Antoni Oliver

✉ aoliverg@uoc.edu | 📱 +34 639 38 47 41 | ORCID: 0000-0001-8399-3770 | 🐙 aoliverg

PERSONAL AND PROFESSIONAL DATA

Name and surname: Antoni Oliver

ID card: 46232691T

Nationality: Spanish

Birth date: 23/05/1969

Home address: Carrer Sostres 7 1r 2a 08024 Barcelona

Current position: Associate professor (*professor agregat*) at the Arts and Humanities Department

Institution: Universitat Oberta de Catalunya (UOC)

SUMMARY

I have a **highly interdisciplinary profile** with a university education including both technical degrees (Telecommunications Engineering and Master in Free Software), as well as education in humanities (degree in Slavonic Philology and PhD in Linguistics). My research is also interdisciplinary, the main area being computational linguistics, and more specifically machine translation and the automatic creation of lexical computational resources, as well as the computer-assisted translation tools. Regarding machine translation, I'm particularly interested in two aspects: (a) the training and integration of neural machine translation for low-resource language pairs; and (b) the use of tailored neural machine translation systems for literary works, and their use in two applications: (I) the creation of bilingual electronic books to promote reading in the original language as a tool for language learning; (II) the promotion of literary works from smaller languages, whose authors have difficulty in reaching an international market.

I would also like to highlight my **international profile** with my participation in teaching and research activities in several countries (see *Teaching and research stays* section in this CV), programme committees of international conferences (see *Programme and scientific committees of international conferences* in this CV) and invited conferences and workshops (see *Invited conferences and workshops* in this CV).

As far as my **leadership skills** are concerned, I can provide evidence in the field of teaching and research. I'm the director of the Master in Translation and Technologies at the UOC, where I coordinate a team of 4 professors and about 10 external lecturers. In terms of research, I'm the principal investigator of a national project, where I coordinate a team of 3 members of the research team and 5 members of the working team. I've also been the principal investigator of 10 technology transfer projects with companies and institutions over the last 5 years. In these projects I had to coordinate with teams in the companies and institutions.

Regarding the **training of young researchers**, I have three thesis directions read and one thesis in progress. From the direction of theses, I would like to highlight the number of published articles derived from theses. With some of my students we continue to collaborate actively in research after they have completed their PhD. One of my doctoral students, Clara Ginovart, obtained the extraordinary thesis award from the Universitat Pompeu Fabra. I have also supervised a total of 19 Master's thesis supervisions at the UOC and UAB.

I have also mentored several research stays by young researchers. Lucia Sanz, a PhD student from the University of Valladolid, spent 1 month at the UOC in 2023. During her stay, I organised a seminar on the training of neural machine translation systems of 10 hours (2 hours a day for 5 days). The seminar was open to the public and was attended by nine researchers. Also in 2023, the young postdoctoral

researcher Bojana Mikelenić from the University of Zagreb will stay for 15 days after receiving a "Research connections" grant from the UOC.

In 2022 and 2023 I collaborated with the EMTTI Master (*European Master in Technologies for Translation and Interpreting*) by mentoring several students for one month during their research internship. I mentored a total of 5 students: M. Guzenko, N. Sugrobova, A. Iletmis, A. Vestel (2023) and L. Kharatyan (2022).

I have a **strong record of publications** in indexed journals, books and book chapters and conference proceedings. Regarding papers in indexed journals, I have 40 papers (11 in the last 5 years) with 33 of them (10 in the last 5 years) in journals indexed in JCR or SJR, and 24 of them (6 in the last 5 years) indexed in WoS collection. Of these papers, in 10 of them I'm the only author and in 12 of them I'm the first author. These data show publication in high impact journals and that there is a reasonable balance between individual work and work in collaboration with other researchers. I also have 20 books or books chapters, with 3 of them in publishing companies in the first quartile of the SPI ranking. I would also like to highlight my 45 papers in conference proceedings (14 in the last 5 years).

In terms of **raising funds** for research, I would like to emphasise that I am the principal investigator of a national project funded by the Ministry of Science and Innovation in the call for Knowledge Generation Projects 2021. I have also participated in several national projects (Know, Know2, Skater, Tuner, Moment and DeepReading), where I have mainly worked on the creation of semantic lexical resources and in several European projects (Metis II, dedicated to statistical automatic translation from monolingual corpora and E-Title, focused on multilingual automatic subtitling). These national and European projects have resulted in numerous publications: 11 articles in indexed journals, 14 articles in conference proceedings and 1 book chapter. I have been the main researcher of an internal teaching innovation project at my university (Avi-Cat, dedicated to the creation of tools to support writing in Catalan) which has resulted in two presentations at conferences and an article indexed in WoS and Scopus. I have participated in two internal projects at my university, which have resulted in 7 articles in indexed journals, 1 paper in conference proceedings and 2 book chapters.

I have also **raised funds** through technology transfer projects with companies and institutions, I have been the IP of 11 non-competitive research projects. Most of these projects are in the field of automatic terminology extraction and machine translation and are supported by tools that I have developed and which are distributed under a free licence: MTUOC (a set of tools to facilitate the training and integration of statistical and neural machine translation); and TBXTools (an automatic terminology extraction tool). The transfer projects I have led have also generated a remarkable number of publications (4 papers in indexed journals and 3 papers in conference proceedings).

In terms of teaching, I'm an associate professor (*professor agregat*) at the Arts and Humanities department of the Universitat Oberta de Catalunya, where I coordinate several subjects at both undergraduate and masters level. I am also an associate professor at the Universitat Autònoma de Barcelona (UAB) where I teach subjects related to translation technologies and supervise master's degree theses. I have won two scholarships (2020 and 2021) to be a visiting scholar in the European Masters in Technology for Translation and Interpreting (EMTTI) at the University of Wolverhampton (United Kingdom).

