Roque Enrique López Condori

Curriculum Vitae

▬ Personal Information

Name	Roque Enrique López Condori
Nationality	Peruvian
Address	Nice, France
Website	http://roquelopez.com/
G. Scholar	http://scholar.google.com/citations?user=eCBvlNMAAAAJ
LinkedIn	http://linkedin.com/in/roquelopez
Blog	http://medium.com/@roquelopez
Github	http://github.com/roquelopez

▬ Education

2013–2015	**MSc in Computer Science**, *São Paulo University*, São Paulo, Brazil.
2006–2010	**BSc in Computer Science**, *San Agustin National University*, Arequipa, Peru.
2001–2005	**High School Diploma**, *Colegio Técnico Parroquial San Martín de Porres*, Arequipa, Peru.

▬ Awards

- 2016, Finalist for the Best MSc Dissertation in IEEE Latin American Conference on Computational Intelligence, Colombia.
- 2014, Young Scientific Best Paper in VII International Conference on Intelligent Information and Engineering Systems, Poland.
- 2014, Master Scholarship granted by Samsung Brazil (2014-2015).
- 2013, Top 5 undergraduate thesis in V Simposio Peruano de Inteligencia Artificial, Peru.
- 2010, Placed in top 5 in the undergraduate, Peru.
- 2005, First place and scholarship High School Diploma, Peru.

▬ Experience

Industrial Work

Feb 2017 present	**Research Engineer**, *Inria*, Sophia Antipolis. This project will enable robots to tap into the ever-growing amount of knowledge available on the Web, by learning from there about the meaning of previously unseen objects, expressed in a form that makes them applicable when acting in situated environments.

Mar 2016 present	**CEO, Co-Founder**, *SimiLabs*. Peruvian NLP startup (`http://simi-labs.com/`) dedicated to create tools and resources in order to analyze opinions on the Internet about companies, products and/or services.
July 2015 Jan 2017	**Researcher, Software Engineer**, *Elabora Consultoria*, São Paulo. I researched some information retrieval methods and word embedding representations to create hierarchies of terms in collections of patent about different topics. In this project, I used Spark, HBase and Python.
Oct 2013 Dec 2013	**Software Engineer**, *Dicionário Criativo*, São Paulo. I implemented a crawler for Twitter in Python. This crawler collects public comments about music, books and films. I also developed a module to normalize these comments (spell checker, etc.). With this data and other resources the database was increased.
Jan 2012 Dec 2012	**Researcher, Software Engineer**, *Lindexa - Wayra Telefónica*, Lima. I researched and developed some automatic techniques of Opinion Mining and Sentiment Analysis for Spanish language. In addition, I implemented a crawler for Facebook and Twitter in Python. This crawler collects all public comments written on these social networks. Lindexa is a winner startup of the first edition of Wayra Peru.
Jan 2011 Dec 2011	**Technical Researcher, Developer**, *Cátedra CONCYTEC*, Arequipa. In general, we researched techniques to generate extractive summaries of many texts. These summaries represent the most important data. I also researched and implemented Text Classification and Clustering algorithms.
Sept 2010 Dec 2010	**Interface Designer**, *Software Solutions S.A.C.*, Arequipa. We developed the Graphical User Interfaces of the Credits and Collections module, it includes the elaboration of the navigability documents and prototypes. I also developed the executable prototype.
Sept 2009 Dec 2009	**Analyst, Developer**, *EPISUNSA Development Centre*, Arequipa. Together with a group of colleagues, we worked on the design and development of the website of the Professional School of Systems Engineering.

Academic Work

Feb 2014 July 2014	**Teaching Assistant**, *Course: Professional Information in Computer Science*, with Dr. Thiago Pardo at São Paulo University.
June 2014 Dec 2014	**Associate Editor**, *CompuScientia*, Peruvian Electronic Journal of the Students Society of Computer Science.

Research Projects

Feb 2017 present	***ALOOF: Autonomous Learning of the Meaning of Objects***, *Inria*, Sophia Antipolis, France. The goal of ALOOF is to significantly advance the ability of today's autonomous systems to adapt to ever changing, dynamic real world environments by enabling them to learn about the meaning of objects from resources accessible through the Web.
July 2015 Feb 2017	***Patent Faceted Searches Using Topic Modeling***, *Elabora Consultoria*, São Paulo, Brazil. This project aimed to integrate advanced techniques of information retrieval for making faceted searches over big collections of patents about different topics.

