Sacha Krstulović Dipl. Eng., PhD, IEEE Senior Member

Head of Technology and Innovation
Machine Learning / Artificial Intelligence for Audio and Speech

+44 (0)7903 828017 | dr.sacha.k@gmail.com | https://sacha.today | LinkedIn | Cambridge, UK

Key skills

I build and drive high-performing and happy innovation teams. I help you to understand, strategise and execute innovation in the AI / machine learning space, with a specialism in audio, speech and music processing. I help you to build your intellectual property policy, portfolio and ROI. I help you to understand and formalise creative processes.

Areas

Machine Learning, Artificial Intelligence, Music Information Retrieval, Automatic Sound Event/Scene Recognition, Automatic Speech/Speaker Recognition, Text-To-Speech Synthesis, Automatic Analysis of Digital Audio Signals.

Methods

Management of technology and innovation, R&D, team leadership, coaching, patenting and intellectual property, transfer of technology from research to products, fundamental research, software engineering, scrum at scale, lean startup, communication, teaching/pedagogy, theory/history/philosophy of creativity.

Professional experience

Head of Artificial Intelligence Research   Jan 2022 – May 2023

Music Tribe   AI for music
Manchester, UK

Digital transformation in the multi-million dollar audio equipment industry: I drove R&D on AI, Machine Learning and advanced Digital Signal Processing; redefined the research process and defined+resourced the research roadmap; doubled the size of the AI/ML/DSP research team from 7 to 14 people, reinforced key R&D support functions, developed key academic partnerships.
3 proofs of concept from my team's research portfolio achieved product integration agreements at the end of 2022. All R&D activities were shut down in 2023Q2 due to a change of company management and strategy.

Innovation leadership positions at   Nov 2012 – Dec 2021

Audio Analytic   Machine Learning for Sound Recognition
Cambridge, UK   Computer Hearing as a Service

Our startup Audio Analytic got acquired by Meta in 2022. The company's ai3™ technology gives consumer products and digital assistants a sense of hearing and the ability to react to the world of sounds around us, thus helping to satisfy our entertainment, safety, security, wellbeing and communication needs.

Director of Innovation   Dec 2016 – Dec 2021

I increased company value through innovation:
- directing in-house R&D of new sound recognition algorithms, techniques and products,
- patenting, managing the company's intellectual property strategy,
- technical evangelist in conferences, academic publications and online communications (e.g. blog posts),
- building academic partnerships, influencing the global sound recognition research community,
The company raised $12m of Series B VC funding in June 2019, for a total of $25m, and was acquired by Meta in 2022.
The technology whose development I oversaw is deployed onto millions of consumer devices.
I have recruited 15 people, 8 of which to build the R&D team, while the company grew to 50+ people.
15 patents, 1 book chapter, 2 conference articles.

V.P. of Technology   Nov 2014 – Dec 2016

I drove the whole technology chain to meet business demands:
- executing in-house R&D of machine learning algorithms and techniques for sound recognition,
- surveying state of the art algorithms, liaising with academic researchers,
- managing big data for machine learning; developing and evaluating the technology over real life problems,
- delivering the technology to customers in the form of efficient embedded and platform-independent software.
The company raised $5.5m of Series A VC funding in December 2016.
I grew the R&D team to 5 people while the company grew to 20 people.
1 journal article, 1 conference publication, 3 patents.

Lead R&D Engineer   Nov 2012 – Nov 2014

I managed core sound recognition technology on both the short term, by delivering solutions to customers, and the long term, by improving the core technology and developing in-house R&D:
- redesigned the core sound recognition engine, tools and algorithm, "made it work",
- developed the recognition of sound events (e.g. smoke alarm, dog bark...) for professional and consumer applications,
- managed audio data collection campaigns,
- integrated and tested for various customers on various platforms, incl. embedded ARM, PCs, Android apps etc.
- drove academic partnerships.
I grew the R&D team to 3 people while the startup grew from 3 to 7 people.
The company raised $1m of Angel funding in 2013.
I also won two collaborative InnovateUK grants for the company, worth a total of £380k, which supported a partnership with Prof. Mark Plumbley of Queen Mary University London (now at University of Surrey).

Senior Research Engineer - Advanced Speech Group   Sep 2011 – Nov 2012

Nuance
Cambridge, UK   **Automatic Speech Recognition**

Automatic transcription of voice mail messages (VM2T: VoiceMail-To-Text).
Successfully delivered on a variety of short R&D projects: unsupervised model training; active learning; Vocal Tract Length Normalisation (VTLN); privacy-aware training (data retention vs voicemail privacy).

Research Engineer - Speech Technology Group   Apr 2007 – Sep 2011

Toshiba Research Europe
Cambridge, UK   **Text-To-Speech Synthesis**

Paradigm-shifted Toshiba's Text-To-Speech (TTS) technology from unit selection to HMM-based speech synthesis.
My prototype achieved 98% preference over competitors' commercial TTS software.
Transferred the new technology from research to production, implemented additional software toolkits to produce synthetic voices for European and North American languages.
Managed two contractors to accelerate prototype development.
3 conference publications, 1 journal article, 1 patent.

Researcher - Language Technology Team   Jun 2006 – Mar 2007

DFKI (Deutsches Forschungszentrum für Künstliche Intelligenz)
Saarbrücken, Germany   **Text-To-Speech Synthesis**

Researched and developed parametric speech synthesis into the MARY open-source TTS platform.
First ever high-quality TTS synthesis system for the German language based on HMMs.
Improved the expressiveness of Text-To-Speech Synthesis through the parametrisation of prosody and voice quality, first ever expressive synthesis for German.
2 conference publications, 1 book chapter.