As other merits, I would like to highlight that I have the advanced research accreditation (2023), the research accreditation (2019), two recognized research sexennials and that I am a member of the GRIAL research group, a consolidated research group of the Generalitat de Catalunya (2021 SGR 00581). I have 16 invited conferences and a large number of presentations at conferences, most of which have led to articles in proceedings (a total of 40 articles with 25 of them indexed in Scopus). I have done 4 short research stays (U. Surrey 2022, U. Wolverhampton 2018, U. Vrije Amsterdam 2014 and U. Zagreb 2014). As regards positions of scientific responsibility, I am a member of the scientific committee of the journal Linguamatica (indexed in SJR, Scopus and CIRC A, among others). I have also been a member of 17 scientific and programme committees for conferences and workshops.

Education

phD in Linguistics	Universitat de Barcelona (UB)	27/07/2004
Master degree in Free Software	Universitat Oberta de Catalunya (UOC)	15/07/2011
Degree in Slavonic Philology	Universitat de Barcelona (UB)	30/06/1998
Telecommunications engineering	Universitat Politècnica de Catalunya (UPC)	30/06/1991

PhD Thesis

Title: Adquisició d'informació lèxica i morfosintàctica a partir de corpus sense anotar: aplicació al rus i al croat
Supervisors: Irene Castellón (UB), Lluís Màrquez (UPC)
University: Universitat de Barcelona (UB)
Department: General Linguistics
Date: 27/07/2004
Qualification: Excellent cum laude unanimously

Work Experience

Associate professor	Universitat Oberta de Catalunya	2006 -
Researcher	Universitat Pompeu Fabra	2004-2006
Researcher	Universitat Oberta de Catalunya	2002-2003
Computational linguist	Universitat Oberta de Catalunya	2004-2005

Projects

Title: Análisis multidimensional de la traducción automática en el tercer sector social: el discurso traducido y poseditado
Founded by: Vicerectorado de Investigación de la Universidad de Alcalá
Reference: PIUAH23/AH-11
Principal Investigator: María del Mar Sánchez Ramos
Start date: 1/01/2024   **End date**: 28/02/2025

Title: TAN-IBE: Neural Machine Translation for the Romance languages of the Iberian Peninsula
Founded by: Spanish Ministry of Science and Innovation
Reference: PID2021-124663OB-100
Principal Investigator: Antoni Oliver
Start date: 1/09/2022   **End date**: 31/08/2025
Granted budget: 69,251 €

Title: Imprentas - SARDWARE
Founded by: Regione Autònoma de Sardigna
Reference: Imprentas - Linia D - Alegato A
Principal Investigator: Adrià Martín (UAB)
Start date: 1/01/2021   **End date**: 31/12/2021
Granted budget: 18,000 €

Title: DeepReading: Mining, Understanding, and Reasoning with Multilingual Content
Founded by: MCIU/AEI/FEDER/UE
Reference: RTI2018-096846-B-C21
Principal Investigator: German Rigau (UPV/EHU)

Start date: 1/01/2019 **End date**: 31/12/2021
Granted budget: 146,894 €

Title: CRISIS-MT: Traducción y tecnología en la comunicación multilingüe para las ONGs en situaciones de emergencia
Founded by: Universidad de Alcalà
Reference: CCG2018/HUM-043
Principal Investigator: María del Mar Sánchez (UAH)
Start date: 20/12/2018 **End date**: 19/12/2019
Granted budget: 7,500 €

Title: MOMENT project: Metaphors of severe mental disorders Discourse analysis of affected people and mental health professionals:
Founded by: Spanish Ministry of Economy, Industry and Competitivity.
Reference: FPI2017-86969-R
Principal Investigator: Marta Coll-Florit (UOC)
Start date: 1/01/2018 **End date**: 31/12/2020
Granted budget: 56,991 €

Title: TUNER: Automatic domain adaptation for semantic processing
Founded by: Spanish Ministry of Science and Innovation
Reference: TIN2015-65308-C5-1-R
Principal Investigator: German Rigau (UPV/EHU)
Start date: 1/01/2016 **End date**: 31/12/2018
Granted budget: 191,906 €

Title: SKATER: Scenario knowledge acquisition by textual reading: linguistics and cognition
Founded by: Spanish Ministry of Economy and Competitivity
Reference: TIN2012-38584-C06-06
Principal Investigator: Irene Castellón (UB)
Start date: 2013 **End date**: 2016
Granted budget: 8,000 €

Title: Anàlisi forense: detecció d'autoria a partir de patrons d'escriptura
Founded by: Universitat Oberta de Catalunya
Reference: Aplica 2015
Principal Investigator: David Bañeres (UOC)
Start date: 2015 **End date**: 2016
Granted budget: 3,000 €

Title: InLéctor: llibres electrònics bilingües i amb suport d'àudio
Founded by: Universitat Oberta de Catalunya
Reference: Aplica 2013
Principal Investigator: Marta Coll-Florit
Start date: 1/03/2013 **End date**: 31/12/2013
Granted budget:

Title: Expresso: eina de suport a l'expressió escrita (avaluació i formació)
Founded by: Universitat Oberta de Catalunaya
Reference: Aplica 2012
Principal Investigator: Ona Domènech (UOC)
Start date: 1/03/2012 **End date**: 31/12/2012

Granted budget: 3,000 €

Title: KNOW2: Language understanding technologies for multilingual domain-oriented information access
Founded by: Spanish Ministry of Science and Innovation
Reference: TIN2009-14715-C04-03
Principal Investigator: Salvador Climent (UOC)
Start date: 1/01/2010   **End date**: 31/12/2012
Granted budget: 54,900 €

Title: AVI.cat: assessor virtual per a la millora de la redacció en català
Founded by: Universitat Oberta de Catalunya
Reference: Aplica 2013
Principal Investigator: Antoni Oliver
Start date: 1/03/2011   **End date**: 31/12/2011
Granted budget: 3,000 €

Title: TINM: Tratamiento de la información multilingüe y multimodal
Founded by: Spanish Ministry of Education and Science
Reference: TIN2006-28218-E
Principal Investigator: Alfonso Ureña
Start date: 1/07/2006   **End date**: 1/07/2011
Granted budget: 40,000 €