Jan 2014 June 2015	***Semantic Text Processing in Brazilian Portuguese***, *Samsung – ICMC/USP*, São Paulo, Brazil. The overall objective of this project was to advance the state of the art in semantic processing of texts written in Brazilian Portuguese. With these resources, we built mining and summarization texts applications, focused on Internet opinions.
Jan 2011 Dec 2011	***Automatic Summarization of Medical Records using Natural Language Processing techniques***, *Cátedra CONCYTEC*, Arequipa, Peru. The goal of this project was to investigate and apply techniques to generate automatic summaries of many documents. We developed different methods to extract, select and organize the most important information.

Monographs

MSc Thesis	*Title:* Automatic Aspect-Based Opinion Summarization Methods *Description:* In this master's project are presented investigations to generate extractive and abstractive summaries of opinions using an aspect-based approach. Besides using known methods in the area, it also was proposed two new methods for Portuguese language which got the best performance in the experiments.
BSc Thesis	*Title:* Medical Documents Classification based on Keywords using Semantic Information *Description:* In this monograph, it is presented a method to classify medical documents which improves the results of Naive Bayes and Rocchio algorithm. This method, in addition to considering statistical information, taking into account the semantic relatedness between the keywords of medical documents.

Publications

Journal Articles

July 2017	**Roque López and Thiago Pardo**, *Opinion Summarization Methods: Comparing and Extending Extractive and Abstractive Approaches*, Expert Systems with Applications (ESWA), volume 78.

Conference Papers

Nov 2016	**Lucia Castro, Roque López, Gabriel Cavalcante and Luiz Lapolla**, *Towards Automatic Building of Term Hierarchies from Large Patent Datasets*, In the Proceedings of the V Artificial Intelligence and Natural Language Conference (AINL), Saint-Petersburg, Russia.
Apr 2015	**Roque López and Thiago Pardo**, *Experiments on Sentence Boundary Detection in User-Generated Web Content*, In the Proceedings of the XVI International Conference on Intelligent Text Processing and Computational Linguistics (CICLING), Cairo, Egypt.
Oct 2014	**Verônica Agostini, Roque López and Thiago Pardo**, *Automatic Alignment of News Texts and their Multi-document Summaries: Comparison among Methods*, In the Proceedings of the XI International Conference on Computational Processing of Portuguese (PROPOR), São Carlos, Brazil.
Oct 2014	**Roque López, Javier Tejada and Mikhail Alexandrov**, *Medical Texts Classification based on Keywords using Semantic Information*, In the Proceedings of the VII International Conference on Intelligent Information and Engineering Systems (INFOS), Rzeszow-Krynica, Poland.

Sept 2012 **Roque E. López, Javier Tejada and Mike Thelwall**, *Spanish Sentistrength as a Tool for Opinion Mining Peruvian Facebook and Twitter*, In the Proceedings of the V International Conference on Intelligent Information and Engineering Systems (INFOS), Rzeszow-Krynica, Poland.

Sept 2012 **Angels Catena, Mikhail Alexandrov and Roque E. López**, *Parameterization of comments from Peruvian Facebook and Twitter: Lexical Resources and Algorithm*, In the Proceedings of the V International Conference on Intelligent Information and Engineering Systems (INFOS), Rzeszow-Krynica, Poland.

Oct 2012 **Alessandro Bokan, Roque E. López**, *Método No Supervisado para la sugerencia de Tags utilizando información semántica basada en conocimiento*, In the Proceedings of the IV Congreso Internacional de Computación y Telecomunicaciones (COMTEL), Lima, Peru.

Nov 2011 **Roque E. López, Dennis Barreda and Javier Tejada**, *MFSRank: An Unsupervised Method to Extract Keyphrases Using Semantic Information*, In the Proceedings of X Mexican International Conference on Artificial Intelligence (MICAI). Lecture Notes in Artificial Intelligence. Springer, Puebla, Mexico.

Nov 2011 **Roque E. López, Dennis Barreda, Javier Tejada and Luis Alfaro**, *Método supervisado orientado a la clasificación automática de Historias Clínicas*, In the Proceedings of the XXIII Encuentro Chileno de Computación (ECC), Curicó, Chile.

