# Michael Penkov

[michael.penkov@gmail.com](mailto:michael.penkov@gmail.com)

Data Scientist with a strong professional Software Engineering experience.

## Education

2010-2015

**PhD, Information Science and Technology**; Hokkaido University (Sapporo, Japan)

*Thesis title: Estimating Video Authenticity via the Analysis of Video Quality and Video Structures*

2002-2006

**BEng, Software Engineering**; University of New South Wales (Sydney, Australia)

## Professional Experience

**Data Scientist at Profound Networks (2012 to present date)**

* Actively led a team to rapidly develop a system for gathering business intelligence via the Web.
* Implemented and enforced baseline code and documentation quality levels.
* Developed machine learning solutions for Web domain classification & postal address extraction.
* Developed an ElasticSearch-based internal search system.
* Received an “Outstanding Developer Achievement Award” for the above efforts.

**Software Engineer at Canon Information Systems Research, Australia (2005 to 2010)**

* Awarded for an effective demonstration of Canon technology at a conference.
* Authored 3 patents with the Australian patent office.

## Skills

Programming Languages

**Python:** my primary tool for development and data analysis. Closely familiar with the standard library, as well as **nltk**, **libsvm**, **requests** and others. I’m very passionate about Python, and frequently attend PyCon conferences around the world.

**R:** used to complement my Python skills.

Intermediate knowledge of **C**, **bash** and **other command-line tools**.

Open Source

I’ve contributed to [ffmpeg](https://github.com/mpenkov/ffmpeg-tutorial), [smart\_open](https://github.com/RaRe-Technologies/smart_open), [flake8](https://github.com/PyCQA/flake8), [pandas](https://github.com/pandas-dev/pandas) and [piexif](https://github.com/hMatoba/Piexif). See my [github profile](https://github.com/mpenkov) for a full list.

Languages

Native speaker of English and Russian

Business-level proficiency in Japanese (JLPT Level 1)