Title: Multilingual Central Repository 2.0: Grial
Founded by: Spanish Ministry of Science and Innovation
Reference: MICINN-FFI2009-08466-E/FILO
Principal Investigator: Irene Castellón (UB)
Start date: 2009   **End date**: 2010
Granted budget: 25,000 €

Title: KNOW: Developing large-scale multilingual technologies for language understanding
Founded by: Spanish Ministry of Science and Innovation
Reference: TIN2006-15049-C03
Principal Investigator: Irene Castellón (UB)
Start date: 1/10/2006   **End date**: 1/10/2009
Granted budget: 42,000 €

Title: METIS-II - Statistical Machine Translation Using Monolingual Corpora: from concept to implementation
Founded by: European Union
Reference: Future and emerging technologies (FET) unit of the IST programme
Principal Investigator: Toni Badia (UPF)
Start date: 1/10/2004   **End date**: 1/10/2007
Granted budget: 1,000,000 €

Title: AMEDIDA: Desarrollo de un sistema de transformación de contenidos a medida de los usuarios
Founded by: Spanish Ministry of Industry, Turism and Commerce
Reference: FIT-350201-2004-6
Principal Investigator: Ramon Recio
Start date: 2004   **End date**: 2005

Granted budget: 536,628 €

Title: e-Title: European Multilingual Transcriptionsand Subtitling Services for Digital Media
Founded by: European Comission
Reference: e-content programme EDC-22160
Principal Investigator: Toni Badia (UPF)
Start date: 1/03/2003   **End date**: 31/12/2005
Granted budget: 1,350,000 €

Title: Interlingua (Desenvolupament, avaluació, experimentació i aplicació de tècniques per a la traducció automàtica no supervisada i el tractament de la informació textual)
Founded by: Internet Interdisciplinary Institute (IN3)
Reference: IN3-IR226
Principal Investigator: Salvador Climent (UOC)
Start date: 2002   **End date**: 2003
Granted budget: 43,800 €

PUBLICATIONS

Summary:

	Overall	Last 5 years
Papers in indexed journals	40	11
Indexed in JCR/SJR	33	10
Indexed in WoS Core Collection	6	24
Unique autorship	10	2
First author	12	-
Books and book chapters	20	6
Q1 SPI	3	3
Papers in conference proceedings	48	17

Statistics Google Scholar

	All	From 2019
Citations	1,024	482
h index	16	11
i10 index	32	16

Papers in indexed journals

Oliver, A. (2024). Corpora and rhetorically informed text analysis (book review). *ES Review: Spanish journal of English Studies*, (45), 296–300.

Oliver, A. (2022). El corpus paral·lel del Diari Oficial de la Generalitat de Catalunya. *Linguamática, 14* (2), 75–81. https://doi.org/10.21814/lm.14.2.380

Álvarez-Vidal, S., Oliver, A., & Badia, T. (2021). What do post-editors correct? A fine-grained analysis of SMT and NMT errors. *Tradumàtica: tecnologies de la traducció*, (19), 131–147. https://doi.org/10.5565/rev/tradumatica.286

Coll-Florit, M., Oliver, A., & Climent, S. (2021). Metaphors of mental illness: A corpus-based approach analysing first-person accounts of patients and mental health professionals. *Culture, Language & Representation/Cultura, Lenguaje y Representación*, *25*(2), 85–104. https://doi.org/10.6035/clr.2021.25.5

Mauri, F., Sánchez-Gijón, P., & Oliver, A. (2021). Cadlaws–an english–french parallel corpus of legally equivalent documents. *Mutatis Mutandis: Revista Latinoamericana de Traducción*, *14*(2), 494–508. https://doi.org/10.17533/udea.mut.v14n2a10

Mikelenić, B., & Oliver, A. (2021). El complemento directo y de régimen en español y sus equivalencias en croata: Un estudio basado en corpus. *Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics*, *34*(1), 255–278. https://doi.org/10.1075/resla.18053.mik

Álvarez-Vidal, S., Oliver, A., & Badia, T. (2020). Post-editing for professional translators: Cheer or fear? *Tradumàtica*, (18), 49–69. https://doi.org/10.5565/rev/Tradumatica.275

Ginovart, C., Colominas, C., & Oliver, A. (2020). Language industry views on the profile of the post-editor. *Translation Spaces*, *9*(2), 283–313. https://doi.org/10.1075/ts.19010.cid

Oliver, A. (2020). Traducción automática para las lenguas románicas de la penınsula Ibérica. *Studia Romanica et Anglica Zagrabiensia: Revue publiée par les Sections romane, italienne et anglaise de la Faculté des Lettres de l'Université de Zagreb*, *65*, 367–375. https://doi.org/10.17234/SRAZ.65.45

Rivera, A. T., Oliver, A., & Coll-Florit, M. (2020). Metaphoricity detection in adjective-noun pairs. *Procesamiento del Lenguaje Natural*, *64*, 53–60. https://doi.org/10.26342/2020-64-6

Vázquez, M., Oliver, A., & Casademont, E. (2019). Using open data to create the catalan IATE e-dictionary. *Terminology*, *25*(2), 175–197. https://doi.org/10.1075/term.00035.vaz

Coll-Florit, M., Climent, S., Correa-Urquiza, M., Hernánndez, E., Oliver, A., & Pié, A. (2018). MOMENT: Metáforas del trastorno mental grave. análisis del discurso de personas afectadas y profesionales de la salud mental. *Procesamiento del Lenguaje Natural*, *61*, 139–142. https://doi.org/10.26342/2018-61-17

Vázquez, M., & Oliver, A. (2018). Improving term candidates selection using terminological tokens. *Terminology*, *24*(1), 122–147. https://doi.org/10.1075/term.00016.vaz

Vázquez, M., Oliver, A., & Ubide, G. (2018). La terminologíaa jurídica del IATE en català. *Revista de Llengua i Dret*, (69), 139–153. https://doi.org/10.2436/rld.i69.2018.3010

