

# Leopoldo Pla Sempere

---

## Education (Science)

- 2014–2015 **Master's Degree**, *Universitat Politècnica de València*, Valencia, *Master's Degree in Artificial Intelligence, Pattern Recognition and Digital Imaging*.  
Specialised in natural language processing and deep learning. Enroled in machine learning MOOCs of Andrew Ng (Stanford) and Yaser Abu-Mostafa (Caltech)
- 2010–2014 **Degree**, *University of Alacant*, San Vicent del Raspeig, *Computer Science*.  
B.Sc. specialized in Informatics (robotics, machine learning, artificial vision, compilers) in the group of high profile academics. Awarded with the highest mark in thesis (outstanding with honors) and best academic record
- 2008–2010 **GCE**, *Instituto IES Cap de l'Aljub*, Santa Pola, *Science*.  
Including mathematics, physics, chemistry and computer programming class
- 2004–2008 **GSCE**, *C.P. Azorín and Instituto IES Cap de l'Aljub*, Santa Pola, .  
Including English and mathematics.

## Education (Music)

- 2004–2010 **Professional degree**, *Elche School of Music*, Elche, *Clarinet*.  
Specialized in jazz improvisation with Miguel García Ferrer. Courshed MOOCs about jazz improvisation with Gary Burton (Berklee School of Music) and fundamentals of rehearsing music ensembles with Dr. Evan Feldman from the University of North Carolina at Chapel Hill
- 2000–2004 **Elemental degree**, *Elche School of Music*, Elche, *Clarinet*.  
Elemental instrument studies of music specialized in clarinet

## Master thesis

- Title *Audio classical composer identification in MIREX 2015: submission based on Structural Analysis of Music*
- Supervisors Roberto Paredes Palacios

## Degree thesis

- Title *Dodecaphonic music composer assistant with OpenMusic*
- Supervisors Carlos Pérez Sancho

## Work experience

- 2019–present **Part-time lecturer**, *University of Alacant*, San Vicent del Raspeig.  
Lecturer in computer science subjects, from the Departament of Software and Computing Systems (University of Alacant). Teaching in several degrees as Mathematics and Computer Science, from programming fundamentals to software specification and planning.

- 2017–present **Natural Language Processing and Machine Learning researcher and developer**, *University of Alacant*, San Vicent del Raspeig.  
 Head developer in EU funded projects ParaCrawl (Provision of Web-Scale Parallel Corpora for Official European Languages), Paracrawl 2 (Broader Web-Scale Provision of Parallel Corpora for European Languages) and Paracrawl 3 (Continued Provision of Web-Scale Parallel Corpora for European Languages) in coordination with University of Edinburgh, Prompsit, TAUS, Omnisic and John Hopkins University. The project focus on creating corpora for all EU official languages (full list at project site) by a broad web crawling effort. State-of-the-art methods are designed and developed for the entire processing chain, from identifying web sites with translated text all the way to cleaned parallel corpora that are ready as training data for CEF Automated Translation. Bitextor (bash, python, snakemake, awk, mgiza, marian...), the production-ready tool that merges all this necessary technology chain for the project, is coordinated and maintained at the Department of Software and Computing systems of the University of Alacant.
- 2015–2017 **Natural Language Processing and Machine Learning researcher and developer**, *Prompsit Language Engineering*, Elche.  
 Developing natural language related projects as Reverso Context, processing full range of corpora and language resources. This precise project, developed for Reverso/Softissimo (approx. rank 300 in Global Alexa Ranking) reaching several million page-views per day, grown in an agile development team, using technologies as SVN, pandas, jupyter, bash (UNIX tools), word2vec, keras, scipy and Apache Solr. Language resources I was developing were in all available languages in the app (Arabic, German, English, French, Hebrew, Italian, Japanese, Dutch, Polish, Portuguese, Romanian and Russian)
- 2013–2013 **Full-stack internship developer**, *Prompsit Language Engineering*, Elche.  
 Developing news sentiment analysis project. Python and bash for crawling and cleaning natural language from all downloaded HTML data; AJAX, jQuery, CSSv3 and PHP for a real-time refreshed interface for dataset development; MySQL for dataset storage.
- 2012-2012 **Web development**, *Grupo Kordino*, Santa Pola.  
 Creation and support of a blog website for the hiking group named Grupo Kordino

