

# JOAN SERRÀ

## Lead Research Scientist in Machine Learning

@joansj@gmail.com    +34 661 658 325    Barcelona, Spain  
serrjoa.github.io    linkedin.com/in/joan-serra-julia    github.com/joansj  
scholar.google.com/citations?user=sZLj96sAAAAJ    twitter.com/serrjoa



## EXPERIENCE

### Head of Applied AI – Machine Learning Research

#### Dolby Laboratories

November 2019 – Current    Barcelona, Spain

- Leading a multi-geo team of 5–9 people. We do research on advanced ML topics, with a focus on audio and multi-modal analysis/synthesis.
- As individual contributor (formerly also Senior Staff Researcher), I work on generative modeling, representation learning, and retrieval.

### Senior Scientific Researcher – Machine Learning

#### Telefónica R&D

March 2015 – October 2019    Barcelona, Spain

- Led the machine learning efforts for internal projects on anomaly detection, recommender systems, and audio/speech processing.
- Also worked on time series classification, mobility estimation, human-computer interaction, and network analytics.

### Postdoctoral Researcher – Artificial Intelligence

#### IIIA-CSIC, Spanish National Research Council

October 2011 – February 2015    Bellaterra, Spain

- Proposed novel evolutionary computation algorithms for large-scale time series mining and retrieval (particle swarms, genetic algorithms).
- Developed ML tools for healthcare applications. Also worked on automatic music segmentation and computational audio analysis.

### PhD in Information and Communication Technologies

#### Music Technology Group, Universitat Pompeu Fabra

October 2007 – September 2011    Barcelona, Spain

- Thesis: “Identification of versions of the same musical composition by processing audio descriptions”. European Doctorate mention.
- Also worked on music mood estimation, general audio classification, and nonlinear time series analysis.

## FURTHER INFORMATION

**Research stays** — Max Planck Institute for Computer Science (2012) and Max Planck Institute for the Physics of Complex Systems (2010).

**Academia** — ELLIS member (2022), part-time associate professor at UVic (2018–2020), tenure-track lecturer accreditation (2014), and assistant professor at UPF (2007–2011).

**Student supervision** — A total of 34 students (including undergraduate, MSc, PhD, and company interns).

**Talks** — I regularly give invited talks and seminars, lately basically related to deep learning and generative models. I've been an active promoter of the Barcelona deep learning community (e.g., DLBCN or bcn.ai).

**Other** — In the past, I was a professional musician, playing solo electric guitar in multiple bands (1996–2012).

## SCIENTIFIC OUTPUT

Publications	133	Citations	6,803
Patents	17	h-index	38
EU(+) projects	13	i10-index	82

NeurIPS	ICLR	ICML	RecSys	IS
ICASSP	WSDM	ACM-MM	AAAI	

## SKILLS

Generative Modeling	●●●●●●
Representation Learning	●●●●●●
Audio Processing	●●●●●●
Time Series	●●●●●●
Statistics	●●●●●●
Algorithms	●●●●●●
Evolutionary Computation	●●●●●●
Information Retrieval	●●●●●●
Complex Systems	●●●●●●

## TOOLS

Python	Numpy/Scipy	Tensorflow
PyTorch	Sklearn	C/C++
Matplotlib	LaTeX	Git
		Assembler

## RECOGNITIONS

- Top 1% performer (11/1528)**  
Kaggle AXA Driver Telematics Challenge (2015)
- First ranked (1/58)**  
Juan de la Cierva - Incorporación postdoctoral fellowship (2015)
- Best algorithm awards (total: 6)**  
MIREX (2008–2012)
- Knowledge transfer award**  
Universitat Pompeu Fabra (2011)

## STRENGTHS

Hard-worker	Fast learner	Creative
Organized	Self-motivated	Proactive
Communication	Team management	

More detail can be found at the links above. My personal web page includes a full list of publications and projects.