Coy Grueso, M., & Oliver, A. (2017). Robinson Crusoe: Creación de la edición electrónica bilingüe. el proyecto InLéctor. *Estudios de Traducción*, *7*, 205–217. https://doi.org/10.5209/ESTR.57457

Oliver, A. (2017a). El corpus paral·lel del Diari Oficial de la Generalitat de Catalunya: Compilació, anàlisi i exemples d'ús. *Zeitschrift für Katalanistik*, *30*, 269–291. https://doi.org/10.46586/zfk.2017.269-291

Oliver, A. (2017b). InLéctor: Automatic creation of bilingual e-books. *Tradumàtica: traducció i tecnologies de la informació i la comunicació*, (15). https://doi.org/10.5565/rev/tradumatica.191

Oliver, A. (2017c). A system for terminology extraction and translation equivalent detection in real time: Efficient use of statistical machine translation phrase tables. *Machine Translation*, *31*(3), 147–161. https://doi.org/10.1007/s10590-017-9201-7

Oliver, A., Vázquez, M., & Ubide, G. (2017). Estudi de la fiabilitat de la viquipèdia com a recurs terminològic. *Tradumàtica: traducció i tecnologies de la informació i la comunicació*, (15), 10–20. https://doi.org/10.5565/rev/tradumatica.193

Oliver, A. (2015). TMX: Intercanvi de memòries de traducció. *Tradumàtica: traducció i tecnologies de la informació i la comunicació*, (13), 608–622.

Gómez Guinovart, X., & Oliver, A. (2014). Methodology and evaluation of the Galician WordNet expansion with the WN-Toolkit. *Procesamiento del Lenguaje Natural*, (53), 43–50.

González, A. O. (2012). WN-Toolkit: Un toolkit per a la creació de wordnets a partir de diccionaris bilingües. *Linguamática*, *4*(2), 93–101.

Oliver, A., & Abuin Castro, M. (2012). inLéctor: Creación de libros electrónicos bilingües interactivos. *Linguamática*, *4*(2), 57–63.

Oliver, A., & Climent, S. (2012a). Parallel corpora for wordnet construction: Machine translation vs. automatic sense tagging. *Computational Linguistics and Intelligent Text Processing: 13th International Conference, CICLing 2012, New Delhi, India, March 11-17, 2012, Proceedings, Part II 13*, 110–121.

Oliver, A., & Climent, S. (2012b). Ús de la Viquipèdia per al desenvolupament de recursos lingüístics: El WordNet 3.0 català i castellà. *Digithum*, (14), 15–24.

Oliver, A., Climent, S., & Coll-Florit, M. (2012). InLéctor: Sistema de lectura bilingüe interactiva. *Procesamiento del Lenguaje Natural*, *49*, 203–207.

Oliver, A., Climent, S., & Contreras, M. (2012). Revisión de técnicas para la construcción de WordNets mediante la estrategia de expansión. *Procesamiento del Lenguaje Natural*, (49), 119–127.

Oliver, A., & Climent, S. (2011). Construcción de los WordNets 3.0 para castellano y catalán mediante traducción automática de corpus anotados semánticamente. *Procesamiento del Lenguaje Natural*, *47*, 293–300.

Oliver, A., Climent, S., & Coll-Florit, M. (2011). AVI.cat: Asistente virtual para la mejora de la redacción en catalán. *Procesamiento del Lenguaje Natural*, *47*, 313–314.

Oliver, A., & Borrell, C. (2008). Com es treballa en una aula virtual multilingüe? Un entorn d'aprenentatge amb Moodle i Apertium. *UOC Papers: revista sobre la societat del coneixement*, (6), 6.

Oliver, A., Vázquez, M., & Moré, J. (2007). Linguoc LexTerm: Una herramienta gratuita de extracción automática de terminología. *Translation Journal*, *4*(11), 1–5.

Oliver, A. (2006). La traducció automàtica a Internet. *Tradumàtica: traducció i tecnologies de la informació i la comunicació*, (4).

Badia, T., Boleda, G., Melero, M., & Oliver, A. (2005). El proyecto METIS-II. *Procesamiento del Lenguaje Natural*, (35), 443–444.

Climent, S., Moré, J., Oliver, A., Salvatierra, M., Sánchez, I., & Vázquez, M. (2005). Tecnologies de la traducció per a la gestió de la doble oferta docent en català i castellà a la UOC. *Zeitschrift für Katalanistik/Revista d'Estudis Catalans*, (18), 31–57.

Moré, J., Climent, S. C. R., Oliver, A., & Taulé Delor, M. (2005). Análisis de los fenómenos lingüísticos de los mensajes de correo electrónico en catalán desde la prespectiva de la traducción automática. *Procesamiento del Lenguaje Natural*, (35), 45–50.

Oliver, A. (2005). Adquisició d'informació lèxica i morfosintàctica a partir de corpus sense anotar aplicació al rus i al croat. *Procesamiento del Lenguaje Natural, 34*, 115–116.

Oliver, A., Boleda, G., Melero, M., & Badia, T. (2005). Traducción automática estadística basada en n-gramas. *Procesamiento del Lenguaje Natural, 35*.

Oliver, A. (2004). Qualitat i traducció. *Llengua i ús: Revista tècnica de política lingüística*, (29), 31–36.

Climent, S., Moré, J., Oliver, A., Salvatierra, M., Sánchez, I., Taulé, M., & Vallmanya, L. (2003). Bilingual newsgroups in Catalonia: A challenge for machine translation. *Journal of Computer-Mediated Communication, 9*(1).

Oliver, A., Castellón, I., & Márquez, L. (2003). Uso de internet para aumentar la cobertura de un sistema de adquisición léxica del ruso. *Procesamiento del Lenguaje Natural*, (31), 269–276.

Oliver, A., Castellón, I., & Márquez, L. (2002). Adquisición automática de información léxica y morfosintáctica a partir de corpus sin anotar: Aplicación al serbocroata y ruso. *Procesamiento del Lenguaje Natural*, (29), 97–104.

Books and book chapters

Oliver, A. (2023a). Author tailored neural machine translation systems for literary works. In A. Rothwell, A. Way, & R. Youdale (Eds.), *Computer-assisted literary translation*. Routledge – Taylor Francis Group.