## Papers

- Title *ParaCrawl: Web-Scale Acquisition of Parallel Corpora*
- Authors Marta Bañón, Pinzhen Chen, Barry Haddow, Kenneth Heafield, Hieu Hoang, Miquel Esplà-Gomis, Mikel L. Forcada, Amir Kamran, Faheem Kirefu, Philipp Koehn, Sergio Ortiz Rojas, Leopoldo Pla Sempere, Gema Ramírez-Sánchez, Elsa Sarrías, Marek Strelec, Brian Thompson, William Waites, Dion Wiggins, and Jaume Zaragoza.
- Info ACL, Online, 6 July, 2020.

## Languages

- English **Proficient**  
 Spanish **Bilingual**  
 Catalan **Bilingual**

## Computer skills

- |             |                             |        |                      |
|-------------|-----------------------------|--------|----------------------|
| Prog. Lang. | C++, Java, Golang           | Vision | ROS, OpenGL, Blender |
| Scripting   | Python, Lisp, sh, AWK, Perl | Server | Scala, SQL, Solr     |

✉ [b64encoded: bGVvcG9sZG9wbGFzZW1wZXJlQGdtYWIsLmNvbQ==](mailto:b64encoded:bGVvcG9sZG9wbGFzZW1wZXJlQGdtYWIsLmNvbQ==@lpla.github.io)  
 📄 [lpla.github.io](https://lpla.github.io)

ML/NLP Tensorflow, Moses, Marian

Web/Cloud CSS, JavaScript, AWS/Azure

---

## Extracurricular activities and interests

### Music

Experience in several youth orchestras, symphonic wind orchestras and rock, swing and jazz bands, writing arrangements, improvising, while playing clarinet and piano as soloist. Non-professional experience with saxophone, guitar, ukulele, bass and drums.

2012–present **Principal clarinet**, *Banda Sociedad Unión Musical de Crevillent*, Crevillent.

2016–2019 **Bass clarinet**, *Banda Jove de la FSMCV*, Valencia.

Conducted by Pablo Marqués, Pere Vicalet and Saül Gómez. Rehearsed with Cristóbal Soler, Roberto Turlo and Santiago Pérez. Performed in ADDA (Alicante), Palau de la Música (Valencia) and honor guest at the CIBM (Valencia)

2014–2014 **Clarinet teacher**, *SUM Crevillent*, Crevillent.

Volunteering clarinet classes in summer camp of "Sociedad Unión Musical de Crevillent"

2006–2012 **Principal clarinet**, *Banda Unión Musical de Santa Pola*, Santa Pola.

Remarkable performance at the Palau de la Música (Valencia)

2011–2012 **Principal clarinet**, *Orquesta de Jóvenes de la Provincia de Alicante*, Alicante.

Remarkable performance at Hamburg "Das Schleswig-Holstein Musik Festival" (SHMF)

### Retro-videogames

Electronic repairs, modding and programming on old video-game platforms. Winner of the "best sound" award at "Retroconsolas Alicante 2013" for my Amstrad CPC 464 entry.

### Hardware

Open source 3D design (Autodesk), modeling and printing. Volunteer with "3D Makers Elche" during COVID-19 lockdown by printing more than two hundred face-masks for near hospitals.

Open source PCB layout design (EAGLE, KiCad) and production (OSHPark, JL-CPCB) for Game Boy devices

### Sports

In my spare time, I like running, swimming, enjoying the mountain on my bike and the coast with my long-board and roller skates.

---

## References

Available on request

✉ `b64encoded: bGVvcG9sZG9wbGFzZW1wZXJlQGdtYWIsLmNvbQ==`  
📄 [lpla.github.io](https://lpla.github.io)