# SELECTED PUBLICATIONS

---

## Audio, Speech, Music

- E. Postolache, J. Pons, S. Pascual, & J. Serrà. Adversarial permutation invariant training for universal sound separation. *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, in press. Jun 2023.
- J. Serrà, S. Pascual, J. Pons, R.O. Araz, & D. Scaini. Universal speech enhancement with score-based diffusion. *ArXiv*, 2206.03065. Jun 2022.
- J. Serrà, S. Pascual, & C. Segura. Blow: a single-scale hyperconditioned flow for non-parallel raw-audio voice conversion. *Advances in Neural Information Processing Systems (NeurIPS)*, 32: 6790-6800. Dec 2019.
- S. Pascual, M. Ravanelli, J. Serrà, A. Bonafonte, & Y. Bengio. Learning problem-agnostic speech representations from multiple self-supervised tasks. *Proc. of the Conf. of the Int. Speech Communication Assoc. (INTERSPEECH)*, 1791-1795. Sep 2019.
- S. Pascual, A. Bonafonte, & J. Serrà. SEGAN: speech enhancement generative adversarial network. *Proc. of the Conf. of the Int. Speech Communication Assoc. (INTERSPEECH)*, 3642-3646. Aug 2017.
- J. Serrà, M. Müller, P. Grosche, & J.L. Arcos. Unsupervised detection of music boundaries by time series structure features. *Proc. of the AAAI Int. Conf. on Artificial Intelligence (AAAI)*, 1613-1619. Jul 2012.
- J. Serrà, A. Corral, M. Boguñá, M. Haro & J.L. Arcos. Measuring the evolution of contemporary western popular music. *Nature Scientific Reports*, 2: 521. Jul 2012.

## Machine Learning, Recommenders, Other

- F. Yesiler, M. Miron, J. Serrà & E. Gómez. Assessing algorithmic biases for musical version identification. *ACM Int. Conf. on Web Search and Data Mining (WSDM)*, 1284-1290. Feb 2022.
- J. Serrà, D. Álvarez, V. Gómez, O. Slizovskaia, J.F. Núñez, & J. Luque. Input complexity and out-of-distribution detection with likelihood-based generative models. *Proc. of the Int. Conf. on Learning Representations (ICLR)*, Apr 2020.
- J. Serrà, D. Surís, M. Miron, & A. Karatzoglou. Overcoming catastrophic forgetting with hard attention to the task. *Proc. of the Int. Conf. on Machine Learning (ICML)*, 80: 4555-4564. Jul 2018.
- J. Serrà & A. Karatzoglou. Getting deep recommenders fit: Bloom embeddings for sparse binary input/output networks. *Proc. of the ACM Conf. on Recommender Systems (RecSys)*, 279-287. Aug 2017.
- J. Serrà, I. Leontiadis, A. Karatzoglou, & K. Papagiannaki. Hot or not? Forecasting cellular network hot spots using sector performance indicators. *Proc. of the IEEE Int. Conf. on Data Engineering (ICDE)*, 259-270. Apr 2017.
- J. Serrà & J.L. Arcos. Particle swarm optimization for time series motif discovery. *Knowledge-Based Systems*, 92: 127-137. Jan 2016.
- F. Font, J. Serrà, & X. Serra. Analysis of the impact of a tag recommendation system in a real-world folksonomy. *ACM Trans. on Intelligent Systems and Technology*, 7(1): 6. Oct 2015.

# SELECTED EU/NATIONAL PROJECTS

---

- BISON (2015-2017):** Big speech data analytics for contact centers. European Commission: ICT-2014-15-645323.
- BuscaMedia (2010-2011):** Automatic generation of audiovisual narrative. Spanish Government, Ministry of Science and Innovation: CENIT-2009-1026.
- DRIMS (2009-2012):** Description and retrieval of music and sound information. Spanish Government, Ministry of Science and Innovation: TIN-2009-14247-C02-01.
- PHAROS (2007-2009):** Platform for search of audiovisual resources across on-line spaces. European Commission: IST-2006-045035.
- SALERO (2006-2009):** Semantic audiovisual entertainment reusable objects. European Commission: IST-2007-0309BSCW.

# REFERENCES

---

- Vivek Kumar (vivekk.chugh@gmail.com). Senior Manager at Google Research, Mountain View, USA. Formerly, Director of Applied AI at Dolby Laboratories.
- Alexandros Karatzoglou (alexandros.karatzoglou@gmail.com). Research Scientist at Google Deepmind, Zurich, Switzerland. Formerly, Research Director at Telefonica R&D.
- Josep Lluís Arcos (jlarcos.cat@gmail.com). Chief Scientific Officer at MJN-Neuro, Blanes, Spain. Formerly, Head of the Learning Systems Dept. at IIIA-CISC.
- Xavier Serra (xavier.serra@upf.edu). Director of the Music Technology Group, Universitat Pompeu Fabra, Barcelona, Spain.