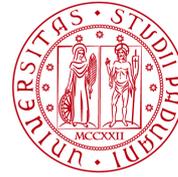




DEI
DIPARTIMENTO DI
INGEGNERIA DELL'INFORMAZIONE



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

via Gradenigo, 6/B
35131 Padova, Italy
tel +39 049 8277600
fax +39 049 8277699
info@dei.unipd.it
www.dei.unipd.it

CF 80006480281
P.IVA 00742430283

PROGETTO PER LA COSTITUZIONE DEL

Centro Studi in Terminologia Computazionale (CENTRICO)

MANIFESTO D'INTENTI

1. Tipologia di progetto: Centro Studi
2. Titolo della proposta: **Centro studi di Terminologia Computazionale (CENTRICO)**
3. Proponenti:
 - Giorgio Maria Di Nunzio (**Principal Investigator**), DEI, Prof. Associato
 - Federica Vezzani, DiSLL, RTDa PON
4. Sede: Dipartimento di Ingegneria dell'Informazione
5. Durata del progetto: 4 anni
6. Settori Scientifico Disciplinari coinvolti: ING-INF/05, INF/01, L-LIN/04, L-LIN/07, L-LIN/09, L-LIN/12, L-LIN/14, L-LIN/01, L-FIL-LET/12, L-FIL-LET/14
7. Oggetto del progetto:

Il centro studi in terminologia computazionale si propone di consolidare le attività di terza missione e di ricerca nate dalla collaborazione i) all'interno dell'Università di Padova tra i Dipartimenti di Ingegneria dell'Informazione e il Dipartimento di Studi Linguistici e Letterari, e ii) all'esterno con l'Unità di coordinamento della terminologia del Parlamento Europeo (TermCoord)¹.

In particolare, i lavori e le iniziative proposte dal centro studi ruoteranno attorno all'applicazione Web FAIRterm² progettata e sviluppata a partire dal 2020. Lo strumento nasce dalla collaborazione dei due proponenti del centro studi e diventerà, a partire dal secondo semestre dell'A.A. 2022, l'applicazione Web ufficiale per la raccolta di dati

¹ <https://termcoord.eu>

² <http://purl.org/fairterm>

terminologici multilingue dell'iniziativa internazionale "Terminology Without Borders" (YourTerm)³ promossa dal TermCoord.

Il centro studi offrirà la possibilità di integrare le competenze multidisciplinari al fine di:

- 1) Stimolare e sostenere attività di terza missione e ricerca che includono la divulgazione della conoscenza di domini specializzati;
- 2) Favorire l'integrazione e l'accesso alle competenze e alle risorse digitali dell'Università di Padova;
- 3) Incentivare la collaborazione e trasferimento di conoscenze tra Università e territorio.

OBIETTIVI DEL CENTRO STUDI

- 1) Consolidare, ampliare e formalizzare la collaborazione con il TermCoord.
- 2) Favorire le attività di terza missione e ricerca che riguardano l'accesso della conoscenza specializzata multilingue per il grande pubblico.
- 3) Promuovere attività di ricerca sulla modellazione di dati linguistici standardizzata secondo le direttive degli ISO pertinenti e su modelli teorici avanzati della terminologia e lessicologia specializzata multilingue.

CONTESTO EUROPEO DEL CENTRO STUDI

Il Parlamento Europeo produce traduzioni di documenti scritti in tutte le 24 lingue dell'UE, consentendo la comunicazione con i cittadini nella loro lingua madre. Il suo servizio di traduzione interno è uno dei più grandi servizi multilingue con circa 1.100 persone tra traduttori, terminologi, assistenti e altro personale di supporto che consentono una comunicazione multilingue al Parlamento europeo rispettando gli elevati standard di qualità imposti dalle procedure parlamentari.

