Ruth García-Gavilanes

1 St. Giles
Oxford OX1 3JS
United Kingdom
✆ +44 7494176970
✉ ruthygarciag@gmail.com
⌂ www.ruthygarcia.com

Research Interests

I am interested in understanding human behaviour from online footprints. To achieve this, I examine big data sets using methods such as web analytics, statistical analysis, machine learning techniques and complex network analysis. Some of the tools that I use to get meaningful results from big data are: Hadoop, Pig, Python, Java, R and Matlab. The online data that I have used so far was collected from Twitter, Google Trends, Wikipedia and Wikidata.

Research Experience

Since 06/15 — Postdoctoral Researcher.
Oxford Internet Institute (University of Oxford), United Kingdom

Collective memory and attention based on the information seeking patterns of a large number of individuals on the web (Google funding). Dynamics of disagreement in the interactions of online editors and bots.

Methods: Time-series analysis (segmented regression, ARIMA models), clustering, machine learning, visualisation. Tools: R, Python. Datasets: Wikipedia, everyday sexism site, google trends.

01/15–06/15 — Visiting Scientist.
Barcelona Media Foundation (now part of Eurocat), Spain

Collaborating in the European project SUPER on crisis event detection, prevention and response in social media.

Methods: Clustering, machine learning, visualisation. Tools: R, Python, Pig. Datasets: Twitter (API and Firehose), Wikipedia.

10/10–12/14 — Ph.D. Researcher.
Yahoo Labs Barcelona, Spain

Extensive experience in crawling, collection, parsing and processing of web-scale large datasets in distributed computing environments (e.g., MapReduce, Apache Hadoop) to study the role of culture in social media. Tools: Pig, R, Java, WEKA, Matlab. Datasets: Twitter Firehose.

01/12–05/12 — Research Associate.
Qatar Computing Research Institute, Qatar

SunFlower demo: a system that employs collaborative editing to summarise large corpora of literature articles. Tools: Java, R, Matlab. Datasets: Arxiv

01/10–10/10 — Master Researcher.
Telefónica R+D-Multimedia Group, Barcelona-Spain

Research on Recommender Systems: hybrid recommendation system of friends in Twitter. Tools: Java, Matlab. Dataset: Twitter

Teaching Experience

04/16–05/16 — Workshop in Basic Machine Learning, University of Oxford, United Kingdom.
Planning, preparing and delivering lessons on basic machine learning using WEKA (9 hours).

10/10–06/13 — Teaching Assistant for Information Retrieval, Universitat Pompeu Fabra, Spain.

Planning, preparing and delivering lessons of Information Technology to 2nd year undergraduate students. Supervising and monitoring pupils laboratory (3 years).

Education

10/10–02/15	**Ph.D. in Information, Communication and Audiovisual Media**.
	Universitat Pompeu Fabra, Barcelona, Spain
	Grade: Excellent with summa cum laude.
10/09–10/10	**Master in Information, Communication and Audiovisual Media**.
	Universitat Pompeu Fabra, Barcelona, Spain
	GPA 87.4% or 3.3
03/01–06/07	**Computer Engineering**.
	Escuela Politécnica del Ejército, Quito, Ecuador
	GPA 93.5% or 4.0

Awards and Honors

2015	Finalist of the Ph.D. Big Data Talent award of the Big Data Congress in Barcelona
2014	Travel/accom. grant: The 17th Computer Supported Cooperative Work and Social Computing
2013	Best paper award: The 5th International Conference on Social Informatics (SocInfo)
2013	Grace Hopper ACM Scholarship recipient
2009	Fundación Carolina Master Scholarship recipient

