



Universitatea Națională de Știință și Tehnologie
POLITEHNICA București
Facultatea de TEOLOGIE, LITERE, ISTORIE ȘI ARTE



FIȘA DISCIPLINEI

1. Date despre program

1.1 Instituția de învățământ superior	Universitatea Națională de Știință și Tehnologie POLITEHNICA din București
1.2 Facultatea	Teologie, Litere, Istorie și Arte
1.3 Departamentul	Departamentul de Limbi Străine Aplicate
1.4 Domeniul de studii universitare	Filologie Inginerie electronică, telecomunicații și tehnologii informaționale
1.5 Programul de studii universitare	<i>Plurilingvism și inteligență artificială</i>
1.6 Ciclu de studii universitare	Master
1.7 Limba de predare	Română
1.8 Locația geografică de desfășurare a studiilor	Pitești

2. Date despre disciplină

2.1 Denumirea disciplinei/	<i>Introducere in lingvistica corpusului. Instrumente de manipulare a corpusurilor și extracție terminologică</i>						
2.2 Titularul/ii activităților de curs	Conf.univ.dr. Ilinca Cristina						
2.3 Titularul/ii activităților de seminar / laborator/proiect							
2.4 Anul de studiu	1	2.5 Semestrul	I	2.6. Tipul de evaluare	E	2.7 Statutul disciplinei	Ob ¹
2.8 Categoria formativă ²	DS	2.9 Codul disciplinei	P.M.23.F.11.I.Ob.03.				

3. Timpul total (ore pe semestru al activităților didactice)

3.1 Număr de ore pe săptămână	2	Din care: 3.2 curs	2	3.3 seminar/laborator/proiect	
3.4 Total ore din planul de învățământ	28	Din care: 3.5 curs	28	3.6 seminar/laborator/proiect	
Distribuția fondului de timp:					ore
Studiul după manual, suport de curs, bibliografie și notițe Documentare suplimentară în bibliotecă, pe platformele electronice de specialitate/Pregătire seminarii/ laboratoare/proiecte, teme, referate, portofolii și eseuri					110
Tutorat					10
Examinări					2
Alte activități (dacă există):					
3.7 Total ore studiu individual	122				
3.8 Total ore pe semestru	150³				
3.9 Numărul de credite	6⁴				

4. Precondiții (acolo unde este cazul)

¹ Obligatorie / Opțională / Facultativă – Se va completa conform planului de învățământ.

² Fundamentală / de specializare/ complementare – Se va completa conform planului de învățământ.

³ Se va calcula ținând cont că se acordă un credit pentru volumul de muncă care îi revine unui student cu frecvență la zi pentru a echivala 25 de ore de pregătire pentru dobândirea rezultatelor învățării.

⁴ Se va completa conform planului de învățământ.



Universitatea Națională de Știință și Tehnologie
POLITEHNICA București
Facultatea de TEOLOGIE, LITERE, ISTORIE ȘI ARTE



4.1 de curriculum	Parcurea și/sau promovarea următoarei discipline: <i>Lingvistică generală</i>
4.2 de rezultate ale învățării	Cunoștințe TIC minimale

5. Condiții necesare pentru desfășurarea optimă a activităților didactice (acolo unde este cazul)

5.1 de desfășurare a cursului	Cursul se va desfășura într-o sală dotată cu videoproiector și computere.
5.2 de desfășurare a seminarului/laboratorului/proiectului	-

6. Obiectiv general

Această disciplină se studiază în cadrul domeniului *Filologie*, programul de studii *Plurilingvism și inteligență artificială* și are următorul obiectiv **general** :

- Dezvoltarea de competențe care să permită analiza unei cantități mari de date lingvistice (corpusuri) cu ajutorul noilor tehnologii informatice.
- Prezentarea lingvisticii corpusului ca metodă de cercetare în domenii precum studiul limbilor, analiză de discurs, traducere terminografie, didactică.

Obiective specifice:

- Dezvoltarea abilității de creare și exploatare a unui corpus în funcție de obiectivele de cercetare lingvistică;
- Dezvoltarea deprinderilor de a utiliza instrumente de manipulare a corpusurilor pentru extragerea și analiza datelor lingvistice, în special pentru activități de traducere și terminografie;
- Dezvoltarea abilităților de cercetare interdisciplinară.

7. Rezultatele învățării

Competențele asociate disciplinei *Introducere în lingvistica corpusului. Instrumente de manipulare a corpusurilor și extracție terminologică* sunt corelate cu următoarele rezultate ale învățării.