³ <https://yourterm.eu>

L'Unità di coordinamento della terminologia (TermCoord) è un'unità di supporto alla Direzione generale della Traduzione (DG TRAD)⁴ del Parlamento europeo. Il ruolo principale del TermCoord è quello di assistere i traduttori e facilitare la ricerca e la gestione terminologica nelle unità di traduzione, nonché di incrementare il contributo del Parlamento europeo alla banca dati terminologica dell'UE IATE (Interactive Terminology for Europe)⁵, la più grande risorsa terminologica multilingue disponibile attualmente a livello mondiale.

Tra le varie attività si annovera il progetto Terminology Without Borders, abbreviato con YourTerm, che riflette e sostiene l'obiettivo della DG TRAD di comunicare con i cittadini in un linguaggio chiaro. L'obiettivo principale di YourTerm è quello di migliorare la comunicazione tra gli esperti del dominio e il grande pubblico in diversi settori specializzati attraverso dei sotto-progetti che coinvolgono direttamente le Università europee ed extraeuropee. Ognuno di questi sotto-progetti produce risorse terminologiche multilingue di alta qualità grazie alla collaborazione e al coordinamento di numerosi utenti provenienti da diverse università e istituti di ricerca.

Dal 2018, i due proponenti del centro studi CENTRICO hanno lavorato insieme al Presidente del TermCoord, Rodolfo Maslias, e ai coordinatori di YourTerm per la progettazione e l'implementazione di un sistema collaborativo di gestione terminologica multilingue per supportare il lavoro dei partner in ogni sottoprogetto. FAIRterm è l'applicazione Web che è stata scelta nel 2022 dal TermCoord come strumento ufficiale per gestire la terminologia prodotta da YourTerm.

8. Costi del progetto:

Il Centro Studi deve garantire la funzionalità 24/7 del software attualmente sviluppato utilizzando la piattaforma Cloud shinyapps.io⁶ per l'esecuzione del codice e la gestione degli accessi degli utenti all'applicazione Web.

Il costo minimo previsto è l'account "basic" pari a 440\$ all'anno, per un totale di 1.760\$. Se il numero di partecipanti ai progetti YourTerm dovesse aumentare considerevolmente, si passerebbe ad un account "standard" pari a 1,100\$ annuali.

⁴ <https://www.europarl.europa.eu/translation/>

⁵ <https://iate.europa.eu/>

⁶ <https://www.shinyapps.io>

Inoltre, l'obiettivo è quello di progettare e sviluppare una nuova applicazione Web che possa essere portata sui server UNIPD per una gestione più efficace e per uno sviluppo modulare più avanzato.

9. Risorse per il finanziamento dell'unità di progetto:

Le iniziative del Centro Studi CENTRICO sono volte a potenziare le attività di terza missione e di ricerca secondo gli obiettivi indicati.

Le iniziative saranno sostenute economicamente secondo le modalità indicate nelle linee di indirizzo dell'Ateneo grazie a: assegnazioni libere dei ricercatori aderenti; assegnazioni da parte dei Dipartimenti, previa delibera del Consiglio che ne motivi l'interesse scientifico e didattico per il Dipartimento stesso; fondi derivanti da convenzioni e contratti con soggetti terzi che prevedano l'esplicita destinazione delle risorse alle attività del progetto.

Ulteriori fondi e risorse necessari allo svolgimento delle attività saranno reperiti e messi a disposizione, garantendo la tracciabilità delle entrate e delle spese destinate alla realizzazione del progetto.