Technical Skills

Data Mining Skills	Web Analytics, Web Mining [excellent]
	Machine Learning, Natural Language Processing, Regression Models [good]
Data Storage and Access	MySQL, SQLite, PostgreSQL, HDFS [excellent]
Data Analytics	Apache Pig, R, Java, Python, Eclipse, Shell scripts [excellent]
	WEKA, MapReduce, Gephi [good]
Visualization	ggplot2/R, OminiGraffle, Illustrator, Photoshop [excellent]
Additional	Perl, PHP, PBS, VB Script, JavaScript, HTML [excellent]
	Shell scripting C, C#, C++ [good]

Languages

- Spanish: **Mother tongue**
- English and Portuguese*: **Fluent**
- French** and Mandarin***: **Basic notion**

* tested by the Brazilian Ministry of Education (MEC)/TOEFL 2008

**B2 level by the Common European Framework of Reference for languages.

*** 1 year (2 levels) at Beijing Language and Culture University (BLCU)

Invited Talks/Oral Presentations

2016/06	Online Collective Memory, 2nd IC2S2, Evanston, IL, USA
2016/04	Online Collective Memory, FutureEverything, Manchester, United Kingdom
2016/03	Cultural Trails and Collective Memory in Online Data, Oxford Brookes University, Oxford, United Kingdom

Peer Reviewing

2017	Senior PC of the 11th International AAAI Conference on Weblogs and Social Media
2016	PC of the 2nd International Conference on Computational Social Science

2014/2016 PC of the 8th and 10th International AAAI Conference on Weblogs and Social Media

2014 PC of the 25th ACM Conference on Hypertext and Social Media

2014 PC of the 7th ASE International Conference on Social Computing

2013/2014 PC of the 5th and 6th International Conference on Social Informatics [honorable mention on the 6th]

Proposals for External Funding

2016/06 **A. Jungherr, J. An, R. García-Gavilanes** *Communicative Power in Hybrid Media Systems: The Relative Prominence and Evolution of Frames.* Submitted to Volkswagen Foundation (Line 2: International Cooperation of Postdoctoral Scholars)

Publications

Theses

[1] **R. García-Gavilanes** *Ph.D. thesis: User Behavior in Microblogs with a Cultural Emphasis.* (Universitat Pompeu Fabra, Department of Information and Communication Technologies, Feb 25, 2015. Barcelona-Spain)

[2] **R. García-Gavilanes** *Master thesis: Hybrid Methods for Recommending Connections in Online Social Networks.* (Universitat Pompeu Fabra, Department of Information and Communication Technologies, October, 2010. Barcelona-Spain)

[3] **R. García-Gavilanes** *B.S. thesis: Development of Multimedia Education Software in Mathematics for the Fifth grade of Elementary School.* (Escuela Politécnica del Ejército, Department of Computer Engineering, July, 2007. Quito-Ecuador) Tools: Flash and XML

Conference Papers and Journals

[1] **M. Tsvetkova, R. García-Gavilanes, T. Yasseri** *Dynamics of Disagreement: Large Scale Temporal Network Analysis Reveals Negative Interactions in Online Collaboration.* (Scientific Reports 6, Article number: 36333, November 3, 2016,) (DOI:http://dx.doi.org/10.1038/srep36333)

[2] **R. García-Gavilanes, M. Tsvetkova, T. Yasseri,** *Dynamics and Biases of Online Attention: The Case of Aircraft Crashes.* (The Royal Society Open Science, October 12, 2016.) (DOI: http://dx.doi.org/10.1098/rsos.160460) [Featured on the Guardian, Science and MIT Technology Review]

[3] **M. Tsvetkova, R. García-Gavilanes, L. Floridi, T. Yasseri** *Even Good Bots Fight.* (under review)

[4] **R. García-Gavilanes, A. Mollgaard, M. Tsvetkova, T. Yasseri,** *Memory Remains: Understanding Collective Memory in the Digital Age.* (under review)

[5] **R. García-Gavilanes**, D. Gómes, D. Parra, C. Trattner, A. Kaltenbrunner, E. Graells-Garrido. *Language, Twitter and Academic Conferences.* The 26th ACM HyperText Conference, 2015 September 01-04, Guzelyurt, TRNC, Cyprus