CP5 Dezvoltă baze de date terminologice (2 ECTS)

CP8 Respectă standarde de calitate în traducere (2 ECTS)

CP9 Aplică terminologia TIC (2ECTS)

Cunoștințe	C5. Studentul/absolventul descrie și clasifică termeni, prezintă etimologia și utilizarea acestora. Identifică semnificația cuvintelor în funcție de contextul în care sunt utilizate, originea cuvântului și evoluția acestuia în timp. C8. Studentul/absolventul explică standardele convenite în domeniul serviciilor lingvistice. C9. Studentul/absolventul identifică limbajul utilizat în activitatea de documentare și comunicare care conține termeni specifici TIC.
------------	---



Universitatea Națională de Știință și Tehnologie
POLITEHNICA București
Facultatea de TEOLOGIE, LITERE, ISTORIE ȘI ARTE



Abilități	<p>A5.Studentul/absolventul colectează și prezintă termeni, după verificarea legitimității lor, pentru a crea baze de date terminologice pe o serie de domenii.</p> <p>A8.Studentul/absolventul respectă standardele convenite, cum ar fi standardul european EN 15038 și ISO 17100, pentru a asigura respectarea cerințelor legate de furnizorii de servicii lingvistice și pentru a garanta uniformitatea.</p> <p>A9.Studentul/absolventul utilizează termeni specifici și vocabular TIC într-un mod sistematic și coerent, pentru documentare și comunicare</p>
Responsabilitate și autonomie	<p>RA5.Studentul/absolventul demonstrează autonomie și responsabilitate în identificarea, validarea și organizarea termenilor specifici diverselor domenii, contribuind la elaborarea și actualizarea bazelor de date terminologice.</p> <p>RA8.Studentul/absolventul aplică riguros standardele recunoscute (precum EN 15038 și ISO 17100), contribuind la asigurarea calității, conformității și uniformității serviciilor lingvistice.</p> <p>RA9.Studentul/absolventul folosește în mod consecvent și structurat termeni specifici și vocabular TIC, asigurând claritatea și precizia în activitățile de documentare și comunicare profesională.</p>

8. Metode de predare

Procesul didactic va valorifica o varietate de metode: atât metode expositive (precum prelegerea și expunerea), cât și metode conversative și interactive, care încurajează învățarea prin descoperire. Acestea vor include explorarea directă și indirectă a instrumentelor de manipulare a corpusurilor, precum și metode centrate pe acțiune, cum ar fi exercițiile, activitățile practice și gestiunea de situații complexe în autonomie și în grup. Activitatea de predare va include prelegeri susținute cu ajutorul prezentărilor PowerPoint și al materialelor video, instrucțiuni de folosire aplicații informatice specifice domeniului care vor fi puse la dispoziția studenților. Fiecare curs va începe cu o scurtă recapitulare a capitolelor anterioare, punându-se accent pe conceptele discutate în sesiunea precedentă.

9. Conținuturi

CURS		
Capitolul	Conținutul	Nr. ore
I	Introducere în lingvistica corpusului. Accepțiunile termenului „corpus”. Perspectivă diacronică asupra lingvisticii corpusului. Lingvistica corpusului ca metodă de lucru în diverse ramuri ale studiului limbilor.	2
II	Corpus, corpusuri. Clasificarea corpusurilor după diverse criterii. Corpusuri importante existente pentru diverse limbi și caracteristicile lor. Utilizarea corpusurilor în limbaje specializate, traductologie, terminografie, didactica limbilor străine, lingvistică juridică.	2
III	Proiectarea corpusurilor și colectarea de date. Prezentarea de proiecte de corpusuri pentru diverse limbi (ex. British National Corpus, CEFC, CoRoLa). Stabilire de obiective de cercetare și criterii de colectare date. Probleme etice și legale (drepturi de autor).	2
IV	Statistică și vizualizarea datelor lingvistice. Metode statistice (analiză cluster, analiză factorială, testul de influență T, analiza variației-ANOVA, testul chi-pătrat, analiza de regresie) folosite în analiza datelor lingvistice și corpusurilor. Exemple de aplicare a acestor metode.	2



