# YASSER GONZALEZ

Curriculum Vitae – September 2015

**Website** – http://yassergonzalez.com
**Email** – contact@yassergonzalez.com
**Phone** – 1-647-767-0942


## SUMMARY


I'm passionate about the design and implementation of algorithms to solve challenging problems and draw conclusions from data. My interests include the areas of optimization, machine learning, natural language processing and data analysis – many of which fall within the field of data science. I love working with open-source software, started 10 years ago and haven't stopped since then.


## EDUCATION

2013–2015 **Master's degree, Information Systems & Technology.**
York University, Canada.

- *Thesis:* Efficient Calculation of Optimal Configuration Processes.
- *Courses:* Advanced Topics in Information Technology – Mining of Massive Datasets, Advanced Information Retrieval Systems, Introduction to Computational Linguistics, Software Product Lines, Research Methods in Information Technology.

2014–2015 **Data Science Specialization.**
A non-credit series offered by Johns Hopkins University through Coursera.

- *Courses:* The Data Scientist's Toolbox, R Programming, Getting and Cleaning Data, Exploratory Data Analysis, Reproducible Research, Statistical Inference, Regression Models.

2011–2012 **Graduate Coursework, Mathematics.**
University of Havana, Cuba.

- *Courses:* Multivariate Statistics, Nonparametric Tests – Methods Based on Ranks, Linear Models, Stochastic Simulation, Linear & Integer Programming, Heuristic & Metaheuristics Algorithms, Introduction to Parallel Computing.

2006–2011 **Bachelor's degree, Computer Science.**
University of Havana, Cuba.

- *GPA:* 5.0/5.0 (Summa Cum Laude).
- *Thesis:* Estimation of Distribution Algorithms Based on Copulas and Vines.
- *Courses:* Algebra, Mathematical Analysis, Probability & Statistics, Discrete Mathematics, Theory of Programming Languages, Design & Analysis of Algorithms, Compiler Construction, Numerical Methods, Operating Systems, Database Systems, Software Engineering, Computer Networks, Artificial Intelligence, Information Retrieval Systems, among others.

## EMPLOYMENT

2013–2015 **Research & Teaching Assistant.**
York University, Canada.

- Designed different solution methods for the problem of optimizing the user interaction in a configuration process.
- Developed techniques to improve the performance of search heuristics on multimodal optimization problems.
- Wrote scientific software in Python (using NumPy, pandas, and SciPy), R, and MATLAB/Octave.

- Teaching assistant for the courses ITEC 1620 Object-Based Programming (four sessions), ITEC 2620 Introduction to Data Structures (one session), and ITEC 1000 Introduction to Information Technologies (one session).

2011–2013 **Research Assistant**. Institute of Cybernetics, Mathematics and Physics, Cuba.

- Developed new estimation of distribution algorithms (EDAs) using copulas and vines to model the probability distributions.
- Created a group of R packages available on CRAN, and a C library for dependence modeling using vines.
- Co-supervised a bachelor's thesis in Computer Science.

**TECHNOLOGIES** Python – R – C – Java – HTML – CSS – JavaScript – SQL – Hadoop
Git – LaTeX – GNU/Linux system administration.

**SELECTED OPEN-SOURCE SOFTWARE**

- **configurator** – Python package providing different solutions to the problem of optimizing the user interaction in a configuration process. https://github.com/yasserglez/configurator.
- **copulaedas** – R package for implementing and studying estimation of distribution algorithms (EDAs) based on copulas. https://github.com/yasserglez/copulaedas.
- **vines** – R implementation of the vine graphical model for building high-dimensional probability distributions as a factorization of bivariate copulas and marginal density functions. https://github.com/yasserglez/vines.
- **ngram_profile** – Python library for text classification based on character n-grams. https://github.com/yasserglez/ngram_profile.
- **dml** – C library for dependence modeling using C-vines, D-vines and R-vines. https://github.com/yasserglez/dml.

For more information, please see http://yassergonzalez.com/software.

**SELECTED PUBLICATIONS**

- Y. Gonzalez-Fernandez, S. Chen. (2015). Leaders and Followers – A New Metaheuristic to Avoid the Bias of Accumulated Information. In *IEEE Congress on Evolutionary Computation*, 776–783. IEEE. http://dx.doi.org/10.1109/CEC.2015.7256970.
- Y. Gonzalez-Fernandez, M. Soto. (2014). copulaedas: An R Package for Estimation of Distribution Algorithms Based on Copulas. *Journal of Statistical Software*, 58(9), 1–34. http://www.jstatsoft.org/v58/i09.
- Y. Gonzalez-Fernandez, S. Chen. (2014). Identifying and Exploiting the Scale of a Search Space in Particle Swarm Optimization. In *Conference on Genetic and Evolutionary Computation*, 17–24. ACM. http://doi.acm.org/10.1145/2576768.2598280.
- Y. Gonzalez-Fernandez, D. Carrera, M. Soto, A. Ochoa. (2012). Vine Estimation of Distribution Algorithms. In *VIII Congreso Español sobre Metaheurísticas, Algoritmos Evolutivos y Bioinspirados*, 1–7. http://simd.albacete.org/maeb2012/papers/paper_99.pdf.

For more information, please see http://yassergonzalez.com/publications.