Research Eng. - METISS Team (Audio Signals Processing)   Mar 2004 – May 2006

IRISA (Institut de Recherches en Informatique et Systèmes Aléatoires)
Rennes, France   **Automatic Speaker Verification**
                                                    Speech & Audio Scene Analysis

Fast version of Matching Pursuit for multi-resolution analysis of digital signals; reduced computation time from 240x to 1x real time, thus making the algorithm applicable to real life signals. 5 conference publications.
NEOLOGOS Project: optimising the phonetic coverage of a speech recognition/synthesis database, innovative approach drawing from state-of-the-art speaker verification techniques. 2 conference publications, 1 journal article, 1 book chapter.
NIST Evaluations 2005: speaker recognition eval, 27 participating sites. Implemented and submitted IRISA/METISS's speaker verification system, distilling a GMM into a decision tree. 20x faster and smaller than competing systems.
Outreach/Teaching/Supervision: demonstrations for science fairs & opening days, internal training of researchers through tutorials in summer schools, supervision of Masters student projects.

World Tour   Feb 2002 – Nov 2003
Africa North to South, sailing from Cape Town to Tahiti via Panama, then South America, USA, back to Europe via the TransSiberian train and the Baltic states.

Research Assistant / PhD Student   **Jun 1996 – Dec 2001**

IDIAP (Institut Dalle Molle d'Intelligence Artificielle Perceptive)
Martigny, Switzerland   **Speech Modelling & Speech Recognition**

Research on parametric analysis of speech under articulatory constraints, evaluation on medium vocabulary continuous speech recognition. Parametric signal analysis library, comprising Auto-Regressive modelling and articulatory modelling, developed in C language. PhD thesis, 5 publications.
Teaching assistant, 75 hours delivered to 28 engineering students, 20 pre-doctorate students and 15 post-grad students.

Publications

My 57 publications are available on [Google Scholar](#).

My 18 patents are available on [Google Patents](#).

Studies and professional training

Innovation Management Certificate of Achievement   **Jul 2021**

The Cambridge Creativity Lab   **Jul 2021**
Stimulate personal and collective creativity; importance and relevance of creativity within a business context.

Driving Entrepreneurial Capability: How to Harness Innovation and Agility   **Nov 2020**
Greater levels of innovation and agility: challenge assumptions, identify opportunities, generate alternative approaches.

Strategic Management: Creating and Sustaining Competitive Advantage   **Mar 2019**
Identify opportunities, find new ways to create value for customers and move ahead of competitors. Strategic analysis.

Managing Innovation Strategically   **Nov 2018**
Implement innovation through a set of practical concepts, tools and methods to create and capture new value.

[Cambridge Judge Business School](#)   **Cambridge, UK**

Undergrad Diploma in Creativity (Theory, History & Philosophy)   **Sep 2020**
Portfolio of essays focusing on the management of creativity in entrepreneurship.

[Cambridge University – Institute of Continuing Education](#)   **Cambridge, UK**

Open Course in Technology and Innovation Management   **Apr 2016**
Key tools and techniques for managing and exploiting technological investments and opportunities.

[Cambridge University's Institute for Manufacturing](#)   **Cambridge, UK**

PhD in Technical Sciences   **Dec 2001**
Parametric analysis of speech under articulatory constraints, research within the ARTIST project at IDIAP (see above).

[EPFL](#)   **Lausanne, Switzerland**

Masters of Engineering   **Jun 1996**
Aircraft branch with specialisation in Digital Signal Processing and Machine Learning (Neural Networks).

[ESTACA](#)   **Levallois-Perret, France**

Summer School - Information Transfer Technologies   **Jul – Aug 1995**
Human vocal tract modelling project under MATLAB/Simulink.

[Harvard University](#)   **Cambridge, Massachusetts, USA**

Baccalauréat (A-levels) – Type C: Sciences   **1990**
Lycée du Parc des Loges   **Évry, France**

Tools and systems

Machine Learning	Deep Learning, Keras, Tensorflow, SciKit-Learn, Torch, PyTorch
Speech and audio tech	HTK, Attila, AudioSeg, Alizé, HTS, OpenMARY, Festival, ToSpeak, Librosa, SPTK, EST, ESPS, Praat, Wavesurfer, Audacity, MPTK
Systems	UNIX/Linux, MacOSX, MS Windows, Android, ARM, FreeRTOS, Xenomai
Programming, scientific computing, cloud	Python, C, C++, Java, Perl, Flex, FFTW, BLAS/LAPACK, MATLAB, PyLab, Jupyter, NumPy/SciPy, matplotlib, pandas, Eclipse, SVN, Git/Gerrit, Automake, gcc-arm, Sungrid Engine, AWS, Azure
Project management	MSDevOps, Atlassian JIRA/Agile, Confluence

Miscellaneous

Languages	French (native), English (fluent), Spanish (conversational), Italian (basic), German (basic)
Activities	Zen meditation and mindfulness.
	Cuban Salsa teacher: running my own dance classes and festival events since 2008.
	Running the Cambridge Creativity Meetup: https://www.meetup.com/Cambridge-Creativity/
	Taking care, educating and having fun with my son born in September 2021.