10. Collegio dei ricercatori (Dipartimento, Nome, Posizione)

- DEI, Giorgio Maria Di Nunzio, PA
- DEI, Nicola Ferro, PO
- DEI, Gianmaria Silvello, PA
- DEI, Guglielmo Faggioli, PhD
- DEI, Stefano Marchesin, Post Doc
- DEI/DPCD, Fabio Giachelle, Post Doc
- DiSLL, Federica Vezzani, RTDa
- DiSLL, Geneviève Henrot, PO
- DiSLL, Carmen Castillo Pena, PA
- DiSLL, Maria Teresa Musacchio, PO
- DiSLL, Alessandro Metlica, PA
- DiSLL, Giorgia Gallucci, PhD
- DPCD, Stefania Gialdroni, PO
- DPCD, Davide Basaldella, Post Doc
- DPCD, Denise Bezzina, Post Doc

- DPCD, Davide De Concilio, Post Doc
- DPCD, Jake Dyble, Post Doc
- DPCD, Francesca Fusco, Post Doc

11. Advisory Board

- Rodolfo Maslias, Presidente dell'Unità di coordinamento della terminologia del parlamento europeo
- Tiziana Sicilia, Presidente dell'Associazione Italiana per la Comunicazione Tecnica (COM&TEC), Presidente della Associazione Europea per la Comunicazione Tecnica (tekomp Europe).

12. Pubblicazioni rilevanti

Vezzani F. (2022). Terminologie numérique: conception, représentation et gestion. Bern, Peter Lang, coll. "Linguistics insights". DOI: 10.3726/b19407 - ISBN: 9783034342643

Monaci L., e Vezzani F. (2022). "Pour une étude de la terminologie médicale de Proust : rétro-numérisation et analyse de la Correspondance avec sa mère". In L'ANALISI LINGUISTICA E LETTERARIA vol. 31 (2), pp. 127-140. ISBN: 978-88-9335-982-5 (Rivista di classe A).

Vezzani F. (in stampa 2022). "La gestion de (méta)données terminologiques « FAIR » : le répertoire de catégories de données de la ressource TriMED". In Terminologie & Ontologie : Théories et Applications Actes de la conférence TOTh 2021, Presses Universitaires Savoie Mont Blanc, Collection "Terminologica". ISBN : 978-2-919732- 80-7 ISSN : 2607-5008

Vezzani F. (2022). "La traduction médicale : un panorama de ressources terminologiques multilingues". Volume Approches linguistiques contemporaines de la traduction, Presses de l'Université d'Artois (coll. "Études Linguistiques").

DOI: 10.4000/books.apu.26754. ISBN: 9782848325323.

Vezzani F., e Di Nunzio G.M. (2022). "Elaborazione e gestione di (meta) dati terminologici". In Risorse e strumenti per l'elaborazione e la diffusione della terminologia, Eurac Research, ISBN: 978-88-98857-72- 2.

Piccini S., Vezzani F., e Bellandi A. (2022). "Entre TBX et Ontolex- Lemon : Quelles Nouvelles Perspectives en Terminologie ?" In CEUR Workshop Proceedings (vol. 3161) of the 1st Multilingual Digital Terminology Today Conference (MDTT 2022).

Silecchia S., Vezzani F., e Di Nunzio G.M. (2022). "Knowledge Representation and Language Simplification of Human Rights". In Proceedings of the LREC 2022 Workshop on Terminology in the 21st century: many faces, many places (Term21) - ISBN: 979-10-95546-95-5

Bonato V., Di Nunzio G.M., e Vezzani F. (2021). "Preliminary considerations on a systematic approach to semic analysis: the case study of medical terminology". Umanistica Digitale, (10), AlmaDL Journals, pp. 211-234. ISSN: 2532-8816, <http://doi.org/10.6092/issn.2532-8816/12621> (Rivista scientifica).

Vezzani, F. (2021). "La ressource FAIRterm: entre pratique pédagogique et professionnalisation en traduction spécialisée". Synergies Italie, (17), Groupe d'Études et de Recherches pour le Français Langue Internationale (GERFLINT) pp. 51-64. pp. 51-64. ISSN 1724-0700, ISSN en ligne 2260-8087 (Rivista scientifica).

Yeganova, L.; Wiemann, D.; Neves, M.; Vezzani, F.; et al. (2021). "Findings of the WMT 2021 Biomedical Translation Shared Task: Summaries of Animal Experiments as New Test Set". In Proceedings of the WMT 2021 - 6th Conference on Machine Translation, Association for Computational Linguistics (ACL), pp. 664-683.

