplicații de traducere automată; C4. Înțelegerea tehnicilor și modelelor folosite pentru crearea de modele de limbă, modele de traducere, algoritmi de decodificare.
Competențe transversale	CT1. Identificarea rolurilor și responsabilităților într-o echipă pluridisciplinară și folosirea de tehnici de comunicare, relaționare și muncă eficientă în cadrul unei echipe. CT2. Aplicarea normelor și valorilor de etică profesională CT3. Participarea la proiecte cu caracter științific în condiții de autonomie și independență profesională și demonstrarea capacității de a identifica oportunități pentru propria formare profesională în viitor. CT4. Utilizarea eficientă a surselor informaționale și a resurselor de comunicare și formare profesională asistată, atât în limba română, cât și într-o limbă de circulație internațională.

7. Obiectivele disciplinei (din grila competențelor specifice acumulate)

7.1 Obiectivul general	Prezentarea noțiunilor fundamentale ale domeniului traducere automată, atât din perspectivă teoretică, cât și din punct de vedere al aplicațiilor.
7.2 Obiectivele specifice	La finalizarea cu succes a acestei discipline, studenții vor fi capabili să: <ul style="list-style-type: none">▪ recunoască modulele de bază ale traducerii automate statistice;▪ creeze modele de limbă;▪ înțeleagă diferența între diferite tipuri de modele de traducere (la nivel de cuvânt, de sintaxă, bazate pe exemple, etc.);▪ utilizeze corect sintaxa și termenii specifici domeniului traducerii automate;▪ dezvolte aplicații de traducere automată pentru diferite limbi.

8. Conținut

8.1	Curs	Metode de predare	Observații (ore și referințe bibliografice)
1.	Cap.I. Noțiuni generale despre desfășurarea cursului. Introducere în traducerea automată: istoric, model de limbaj, model de traducere, algoritm de decodare, evaluare	Expunere, prelegere participativă	2
2.	Cap. II. Traducere bazată pe reguli. Traducere bazată pe exemple. Traducere lexicală.	Expunere, prelegere participativă, exemplificare, aplicații	2
3.	Cap. III. Resurse: Corpusuri paralele. Corpusuri comparabile. Aliniere la nivel de cuvânt.	Expunere, prelegere participativă, exemplificare, aplicații	2
4-7.	Capitolul IV. Traducere bazată pe statistică. - Modele de limbă: N-gram, maximum-entropy, evaluare	Expunere, prelegere participativă, exemplificare, aplicații	8



	- Modele de traducere: Modele bazate pe cuvinte (word-based models). - Modele bazate pe grupuri (phrase-based models). - Modele bazate pe sintaxă.		
8.	Recapitularea și evaluarea noțiunilor acumulate.	Discuții și dezbateri, aplicații, testare.	2
9.- 10.	- Modelele IBM 1-5. Algoritmul Expectation Maximization. - Modele de reordonare	Expunere, prelegere participativă, exemplificare, aplicații	4
11- 12.	Capitolul V. Algoritmi de decodare.	Expunere, prelegere participativă, exemplificare, aplicații	4
13.	Capitolul VI. Metrice de evaluare: BLEU, METEOR	Expunere, prelegere participativă, exemplificare, aplicații	2
14.	Capitolul VII. Traducere asistată de calculator. Memorii de traducere	Expunere, prelegere participativă, exemplificare, aplicații	2

Bibliografie

- Philipp Koehn: Statistical machine translation. Cambridge University Press. xii, 433pp, 2009
- Yorick Wilks: Machine translation: its scope and limits. New York: Springer. x, 252pp, 2009
- Daniel Jurafsky and James H. Martin. 2008. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition. Second Edition. Prentice Hall.
- Machine Translation Archive <http://www.mt-archive.info>

Seminar/Laborator

Laboratoarele urmează îndeaproape noțiunile predate la curs. Printre obiectele de studiu se numără familiarizarea cu paradigmele majore din domeniul traducerii automate: traducere bazată pe cuvinte, traducerea bazată pe grupuri și traducerea bazată pe sintaxă. Studenții vor căpăta experiență în construirea de sisteme de traducere automată, lucrând cu date reale, și vor învăța cum să formuleze și să investigheze diferite direcții de cercetare. Cursurile și laboratoarele vor varia de la prelegeri asistate de videoproiector, la prezentări și discuții, studii de caz, programare, experimente cu diferite instrumente existente, iar în a doua jumătate a semestrului lucru la proiecte.

Bibliografie

- Philipp Koehn: Statistical machine translation. Cambridge University Press. xii, 433pp, 2009
- Yorick Wilks: Machine translation: its scope and limits. New York: Springer. x, 252pp, 2009
- Daniel Jurafsky and James H. Martin. 2008. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition. Second Edition. Prentice Hall.
- Machine Translation Archive <http://www.mt-archive.info>
- Jascha Hoffman Me Translate Funny One Day NY Times article, Oct. 19, 2012.

9. Coroborarea conținutului disciplinei cu așteptările reprezentanților comunității, asociațiilor profesionale și angajatorilor reprezentativi din domeniul aferent programului



Conținutul cursului este coroborat cu conținutul cursurilor similare de la diferite universități internaționale de prestigiu. În contextul big data și a multilingvismului, înțelegerea și dezvoltarea de aplicații de traducere automată sunt esențiale pentru comunicarea interumană, precum și pentru extragerea de informații din colecții imense de date, structurate sau nestructurate.

10. Evaluare

Tip activitate	10.1 Criterii de evaluare	10.2 Metode de evaluare	10.3 Pondere în nota finală (%)
10.4 Curs	<ul style="list-style-type: none"> - Corectitudinea și completitudinea cunoștințelor: - Capacitatea de a înțelege și explica modelele de traducere și de limbă. - Înțelegerea modului de decodare pe baza modelelor statistice. - Calitatea formulării răspunsurilor. 	Teste practice (evaluare parțială și evaluare finală), proiecte	30%
10.5 Seminar/ Laborator	<ul style="list-style-type: none"> - Abilitatea de a construi o aplicație de traducere automată care rezolvă probleme de procesarea limbajului. - Folosirea corectă a sintaxei și termenilor specifici. - Însușirea deprinderilor de lucru în echipă și contribuire susținută la diversele etape din dezvoltarea unui proiect. 	Exerciții de laborator, teme	30%
	<ul style="list-style-type: none"> - Însușirea deprinderilor de lucru în echipă și contribuire susținută la diversele etape din dezvoltarea unui proiect. 	Proiect	30%
	<ul style="list-style-type: none"> - Criterii ce vizează aspectele atitudinale: - Calitatea programelor/ modelelor elaborate - Interesul pentru studiu individual 	Participarea activă la laboratoare	10%

10.6 Standard minim de performanță

Pentru promovare trebuie îndeplinite simultan următoarele criterii:

- Utilizarea corectă a sintaxei specifice aplicațiilor de traducere automată;
- Abilitatea de a explica modul de funcționare a modelelor de limbă, de traducere, de reodonare;
- Capacitatea de a implementa o aplicație de traducere automată pentru diverse limbi;
- Minim 50% prezență la laborator;
- Minim 15 puncte din temele de la laborator;
- Minim 15 puncte suma punctajelor testelor de la curs
- Minim 50 puncte în total: (prezențe + teste + teme + bonus + proiect)

Notele se stabilesc conform cu criteriile ECTS

Data completării
20 martie 2018

Titular de curs
Conf.Dr. Diana Trandabăț

Titular de seminar
conf.Dr. Diana Trandabăț

Data avizării în departament

Director de departament
Prof.Dr. Dorel LUCANU