Sept 2011 **Roque E. López, Mikhail Alexandrov and Javier Tejada**, *LexisTerm - The program for term selection by the criterion of specificity*, In the Proceedings of the IV International Conference on Intelligent Information and Engineering Systems (INFOS), Polanczyk, Poland.

Sept 2011 **Ales Bourek, Mikhail Alexandrov and Roque E. López**, *Folksonomy - supplementing RICHE expert based taxonomy by terms from online documents*, In the Proceedings of the IV International Conference on Intelligent Information and Engineering Systems (INFOS), Polanczyk, Poland.

Aug 2011 **Roque E. López, Dennis Barreda, Javier Tejada and Luis Alfaro**, *Clasificación automática de Historias Clínicas basada en Prototipos utilizando técnicas de Procesamiento de Lenguaje Natural*, In the Proceedings of the X Jornadas Peruanas de Computación (JPC), Pucallpa, Peru.

Aug 2011 **Dennis Barreda, Roque E. López, Javier Tejada and Luis Alfaro**, *Un algoritmo genético para la agrupación de documentos aplicado en corpus con características diferentes*, In the Proceedings of the X Jornadas Peruanas de Computación (JPC), Pucallpa, Peru.

Workshop Papers

June 2015 **Roque López, Lucas Avanço, Pedro Balage, Alessandro Bokan, Paula Cardoso, Márcio Dias, Fernando Nóbrega, Marco Sobrevilla, Jackson Souza, Andressa Zacarias, Ariani Di Felippo, Eloize Seno and Thiago Pardo**, *A Qualitative Analysis of a Corpus of Opinion Summaries based on Aspects*, In the Proceedings of the IX Linguistic Annotation Workshop (LAW), co-located with NAACL, Colorado, USA.

Oct 2014	**Márcio Dias, Alessandro Bokan, Carla Chuman, Cláudia Barros, Erick Maziero, Fernando Nobrega, Jackson Souza, Marco Sobrevilla, Marina Delege, Lucía Castro, Naira Silva, Paula Cardoso, Pedro Balage, Roque López, Vanessa Marcasso, Ariani Felippo, Maria Graças, Thiago Pardo**, *Enriquecendo o Corpus CSTNews - a Criação de Novos Sumários Multidocumento*, In the Proceedings of the I Workshop on Tools and Resources for Automatically Processing Portuguese and Spanish (ToRPorEsp), São Carlos, Brazil.

Congress and Scientific Events

2016	**Assistant**, *III Latin American Conference on Computational Intelligence*, Cartagena - Colombia.
2015	**Assistant**, *XVI International Conference on Intelligent Text Processing and Computational Linguistics*, Cairo - Egypt.
2014	**Assistant**, *XI International Conference on Computational Processing of Portuguese*, São Carlos - Brazil.
2014	**Assistant**, *VII International Conference on Intelligent Information and Engineering Systemsn*, Krynica - Poland.
2013	**Assistant**, *V Simposio Peruano de Inteligencia Artificial*, Arequipa - Peru.
2011	**Assistant**, *X Mexican International Conference on Artificial Intelligence*, Puebla - México.
2011	**Assistant**, *X Jornadas Peruanas de Computación*, Pucallpa - Peru.
2010	**Assistant**, *III Simposio Peruano de Inteligencia Artificial 2010*, Arequipa - Peru.
2010	**Assistant**, *Congreso Internacional de Dirección de Proyectos 2010*, Arequipa - Peru.
2010	**Assistant**, *XVI Congreso Internacional INFOTEK 2010 y FETEC 2010*, Arequipa - Peru.
2010	**Assistant**, *Footprint techniques, Scanning, Sniffing, Vulnerability Analysis and Countermeasures*, Arequipa - Peru.

Languages

Spanish	Fluent	*native language*
English	Advanced Level	*completed studies*
Portuguese	Advanced Level	*completed studies*

Programming Knowledge

Programming Languages	- Python - Java - C++
Big Data	- Hadoop - Spark - HBase
Databases	- MySQL - PostGresSQL - SQL Server

Others - Git
- NLTK
- Scikit-learn
- NumPy
- Linux

Main Interests

- Opinion Summarization
- Sentiment Analysis and Opinion Mining
- Natural Language Generation
- Natural Language Processing
- Machine Learning
- Data Science
- Software Development