Di Nunzio G.M., e Vezzani F. (2021). "One Size Fits All: A Conceptual Data Model for Any Approach to Terminology". In Workshop TOTH 2021. Terminology, Interoperability and Data integration: Issues and Challenges.

Di Nunzio G.M., e Vezzani F. (2021). "IMS-UNIPD@ CLEF eHealth Task 2: Reciprocal Ranking Fusion in CHS". In Proceedings of the Working Notes of CLEF 2021 - Conference and Labs of the Evaluation Forum, CEUR-WS, pp. 775-779. ISSN: 1613-0073

Di Nunzio G.M., e Vezzani F. (2021). "On the Reusability of Terminological Data." In the Proceedings of the 10th AIUCD Conference (Associazione per l'Informatica Umanistica e la Cultura Digitale), AlmaDL Journals, pp. 183-186. ISBN 9788894253559. DOI 10.6092/unibo/amsacta/6712

Vezzani F., e Di Nunzio G.M. (2020). "Methodology for the standardization of terminological resources: design of TriMED database to support multi-register medical communication". Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication, 26(2), John Benjamins Publishing Company, pp. 265-297. ISSN: 0929-9971, DOI: <https://doi.org/10.1075/term.00053.vez> (Rivista di classe A).

Di Nunzio G.M., Marchesin S., e Vezzani F. (2020). "A Study on Reciprocal Ranking Fusion in Consumer Health Search. IMS UniPD at CLEF eHealth 2020 Task 2". In Working Notes of CLEF 2020 - Conference and Labs of the Evaluation Forum. ISSN: 1613-0073

Di Nunzio G.M., Dosso D., Fabris A., Faggioli G., Ferro N., Giachelle F., Irrera O., Marchesin S., Piazzon L., Purpura A., Silvello G., e Vezzani F. (2020). "UNIPD at Covid-19 MLIA". In Covid-19 MLIA @ Eval Initiative.

Bawden R., Di Nunzio G.M., Grozea C., Unanue I., et al. (2020) "Findings of the WMT 2020 Biomedical Translation Shared Task: Basque, Italian and Russian as New Additional Languages." In the Proceedings of the 5th Conference on Machine Translation, Association for Computational Linguistics, pp. 660-687.

Vezzani F., e Di Nunzio G.M. (2020). "On the Formal Standardization of Terminology Resources: The Case Study of TriMED". In the Proceedings of the 12th Language Resources and Evaluation Conference (LREC). European Language Resources

Association, pp. 4903–4910. ISBN 979-10-95546-34-4, <https://aclanthology.org/2020.lrec-1.603/>

Vezzani, F., e Di Nunzio, G. M. (2019). “(Not so) Elementary, my dear Watson! A different perspective on medical terminology”. *Umanistica Digitale*, (6), AlmaDL Journals, pp. 59-75. ISSN: 2532-8816, DOI: 10.6092/issn.2532-8816/8632 (Rivista scientifica).

Vezzani F., e Di Nunzio G.M. (2019). “Computational Terminology in eHealth”. In: *Digital Libraries: Supporting Open Science - 15th Italian Research, Conference on Digital Libraries, IRCDL 2019, Pisa, Italy, January 31 - February 1, 2019, Proceedings*, vol. 988, Communications in Computer and Information Science, Springer, pp. 72-85, ISBN: 978- 3-030-11226-4, DOI: 10.1007/978-3-030-11226-4_6

Vezzani F. (2019). “La technicité des termes : le v-tech comme paramètre d'évaluation”. In *Terminologie & Ontologie : Théories et Applications Actes de la conférence TOTh 2019*, Presses Universitaires Savoie Mont Blanc, Collection « Terminologica », pp. 215-227. ISBN: 978-2-919732-80-7, ISSN: 2607-5008.

