

**Sofia Gil-Clavel, Ph.D.**

Computer, Actuarial, and Data Scientist

Machine Learning, Natural Language Processing & Statistical Analysis.  
Expert in Digital Trace Data Analytics. [sofiag1l.github.io](https://www.sofiag1l.github.io) [SofiaG1l](https://github.com/SofiaG1l) [B.S.GilClavel@tudelft.nl](mailto:B.S.GilClavel@tudelft.nl)**EDUCATION****Ph.D. Data Science applied to Population Studies**University of Groningen, NL  
2020 – 2023**M.Sc. Computer Science**Center of Mathematics Research, MX  
2015 – 2017**B.Sc. Actuarial Science**National Autonomous University of Mexico, MX  
2009 – 2014**MEMBERSHIPS**

- Canadian Consortium on Analytics for Data-Driven Decision-Making (CAnD3)
- International Max Planck Research School for Population, Health and Data Science (IMPRS-PHDS)

**SKILLS****Statistical Analysis [R, Python]:**

Frequentist	(★★★★★)
Bayesian	(★★★★★)
Inferential	(★★★★★)
Forecasting	(★★★★★)

**Programming:**

Python	(★★★★★)
R	(★★★★★)
C++	(★★★★★)
VBA	(★★★★★)

**Data visualization and exploration:**

R (shiny, ggplot)	(★★★★★)
Python (matplotlib)	(★★★★★)
Gephi	(★★★★★)

**Use of Application Programming****Interfaces (APIs):**

Facebook	(★★★★★)
Twitter	(★★★★★)
DeepL	(★★★★★)

**Languages:**

Spanish (native)	
English (professional level)	
Dutch (basic-partners' language)	
German (basic)	

**EXPERIENCE****Post Doctoral Researcher**Delft University of Technology, Netherlands  
Jul. 2022 – Today

Natural Language Processing applied to textual data to study societies' drivers of climate change adaptation. Creation of python package to visualize articles' findings using networks. Bayesian Multilevel analysis of data derived from text. Co-supervise Ph.D. students. Research publication in high ranked journal. Research presentation at international conferences.

**Digital Trace Data Analyst**Max Planck Institute for Demographic Research, Germany  
Sep. 2018 – Jun. 2022

Studying Social Interactions Using Digital Trace Data from Facebook and Twitter. Developing of R code to interact with the Facebook API. Processing and management of terabytes of data from the Internet Archive. Derive databases from Digital Trace Data to perform statistical analysis. Obtain statistical insights from the data to perform population studies. Research publication in high ranked journal. Research presentation at international conferences.

**Head of Demographic Analysis Unit**National Population Council, Mexico  
Aug. 2017 – Aug. 2018

Creation of computational tools to visualize and analyze Mexican conciliation data (1950-2015) and the Mexican projected population (2016-2050). Demographic data analysis of population, deaths, births, and migration. Programming and uploading (Linux servers) the first version (2018) of the shiny app to visualize these data.

**Research Support**El Colegio de Mexico, Mexico  
Aug. 2014 – Aug. 2015

Development of the R package "DemosR: Package for the Stochastic Population Forecast" (internal use). Analysis and handling of big data for demographic research.

## Appendix A: Complete list of Professional Activities

Invited speaker at the Department of Methodology and Statistics of Tilburg University to present “Big Data as a Data Source for Research: the Case of Facebook Data”. March 2023.

Invited speaker at the Permanent Seminar ‘Aplicaciones de ciencia de datos en las ciencias sociales y las humanidades’ of El Colegio de México. Presentation: “La asociación entre las políticas de integración cívica y de ciudadanía con el uso de la lengua de destino de los inmigrantes de la UE-15: Uso de datos de Twitter para analizar la adquisición de idiomas” (You can find the video [here](#)). November 2022.

Gave a virtual training Workshop at El Colegio de Mexico “V Coloquio Internacional Estudiantil en Población, Ciudad y Ambiente”: *Uso de APIs para la descarga de datos* ([link](#)). April 2022.

Invited as discussant at the panel “Machine learning in the social sciences” of the *MigrantLife Symposium* ([link](#)). March 2022.

Invited as discussant at the panel “Digital Traces for demographic knowledge” of the *Population Studies Association of Argentina* (AEPA) ([link](#)). October 2021.

Invited to the roundtable Demographic Research with Web and Social Media Data at the *Latin American Population Association*, organized by Diego Alburez-Gutierrez. You can find the video [here](#).

Gave a [training Workshop](#) at IUSSP Digital Demography in the Era of Big Data Pre-Workshop: *Demography Today Training* ([link](#)). June 2019.

Organizer at the ICWSM 2019 Workshop: *Demographic Research with Web and Social Media Data* ([link](#)). June 2019.

### Other Professional Activities:

- |   |                       |
|---|-----------------------|
| • Launched Reading Club at the SC3 Group of TU Delft. | [April 2023 – ...]    |
| • Co-launched Peer-Review Workshop at the IMPRS-PHDS. | [Aug.2020 – Apr.2022] |
| • Elected PhD Core Representative at the IMPRS-PHDS.  | [Nov.2019–Jun.2022]   |
| • Elected PhD External Representative at the MPIDR.   | [Jun.2020 – Jun.2021] |