V	Instrumente de manipulare a corpusurilor. Definirea conceptelor de lucru: frecvență, concordanță, cologații, coligații, N-gram. Prezentarea instrumentelor existente. AntConc, Sketch Engine, Voyant Tools – instalare, utilizare	4
VI	Adnotarea corpusurilor. Metadate. Etape de preprocesare a corpusurilor. Adnotarea morfologică. Adnotare sintactică (de suprafață și de adâncime). Alte adnotări	4
VII	Explorarea cantitativă a corpusurilor. Lista de cuvinte. Frecvență. Dispersie. Cuvinte-cheie; cuvinte-cheie negative	4
VIII	Explorarea lexicală a corpusurilor. Concordanțe. Cologații. Coligații	4
IX	Corpusuri în analiza discursului specializat. Extracție terminologică. Selectarea termenilor-candidați și constituirea de baze de date terminologice.	4
Total:		28

Bibliografie:

Ilinca, Cristina. 2021. *Introducere în lingvistica corpusului. Instrumente de manipulare a corpusurilor și de extragere terminologică*. (note de curs), suport electronic

Barbu Mititelu, Verginica, Irinia, Elena, Tufiş, Dan. 2014. CoRoLa – The Reference Corpus of Contemporary Romanian Language. Proceedings of LREC2014.

Barbu Mititelu, Verginica and Irinia, Elena and Păiş, Vasile and Mitrofan, Maria and Avram, Andrei-Marius and Curea, Eric. 2021. "Linked open data resources for Romanian". *The 16th International Conference on Linguistic Resources and Tools for Natural Language Processing*. pp. 7—19.

Burnard, Lou. *Reference Guide for the British National Corpus*. <http://www.natcorp.ox.ac.uk/docs/URG/>

Frérot, C., & Pecman, M. (éds.). (2021). *Des corpus numériques à l'analyse linguistique en langues de spécialité* (1-). UGA Éditions. <https://doi.org/10.4000/books.ugaeditions.24195>

Ion, Radu and Barbu Mititelu, Verginica and Păiş, Vasile and Irinia, Elena and Badea, Valentin. 2024. "A Cross-model Study on Learning Romanian Parts of Speech with Transformer Models". *Proceedings of the Sixth International Conference on Computational Linguistics in Bulgaria (CLIB 2024)*. pp. 6—13.

Ivanović, Tanja, Stanković, Ranka, Todorović, Branislava Šandrih, Krstev, Cvetana. 2022. "Corpus-based bilingual terminology extraction in the power engineering domain". *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, Volume 28, Issue 2, p. 228 – 263.

Lindquist, Hans, Levin, Magnus, 2018. *Corpus Linguistics and the Description of English*. Edinburgh University Press.

McEnery, Anthony, Hardie, Andrew. 2012. *Corpus Linguistics. Method, Theory and Practice*. Cambridge: Cambridge University Press.

Weisser, Martin. 2016. *Practical Corpus Linguistics: An Introduction to Corpus-Based Language Analysis*. Wiley-Blackwell;

Xiao, Richard, Wei Naixing . 2014. "Translation and contrastive linguistic studies at the interface of English and Chinese". *Corpus linguistics and linguistic theory*, 10, 1 – 10.

10. Evaluare

Tip activitate	10.1 Criterii de evaluare	10.2 Metode de evaluare	10.3 Pondere din nota finală
10.4 Curs	- capacitatea de asimilare și aplicare a cunoștințelor învățate; - gradul de asimilare a limbajului de specialitate și capacitatea de comunicare pe temele cursului; - capacitatea de a opera cu cunoștințele asimilate în activități intelectuale complexe.	Evaluare formativă	
		- Referat	30%
		-Teme	20%
		Evaluare sumativă (la încheierea perioadei de studiu al disciplinei de învățământ) – probă orală	50%



Universitatea Națională de Știință și Tehnologie
POLITEHNICA București
Facultatea de TEOLOGIE, LITERE, ISTORIE ȘI ARTE



10.5 Seminar/laborator/proiect			
10.6 Condiții de promovare			
Exemplu: <ul style="list-style-type: none">• Obținerea a 50% din punctajul total.• Obținerea a 50% din punctajul aferent activității pe parcursul semestrului:• Cunoașterea avantajelor și a limitărilor lingvisticii corpusului, a tipurilor de corpusuri, a etapelor prelucrării acestora și a modalităților de explorare a lor.• Abilitatea de a utiliza un instrument de manipulare a corpusurilor pentru a regăsi diverse informații.			

Data completării	Titular de curs	Titular(ii) de aplicații
23.09.2025	Conf.univ.dr. Elena-Cristina Ilinca	
Data avizării în departament	Director de departament	
24.09.2025	Conf.univ.dr.Laura Cițu	
Data aprobării în Consiliul Facultății	Decan	
29.09.2025	Conf.univ.dr. Constantin Augustus Bărbulescu	