Oliver, A. (2023b). Entrenamiento de motores de traducción automática. In M. del Mar Sánchez & C. Rico (Eds.), *Traducción automática en contextos especializados*. Peter Lang.

Oliver, A. (2023c). Especialista en traducción automática. In S. Queralt (Ed.), *Lingüistas de hoy* (pp. 131–137). Editorial Síntesis.

Ginovart, C., & Oliver, A. (2020). The post-editor's skill set according to industry. In J. Porsiel (Ed.), *Maschinelle Übersetzung für Übersetzungsprofis* (pp. 305–322). BDÜ Fachverlag.

Toral, A., Oliver, A., & Ribas, P. (2020). Machine translation of novels in the age of transformer. In J. Porsiel (Ed.), *Maschinelle Übersetzung für Übersetzungsprofis* (pp. 276–295). BDÜ Fachverlag.

Iribarren, T., Oliver, A., & Peiró, E. (2017). Traduccions a Internet: El cas de L'illa dels pingüins, d'Anatole France, en traducció de JF Vidal Jové. In M. Bacardí & P. Godayol (Eds.), *Traducció i franquisme* (pp. 139–152). Punctum.

Oliver, A. (2016). *Herramientas tecnológicas para traductores*. Editorial UOC.

Oliver, A., Sojat, K., & Srebacic, M. (2016). Automatic expansion of the Croatian Wordnet. In K. C. Kovačević & S. L. Udier (Eds.), *Metodologija i primjena lingvistickih iztrazivanja* (pp. 139–152). Srednja Europa.

Oliver, A., Benvant, F., Gómez, X., Martí, Q., Alonso, J. A., Vàzquez, A., & Domingo, J. (2010). *Llengua catalana i tecnologies digitals*. Fundació per a la Universitat Oberta de Catalunya.

Oliver, A., Xavier, G., & Irene, C. (2010). *Lingüística computacional.* Fundació per a la Universitat Oberta de Catalunya.

Vázquez, M., & Oliver, A. (2010). Ús d'estratègies per a l'extracció automàtica d'unitats terminològiques. In J. Martí & M. Salse (Eds.), *La terminologia i la documentació: Relacions i sinergies* (pp. 75–84). Institut d'Estudis Catalans.

Vázquez, M., & Oliver, A. (2009). Utilisation de stratégies pour la recuperation automatique de termes. In M. van Campenhoudt, T. Lino, & R. Costa (Eds.), *Passeurs de mots, passeurs d'espoir. lixcologies, terminologies et traduction face au défi de la diversité* (pp. 411–428). Editions des archives contemporaines.

Climent, S., Moré, J., Oliver, A., Salvatierra, M., Sánchez, I., Taulé, M., & Vallmanya, L. (2007). Enhancing the status of Catalan versus Spanish in online academic forums: Obstacles to machine translation. In B. Danet & S. C. Herring (Eds.), *The multilingual internet: Language, culture, and communication online* (pp. 209–230). Oxford University Press New York, NY.

Oliver, A., Moré, J., & Climent, S. (2007a). *Tecnologies de la traducció.* Editorial UOC.

Oliver, A., Moré, J., & Climent, S. (2007b). *Traducció i tecnologies.* Editorial UOC.

Oliver, A., Moré, J., & Climent, S. (2007c). *Traducción y tecnologías.* Editorial UOC.

Climent, S., Moré, J., & Oliver, A. (2004). Traducción automática no supervisada de mensajes de correo electrónico. In J. M. Oro Cabanas, J. Anderson, & J. Varela Zapata (Eds.), *Lengua y sociedad: Lingüística aplicada en la era global y multicultural* (pp. 403–418).

Climent, S., Oliver, A., & Moré, J. (2004). *Traducció assistida per ordinador: Programes i recursos lliures i gratuïts.* Fundació per a la Universitat Oberta de Catalunya.

Climent, S., Moré, J., & Oliver, A. (2003). El proyecto Interlingua: Construcción de un entorno de traducción automática de e-mails. In S. Posteguillo, E. Ortells, J. R. Prado, & A. Alcina (Eds.), *Internet in linguistics, translation and literary studies* (pp. 175–189). Publicacions de la Universitat Jaume I.

Papers in conference proceedings

Mikelenić, B., & Oliver, A. (2024). Mikelenić, b., oliver, a. (2024). using a multilingual literary corpus to train NMT systems. *Proceedings of the 1st Workshop on Creative-text Translation and Technology (CTT24)*, 3–11.

Oliver, A. (2024). TAN-IBE participation in the shared task: Translation into low-resource languages of Spain. *Proceedings of the Ninth Conference on Machine Translation*, 871–877. https://aclanthology.org/2024.wmt-1.83

Oliver, A., & Álvarez-Vidal, S. (2024). LitPC: A set of tools for building parallel corpora from literary works. *Proceedings of the 1st Workshop on Creative-text Translation and Technology (CTT24)*, 25–36.

Oliver, A., Álvarez-Vidal, S., Stemle, E., & Chiocchetti, E. (2024). Training an NMT system for legal texts of a low-resource language variety (South Tyrolean german - Italian). *In Proceedings of the 25th Annual Conference of the European Association for Machine Translation (Volume 1)*, 573–579.

Perez-Ortiz, J. A., Sánchez-Martínez, F., Sánchez-Cartagena, V. M., Esplà-Gomis, M., Galiano Jimenez, A., Oliver, A., Aventín-Boya, C., Pardos, A., Valdés, C., Sans Socasau, J. L., & Martínez, J. P.

(2024). Expanding the FLORES+ multilingual benchmark with translations for Aragonese, Aranese, Asturian, and Valencian. *Proceedings of the Ninth Conference on Machine Translation*, 547–555. https://aclanthology.org/2024.wmt-1.41

Sánchez-Martínez, F., Perez-Ortiz, J. A., Galiano Jimenez, A., & Oliver, A. (2024). Findings of the WMT 2024 shared task translation into low-resource languages of Spain: Blending rule-based and neural systems. *Proceedings of the Ninth Conference on Machine Translation*, 684–698. https://aclanthology.org/2024.wmt-1.57

Álvarez, S., & Oliver, A. (2023). PE effort and neural-based automatic MT metrics: Do they correlate? *In Proceedings of the 24th Annual Conference of the European Association for Machine Translation - EAMT 2023*, 315–323.