[6] **R. García-Gavilanes**, A. Kaltenbrunner, D. Saez-Trumper, R. Baeza-Yates, P. Aragon, D. Laniado. *Who are my Audiences? A Study of the Evolution of Target Audiences in Microblogs.* The 6th International Conference on Social Informatics (SocInfo), 2014 November 10-13, Barcelona, Spain

[7] **R. García-Gavilanes**, Y. Mejova, D. Quercia. *Twitter ain't Without Frontiers: Economic, Social, and Cultural Boundaries in International Communication.* The 17th ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW), 2014 February 15-19, Baltimore, Maryland, USA

[8] **R. García-Gavilanes**, N. O-Hare, L. Aillelo, A. Jaimes. *Follow My Friends This Friday! An Analysis of Human-generated Friendship Recommendations.* The 5th International Conference on Social Informatics (SocInfo), 2013 November 25-27, Kyoto, Japan [**Best paper award**]

[9] **R. García-Gavilanes**, D. Quercia, A. Jaimes. *Cultural Dimensions in Twitter: Time, Individualism and Power.* The 7th International AAAI Conference on Weblogs and Social (ICWSM), 2013 - Boston, USA [**Honored mention**]

[10] MC. Marcos, **R. García-Gavilanes**, E. Bataineh, L. Pasarín. *Using Eye Tracking to Identify Cultural Differences in Information Seeking Behavior* Many People, Many Eyes 2013 (ACM CHI workshop) – Paris, France

[11] B. Poblete, **R. García-Gavilanes**, M. Mendoza, A. Jaimes. *Do All Birds Tweet the Same? Characterizing Twitter Around the World* ACM International Conference of Information and Knowledge Management 2011 - Glasgow,UK

[12] **R. García-Gavilanes**, B. Poblete, M. Mendoza, A. Jaimes. *Microblogging without Borders: Differences and Similarities* ACM International Conference on Web Science 2011 - Koblenz, Germany

[13] **R. García-Gavilanes**, X. Amatriain. *Weighted Content Based Methods forRecommending Connections in Online Social Networks* Workshop RECSYS'10 – Barcelona Spain

Referees

Since 06/15   **Postdoctoral Advisor: Dr. Taha Yasseri**.
Dr. Taha Yasseri is a Research Fellow in Computational Social Science at the OII, a Faculty Fellow at the Alan Turing Institute for Data Science, and Research Fellow in Humanities and Social Sciences at Wolfson College, University of Oxford.
e-mail: taha.yasseri@oii.ox.ac.uk
address: 1 St. Giles Oxford OX1 3JS, United Kingdom

10/10-02/15   **PhD Advisor: Dr. Ricardo Baeza-Yates**.
Dr. Ricardo Baeza-Yates is CTO of NTENT, Professor at Universitat Pompeu Fabra (Barcelona) and Professor at Universidad de Chile (Chile). He was the Vice President of Research at Yahoo Labs working in that company for more than 10 years.
e-mail: rbaeza@acm.org, ricardo.baeza@upf.edu

10/12-10/14   **Co-Author: Dr. Daniele Quercia**.
Dr. Daniele Quercia is the Research Team Manager at Nokia Bell Labs in Cambridge, UK. He was Research Scientist at Yahoo Labs, a Horizon senior researcher at The Computer Laboratory of the University of Cambridge, and Postdoctoral Associate at the Massachusetts Institute of Technology.
e-mail: quercia@cantab.net, daniele.quercia@gmail.com
address: Bell Labs Ely Road 3, Milton Cambridge, CB24 6DD

10/12-10/14   **Co-Author: Dr. Yelena Mejova**.
Dr. Yelena Mejova is currently a scientist at Qatar Computing Research Institute, at the Social Computing group. Previously, she was a post-doc at Yahoo! Research Barcelona in the Web Mining group.
e-mail: ymejova@qf.org.qa