

# Gökhan Çiflikli, Ph.D.

[www.gokhan.io](http://www.gokhan.io)

[GitHub](#) | [LinkedIn](#)

Email : [g.ciflikli@lse.ac.uk](mailto:g.ciflikli@lse.ac.uk)

Mobile : +44 7477 104110

Twitter : [@gokhan\\_ciflikli](#)

## EDUCATION

---

- **London School of Economics** London, UK  
*Ph.D. in International Relations* Sep. 2014 – Sep. 2018
- **Uppsala University** Uppsala, Sweden  
*Master of Social Science in Peace and Conflict Studies* Aug. 2010 – June. 2012

## EXPERIENCE

---

- **UCDP** (Remote)  
*Data Scientist [R, Python]* Jan 2019 - Present
  - **Natural Language Processing:** I employ industrial solutions (SpaCy) to tokenise unstructured, multilingual text mined from aggregated news sources and human rights reports.
  - **Quantitative Text Analysis:** I recover meaning from tokenised text (approx. 48 million) that are actionable for policy makers using content/correspondence analysis and topic modelling.
  - **Supervised Learning:** I train ensemble models of binary- and multi-classification algorithms (Naive Bayes, Random Forest, Support Vector Machines, XGBoost) to automate the labelling process using a sample set of annotated events coded by humans.
  - **Active Learning:** I increase the predictive accuracy of the models (~10%) by having human annotators validate the predictions at regular intervals using the Prodigy annotation tool.
- **London School of Economics** London, UK  
*Software Developer [R]* Sep 2018 - Present
  - **Software:** [quanteda](#) is a popular R package (256K total downloads; 16K downloads/month) for quantitative text analysis. As a member of the Quanteda Initiative, I work on a commercial, scalable (up to 1,000 users) graphical user interface (GUI) cloud application using the interactive R-Shiny framework.
  - **Team Management:** I manage contracted R and JavaScript programmers (up to 3-5 at a time) building the infrastructure of the GUI via weekly video conferences and daily GitHub issue/milestone management.
- **Alan Turing Institute** London, UK  
*Software Developer [R, Python]* Dec 2017
  - **Predictive Modelling:** I built an interactive app for the UK Cabinet office for forecasting radical terrorism in the UK based on 300 profiles of convicted terrorists.
- **University College London** London, UK  
*Data Scientist [R]* Sep 2015 - 2017
  - **R package:** I designed an intuitive Random Forest fitter (RebelCast) for time-series-cross-section data to forecast escalation and severity of armed conflict which improves upon baseline models by 3%-26%.

## INTERACTIVE WEB APPS [R SHINY]

---

- **Adavis:** Visualises political polarisation in the US Congress over the last 70 years based on voting patterns consisting of 36K roll-call votes. [\[App\]](#) [\[Code\]](#)
- **Gender Bias:** Breaks down the gender bias in a leading UK higher education institution reading list comprised of 13K publications finding a 80/20 divide favouring male authors. [\[App\]](#) [\[Code\]](#)

## AWARDS, SCHOLARSHIPS, AND GRANTS

---

- **Knowledge Exchange & Impact Strategy Group Fund:** LSE. 2018. *2,000 GBP*.
- **Sir Patrick Gillam Fieldwork Scholarship:** LSE. 2017. *3,000 GBP*.
- **University of California, Berkeley Bursary:** LSE Ph.D. Mobility Program. 2016. *2,300 GBP*.
- **Social Research Methods Grant:** International Relations Department, LSE. *3,000 GBP*.
- **Michael Leifer Ph.D. Scholarship:** LSE. 2014-2017. *100,000 GBP*.
- **‘Facilitating Data Visualisation and Extraction: A New Interface for the Uppsala Conflict Data Program’:** Collaborator. R.J. 2014. *2.4 million SEK (approx. 211,000 GBP)*.

## SKILLS

---

- **Programming:** R (developer), SQL & Postgres (advanced), Python (intermediate)
- **Statistics:** Stata, SPSS (instructor)
- **Languages:** English, Turkish **Front-end:** HTML, CSS,  $\text{\LaTeX}$
- **Presentation:** RMarkdown, Reveal.js **Interactive:** Shiny

Last updated on April 11, 2019