Oliver, A., & Álvarez, S. (2023a). Filtering and rescoring the CC-Matrix corpus for neural machine translation. *In Proceedings of the 24th Annual Conference of the European Association for Machine Translation - EAMT 2023*, 39–45.

Oliver, A., & Álvarez, S. (2023b). Training and integration of neural machine translation with MTUOC. *In Proceedings of the 1st Workshop on Open Community-Driven Machine Translation (crowdMT). 24th Annual Conference of the European Association for Machine Translation - EAMT 2023*, 5–13.

Oliver, A., Vázquez, M., Coll-Florit, M., Álvarez, S., Suárez, V., Aventín-Boya, C., Valdés, C., Font, M., & Pardos, A. (2023a). TAN-IBE: Neural machine translation for the romance languages of the iberian peninsula. *In Proceedings of the Annual Conference of the Spanish Association for Natural Language Processing 2023: Projects and System Demonstrations (SEPLN-PD 2023) co-located with the International Conference of the Spanish Society for Natural Language Processing (SEPLN 2023)*, 45–49.

Oliver, A., Vázquez, M., Coll-Florit, M., Álvarez, S., Suárez, V., Aventín-Boya, C., Valdés, C., Font, M., & Pardos, A. (2023b). TAN-IBE: Neural machine translation for the romance languages of the Iberian Peninsula. *In Proceedings of the 24th Annual Conference of the European Association for Machine Translation - EAMT 2023*, 495–498.

Oliver, A. (2022). MTUOC: A complete toolkit for training and integrating neural machine translation. *In Proceedings of the New Trends in Translation and Technology Conference 2022*, 4–6.

Álvarez, S., Oliver, A., & Badia, T. (2020). Quantitative analysis of post-editing effort indicators for NMT. *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation*, 411–420.

Oliver, A. (2020a). Aliging Wikipedia with Wordnet: A review and evaluation of different techniques. *In Proceedings of the 12th International Conference on Language Resources and Evaluation (LREC 2020)*, 4851–4858.

Oliver, A. (2020b). MTUOC: Easy and free integration of NMT systems in professional translation environments. *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation*, 467–468.

Oliver, A. (2020c). ReSiPC: A tool for complex searches in parallel corpora. *In Proceedings of the 12th International Conference on Language Resources and Evaluation (LREC 2020)*, 7033–7037.

Oliver, A., Alvarez, S., & Badia, T. (2020). PosEdiOn: Post-Editing Assessment in PythOn. *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation*, 403–410.

Oliver, A., & Vàzquez, M. (2020). Termeval 2020: Using tsr filtering method to improve automatic term extraction. *Proceedings of the 6th International Workshop on Computational Terminology*, 106–113.

Rico, C., Sánchez, M. d. M., & Oliver, A. (2020). Inmigra3: Building a case for NGOs and NMT. *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation*, 469–470.

Rivera, A. T., Oliver, A., Climent, S., & Coll-Florit, M. (2020). Neural metaphor detection with a residual biLSTM-CRF model. *Proceedings of the Second Workshop on Figurative Language Processing*, 197–203.

Álvarez, S., Oliver, A., & Badia, T. (2019). Does NMT make a difference when post-editing closely related languages? The case of Spanish-Catalan. *Proceedings of Machine Translation Summit XVII: Translator, Project and User Tracks*, 49–56.

Šojat, K., Filko, M., & Oliver, A. (2018). Further expansion of the Croatian WordNet. *Proceedings of the 9th Global Wordnet Conference*, 352–357.

Kumar, A., Cotterell, R., Oliver, A., & Padró, L. (2017). Morphological analysis of the Dravidian language family. *EACL 2017: 15th Conference of the European Chapter of the Association for Computational Linguistics: Valencia, Spain: April, 3-7 2017: Proceedings of Conference, Volume 1: Long Papers*, 217–222.

Oliver, A. (2016). Extending the WN-Toolkit: Dealing with polysemous words in the dictionary-based strategy. *Proceedings of the 8th Global WordNet Conference (GWC)*, 267–274.

Kumar, A., Padró, L., & Oliver, A. (2015a). Learning agglutinative morphology of Indian languages with linguistically motivated adaptor grammars. *Proceedings of the International Conference Recent Advances in Natural Language Processing*, 307–312.

Kumar, A., Padró, L., & Oliver, A. (2015b). Unsupervised learning of agglutinated morphology using nested Pitman-Yor process based morpheme induction algorithm. *2015 International Conference on Asian Language Processing (IALP)*, 45–48.

Oliver, A., Šojat, K., & Srebačić, M. (2015). Enlarging the Croatian wordnet with WN-Toolkit and Cro-Deriv. *Proceedings of the international conference recent advances in natural language processing*, 480–487.

Oliver, A., & Vàzquez, M. (2015). TBXTools: A free, fast and flexible tool for automatic terminology extraction. *Proceedings of the international conference recent advances in natural language processing*, 473–479.

Oliver, A. (2014). WN-Toolkit: Automatic generation of WordNets following the expand model. *Proceedings of the Seventh Global Wordnet Conference*, 7–15.

Oliver, A., & Climent, S. (2014). Automatic creation of WordNets from parallel corpora. *Proceedings of LREC 2014, Ninth International Conference on Language Resources and Evaluation*, 1112–1116.

Vàzquez, M., & Oliver, A. (2013). Improving term candidate validation using ranking metrics. *Proceedings of the 3rd World Conference on Information Technology (WCIT-2012)*, *3*, 1348–1359.

Lloberes, M., Oliver, A., Climent, S., & Castellón, I. (2012). Tratamiento de multipalabras en WordNet 3.0. *Proceedings of the XXX Congreso de la Asociación Española de Lingüística Aplicada (AELSA)*.