Vezzani F., (2019). “Aménagement de la terminologie spontanée: un cas de collocation”. *Convergences et divergences dans la pratique terminologique. De la terminologie spontanée à la terminologie aménagée*, Délégation générale à la langue française et aux langues de France, Ministère de la Culture, pp. 163-173. ISBN 978-2-11-139384-4

Vezzani F., e Di Nunzio G.M. (2019). “On the Use of Terminological Records in Specialised Translation”. In the *Proceedings of the 8th AIUCD Conference (Associazione per l'Informatica Umanistica e la Cultura Digitale)*, AlmaDL Journals, pp.151-156. ISBN 978- 88-942535-3-5. DOI 10.6092/unibo/amsacta/6361

Di Nunzio G.M., e Vezzani F. (2018). “Using R Markdown for Replicable Experiments in Evidence Based Medicine”. In the *Experimental IR Meets Multilinguality, Multimodality, and Interaction - 9th International Conference of the CLEF Association, CLEF 2018, Avignon, France, September 10-14,*

2018, Springer, pp. 28-39. ISBN: 978-3-319-98931-0. DOI:
10.1007/978-3319-98932-7

Di Nunzio G.M., Maistro M., e Vezzani F. (2018). "A Gamified Approach to Naïve Bayes Classification: A Case Study for Newswires and Systematic Medical Reviews". In the Companion of the The Web Conference 2018 on The Web Conference 2018, WWW 2018, Lyon, France, April 23-27, 2018, International World Wide Web Conferences Steering Committee, pp. 1139--1146. ISBN 9781450356404, DOI: 10.1145/3184558.3191547

Vezzani F., Di Nunzio G.M., e Henrot G. (2018). "TriMED: A Multilingual Terminological Database". In the Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC), European Language Resources Association, pp. 4367-4371. ISBN: 979-10-95546-00-9

Vezzani F., Di Nunzio G.M., e Henrot G. (2018). "TriMED: banca dati terminologica multilingue". In the Proceedings of the 7th AIUCD Conference (Associazione per l'Informatica Umanistica e le Culture Digitali), AlmaDL Journals, pp. 237-241. ISBN 978-88-942535-2-8. DOI 10.6092/unibo/amsacta/5997

Di Nunzio G.M., e Vezzani F. (2018). "A Linguistic Failure Analysis of Classification of Medical Publications: A Study on Stemming vs Lemmatization". In the Proceedings of the Fifth Italian Conference on Computational Linguistics (CLiC-it).

Di Nunzio G.M., Ciuffreda G., e Vezzani F. (2018). "Interactive Sampling for Systematic Reviews. IMS Unipd At CLEF 2018 eHealth Task 2". In the Working Notes of CLEF 2018 - Conference and Labs of the evaluation forum.

Di Nunzio G.M., e Vezzani F. (2018). "A Study on Manual Query Reformulation for Systematic Medical Reviews". In the Proceedings of the 9th Italian Information Retrieval Workshop (IIR).

Di Nunzio G.M., Beghini F., Vezzani F., e Henrot G. (2017). "A Reproducible Approach with R Markdown to Automatic Classification of Medical Certificates in French". In the Proceedings of the Fourth Italian Conference on Computational Linguistics (CLiC-it).

Di Nunzio G.M., Beghini F., Vezzani F., e Henrot G. (2017). "A Lexicon Based Approach to Classification of ICD10 Codes. IMS Unipd at CLEF eHealth Task 1". In the Working Notes of CLEF-2017 - Conference and Labs of the evaluation forum.

Di Nunzio G.M., Beghini F., Vezzani F., e Henrot G. (2017). "An Interactive Two-Dimensional Approach to Query Aspects Rewriting in Systematic Reviews. IMS Unipd At CLEF eHealth Task 2". In the Working Notes of CLEF2017 - Conference and Labs of the evaluation forum.