Oliver, A., & Climent, S. (2012). Building wordnets by machine translation of sense tagged corpora. *Proceedings of the Global WordNet Conference, Matsue, Japan.*

Oliver, A., Climent, S., & Coll-Florit, M. (2012). Avi.Cat: Turning grammar checking into an e-learning experience. *Proceedings of the IADIS International Cconference*, 403.

Vázquez, M., & Oliver, A. (2010). Ús d'estratègies estadístiques per a l'extracció automàtica d'unitats terminològiques. *La terminologia i la documentació: relacions i sinergies*, 75.

Oliver, A., & Borrell, C. (2009). IT: Moving towards real multilingualism. *Proceedings of the SLOVKO 2009 Conference. NLP, Corpus Linguistics, Corpus Based Grammar Research*, 279–286.

Alvez, J., Atserias, J., Carrera, J., Climent, S., Laparra, E., Oliver, A., & Rigau, G. (2008). Complete and consistent annotation of wordnet using the Top Concept Ontology. *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008).*

Álvez, J., Atserias, J., Carrera, J., Climent, S., Oliver, A., & Rigau, G. (2008). Consistent annotation of EuroWordNet with the Top Concept Ontology. *Proceedings of the GWC 2008: Fourth International Wordnet Conference*, 8.

Borrell, C., & Oliver, A. (2008). How do we work in a virtual multilingual classroom? A multilingual environment with Moodle and Apertium. *Proceedings of the CALPIU Conference 2008.*

Oliver, A., & Borrell, C. (2008a). Gestionando Babel: Un aula multilingüe en tiempo real. *Proceedings of the MoodleMoot Barcelona 2008.*

Oliver, A., & Borrell, C. (2008b). Multilingüismo en las instituciones europeas: Herramientas y recursos para la integración del catalán. *Proceedings of the Congreso Mundial de Traducción Especializada.*

Melero, M., Oliver, A., Badia, T., & Suñol, T. (2007). Dealing with bilingual divergences in MT using target language n-gram models. *Proceedings of the METIS-II Workshop: New Approaches to Machine Translation*, 19–26.

Oliver, A., & Vàzquez, M. (2007). A free terminology extraction suite. *Proceedings of Translating and the Computer 29.*

Melero, M., Oliver, A., & Badia, T. (2006). Automatic multilingual subtitling in the eTITLE project. *Proceedings of Translating and the Computer, 28*, 1–18.

Badia, T., Boleda, G., Melero, M., & Oliver, A. (2005). An n-gram approach to exploiting a monolingual corpus for machine translation. *Workshop on example-based machine translation*, 1–7.

Civit, M., Castelvi, J., Morante, R., Oliver, A., & Aparicio, J. (2005). 4LEX: A multilingual lexical resource. *Cross-Language Knowledge Induction Workshop*, 39–45.

Moré, J., Climent, S., & Oliver, A. (2004). A grammar and style checker based on Internet searches. *Proceedings of LREC 2004: 4th International Conference on language resources and evaluation.*

Oliver, A., & Tadic, M. (2004). Enlarging the Croatian morphological lexicon by automatic lexical acquisition from raw corpora. *Proceedings of LREC 2004: 4th International Conference on language resources and evaluation.*

Climent, S., Moré, J., & Oliver, A. (2003). Building an environment for unsupervised automatic e-mail translation. *IN3 Working Paper Series*, (3).

Oliver, A., Castellón, I., & Marquez, L. (2003). Use of internet for augmenting coverage in a lexical acquisition system from raw corpora: Application to Russian. *IESL Workshop of RANLP*, *3*.

Oliver, A., Castellón, I., & Màrquez, L. (2003). Automatic lexical acquisition from raw corpora: An application to Russian. *Proceedings of the 2003 EACL Workshop on Morphological Processing of Slavic Languages*, 17–24.

Oliver, A. (1993). Predicción de la inteligibilidad en recintos cerrados. *Proceedings of Tecniacústica 93*.

Teaching and research stays

- 2023 - Erasmus teaching stay at the University of Zagreb (Croatia) - 5 days

- 2022 - Research stay at the University of Surrey (United Kingdom) - 4 days

- 2021 - Visiting scholar grant at the EMTTI master, University of Wolverhampton (United Kingdom) - performed online due to Covid pandemic

- 202 - Visiting scholar grant at the EMTTI master, University of Wolverhampton (United Kingdom) - performed online due to Covid pandemic

- 2018 - Research stay at the University of Wolverhampton (United Kingdom) - 12 days

- 2015 - Research stay at the Vrije Universiteit Amsterdam (The Netherlands) - 21 days

- 2015 - Teaching stay at the Facultat de traducción, interpretación y ciencias de la comunicación at UNIFE (Universidad Femenina del Sagrado Corazón) Lima (Perú) - 5 days

- 2014 - Research stay at the University of Zagreb (Croatia) - 15 days

- 2014 - Erasmus teaching stay at the University College London (United Kingdom) - 5 days

Editorial commitees of scientific journals

Member of the advisory board of the journal *Procesamiento del Lenguaje Natural* (http://journal.sepln.org/) from 2013.

Programme and scientific commitees of international conferences

- 27th European Conference on Artificial Intelligence

- COLM 2024 - Conference on Language Modelling

- EAMT 2024 - Conference of the European Association for Machine Translation

- SLATE23 - Symposium ont language, applications and technology

- 2023 - TC45: Translating and the Computer - Luxembourg

- 2023 - EAMT 2023: The 24th Annual Conference of the European Association for Machine Translation - Tampere (Finland)

- 2023 - IJCNLP-AACL 2023: The 13th Joint Conference on Natural Language Processing and the 3rd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics - Bali (Indonesia)

- 2023 - EACL 2023 - The 17th Conference of the European Association for Computational Linguistics - Dubrovnik (Croatia)

- 2023 - RANLP 2023: Recent Advances in Natural Language Processing - Varna (Bulgaria)

- 2022 - AACL-IJCNLP 2022 The 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th Joint Conference on Natural Language Processing - Taiwan

- 2022 - EAMT 2022: The 23rd Annual Conference of the European Association for Machine Translation - Ghent (Belgium)

- 2021 - RANLP 2021: Recent Advances in Natural Language Processing - online

- 2020 - EMNLP 2020 - The 2020 Conference on Empirical Methods in Natural Language Processing - online

- 2020 - ACL 2020: The 58th Annual Meeting of the Association for Computational Linguistics - online

- EAMT 2020 - The 22nd Annual Conference of the European Association for Machine Translation - online

- 2020 - SLATE 2020 - Symposium on Languages, Applications and Technologies - online

- 2020 - LREC 2020: Language Resources and Evaluation Conference - online

- 2018 - Globalex 2018: LREC 2018 workshop "Lexicography and WordNets" - Miyazaki (Japan)

- 2018 - SLATE 2018: Symposium on Languages, Applications and Technologies - Minho (Portugal)

- 2016 - LREC 2016: Language Resources and Evaluation Conference - Portoroz (Slovenija)

- 2015 - EAMT 2015: The 18th Annual Conference of the European Association for Machine Translation- Antalya (Turkey)

INVITED CONFERENCES AND WORKSHOPS

- 2024 - Invited talk "Traducción automática en contexts especializados" in the Translating Europe Workshops

- 2023 - Invited Workshop "Creación de corpus paralelos para la enseñanza de la traducción" at the XJ Jornadas Nacionales de la Enseñanza de la Traducción y la Interpretación FORTICH - Copiapó (Chile)

- 2023 - Invited conference "Sistemas de traducción automática libres: uso e integración" organized by Secretaría de Ciencia y Técnica de la Facultad de Lenguas, la Oficina de Diplomaturas y la Diplomatura de Localización en Software, Facultad de Lenguas. Universidad Nacional de Córdoba - Córdoba (Argentina)

- 2023 - Invited conference "Integración de técnicas de inteligencia artificial en entornos productivos de traducción" at ANETI 2023: IV Congreso de la Asociación Nacional de Empresas de Traducción - Madrid (Spain)

- 2022 - Invited plenary conference "Traducción automática neuronal y lenguas romances de la Península Ibérica" in the conference "Traducción, Tecnologías e Interación Humana" - Departamento de Estudios Románicos, Franceses, Italianos y Traducción. Facultad de Filología. Universidad Complutense

- 2022 - Invited workshop "Entrenamiento y evaluación de motores de traducción automática" organized by the Departamento de Fiología Moderna de la Universidad de Alcalá - Alcalá de Henares (Spain)

- 2022 - Inaugural invited conference "Recursos y herramientas libres para la traducción automática

neuronal" at the IV Encuentro de Estudiantes de Traducción e Interpretación Zona Norte. Universidad de Atacama . Copiapó (Chile)

- 2021 - Invited conference "Integración de traducción automática neuronal con MTUOC" at the III Congreso para empresas y profesionales de la traducción de ANETI (Asociación Nacional de Empresas de Traducción e Interpretación" - Madrid (Spain)

- 2020 - Invited conference "Techniques for automatic terminology extraction: implementation into TBXTools" in the Seminar Series of the European Masters in Technology for Translation and Interpreting" - online

- 2020 - Invited conference "Integración de Traducción Automática Neuronal en herramientas TAO con MTUOC". Universidad Nacional de Córdoba (Argentina) - online

- 2020 - Participation in the Expert Meeting "Creative Translation and Technologies". Centre for Translation Studies. University of Surrey - Guilford (United Kingdom)

- 2019 - Invited conference "Traducción y tecnología: ¿enemigos o aliados? Universidad de Atacama. Copiapó. Chile

- 2018 - Invited conference "Avances recientes en traducción automática neuronal ¿cómo puede beneficiar al traductor humano? IX Jornadas Nacionales de la Federación Argentina de Traductores. Córdoba. Argentina.

- 2010 - Invited conference "Automatic detection of translation equivalents of temrs in large parallel and comparable corpora" in the Seminar seiers of the Research Group in Computational Linguistics (RGL) - University of Wolverhampton. United Kingdom.

- 2018 - Invited conference "La Agencia Virtual de Traducción: simulación de un entorno profesional de traducción para la ralización de prácticas". I Jornadas Tecnologías aplicadas a la enseñanza de la traducció e interpretación. Córdoba. Argentina.

- 2017 - Participation in the roundtable "Nuevos marcos docentes, nuevas metodologías docentes, nuevas orientaciones. La formación en terminología: un espacio en movimiento" - XVII Jornadas de la Asociación española de terminología. Universitat Pompeu Fabra. Barcelona

- 2015 - Invited talk "Extending the WN-Toolkit: dealing with polisemous words in the dictionary-based strategy" - Computational Lexicology and Terminology Lab (CLTL). Vrije Universiteit Amsterdam. The Netherlands.

- 2014 - Participation in the roundtable "Passat i futur de les tecnologies de la traducció" - Jornades 20 anys de traducció i interpretació a la UVic-UCC. Vic. (Spain)

- 2014 - Invited conference "TBXTools: Automatic Terminology Extraction" - II Abumatran Summer Workshop. Elche (Spain)

- 2014 - Invited conference "Leer y aprender lenguas: libros electrónicos bilingües con soporte de audio" - IV Jornadas Internacionales de tecnologías aplicadas a la enseñanza de lenguas. Córdoba. Argentina.

- 2011 - Invited talk "WN-Toolkit: automatic creation of WordNets: results for Croatian and Slovene (preliminary) in the NLP-related talks at the University of Ljubljana (Slovenija)

- 20122 - Invited conference "Software libre y didáctica de la traducción" - III Jornadas de tecnologías aplicadas a la enseñanza de lenguas. Córdoba. Argentina.

- 2011 - Invited conference "Software libre y traducción" a the Colegio de traductores públicos de la provincia de Córdoba. Córdoba. Argentina.

Last updated: November 18, 2024