

## Joe Millard—Data Scientist and Computational Ecologist

Address: 15 Tewkesbury Place, Nether Street, Nottingham, NG9 2BA

Email: joeymillard@hotmail.co.uk; Mobile: 0758 4168 355; GitHub: <https://github.com/Joemillard>

### Summary

I'm a data scientist and computational ecologist at UCL interested in the causes and consequences of global pollinator biodiversity change, and the power of online data in quantifying biodiversity awareness. Prior to my PhD I worked at UNEP-WCMC building automated data handling systems for wildlife trade data, and at the journal Scientific Reports in manuscript peer-review and decisions. I'm also an award-winning science writer and R Shiny developer. After my PhD I'm keen to build a research career in biodiversity modelling and conservation culturomics.

### Education and qualifications

#### PhD Computational Ecology – UCL London NERC DTP (2017-present)

Thesis: Pollinator biodiversity change in the Anthropocene

- Undertaking a CASE (Collaborative Award in Science and Engineering) PhD in collaboration with UCL (Dr Tim Newbold), the Institute of Zoology (Dr Robin Freeman), and the RSPB (Prof Richard Gregory)
- My project aims to collate current knowledge on pollinator decline, before building statistical models to predict the response of pollinators to future environmental change. I am also exploring developments in the novel field of conservation culturomics, using Wikipedia web-scraping to develop metrics for biodiversity awareness.
- My PhD applies a broad spectrum of big data tools in R, Python, and SQL.
- Co-supervised 2 masters students working on pollinator biodiversity change and text-mining.

#### MSc Biodiversity and Conservation, Distinction (1<sup>st</sup>) – The University of Leeds (2014-2015)

Thesis: A Satellite Telemetry Investigation of Caspian Seal (*Pusa caspica*) Activity Budgets (77)

- Awarded the second highest overall mark in the year for my data science research project and the highest thesis mark at 80.5%.
- Gained experience in large-scale data analysis using R packages ggplot2, reshape2, lme4, ggmap, and ganimate.

#### BSc Zoology, Upper second-class honours (2:1) – The University of Leeds (2011-2014)

Thesis: Of Molecules and Morphology: An Extended Review of the Eutherian Radiation (69)

### Employment and volunteering summary

#### Postgraduate teaching assistant (2019-2021)

- Taught statistical programming on two computational biology modules, for a total of more than 100 hours. Covered statistical methods (lme4, anova, lm), programming (basic Python), data handling (piping, dplyr verbs), and functional programming in R.

#### Academic tutor – UKonexão (2018-2019)

- Set up a private coaching service for Brazilian students wanting to study in the UK.
- Delivered nine 1 hour 1-1 Skype advice sessions for students at universities in Brazil.
- Provided verbal and written comments on CVs for 9 undergraduate and postgraduate students in Brazil, and gave a talk at the University of São Paulo on studying in the UK.

#### Programme Assistant – UNEP-WCMC (September 2016-June 2017)

- Collaborated with Dr Aly Pavitt to develop two new automated systems in R for the processing of CITES data, replacing a previous system built in VBA.
- Built the front end Shiny app, combined the UI and server side scripts, tested the app in the initial phase, and wrote a guide manual.
- Delivered workshops on Shiny to the Species Programme, the wider centre, and a specialist technical GIS hub.

- Involved in producing a data handling guide for the Ecuadorian environmental ministry and an Amazonian regional trade overview.
- Designed a new interactive data visualisation CITES tool for the WCMC informatics team.
- Wrote ~12 high quality wildlife trade reviews for the European Commission.

#### **Publishing Assistant – Nature Research** (November 2015-September 2016)

- Applied my broad academic knowledge to rapidly and accurately assign manuscripts to the editors at Scientific Reports.
- Trained in the handling and identification of clinical trials and responsible for processing and vetting trials before acceptance.
- Performed initial quality control and liaised with senior colleagues on problem articles.
- Presented a high impact ecological paper to the team and wrote paper summaries for our Facebook page, published to more than 100,000 followers.

#### **Social Media Coordinator and Blogger – Eco Sapien** (November 2014-November 2015) **& Graph Start** (September 2015-September 2018)

- Key team member of Eco Sapien, a non-profit organisation dedicated to ecological education.
- Published more than 50 popular blog posts and 15 extended articles, and occasionally gained over 1000 reblogs and likes for my posts.
- Set up Graph Start, an educational website dedicated to data visualisation in R.
- Wrote blogs on R programming and took on three voluntary team members: one graphic designer to assist with branding and two bio-statisticians to provide R tutorials.

#### **Lead Conference Coordinator and Speaker** (January 2015)

- Won a place as lead coordinator on the first IEMA Student Sustainability Conference, a national event with more than 100 delegates.
- Delivered an introductory speech, ran two conference sessions, devised timetables, booked caterers, researched marketing strategies, and later invited to meet the CEO of IEMA.

#### **Research Assistant – Action for the Wild** (Summer 2014) **& Cetacean Research and Rescue Unit** (June 2016)

- Worked as a field assistant for 2 months on the Umphafa Private Nature Reserve, South Africa, and for 2 weeks in the Moray Firth.
- Gained core field research skills and an ability to lead and assist undergraduate researchers.

### **Publications**

- **Millard, J.**, Freeman, R., Newbold, T., 2020. Text-analysis reveals taxonomic and geographic disparities in animal pollination literature. *Ecography*, 43(1), pp. 44-59.
- **Millard, J.**, Gregory, R., Jones, K., Freeman, R., 2020. The Species Awareness Index: a Wikipedia derived conservation culturomics metric for public biodiversity awareness. *Conservation Biology* (Accepted, selected to appear in an 'Advancing Conservation Culturomics' special section). <https://doi.org/10.1101/2020.08.17.254177>
- **Millard, J.**, Outhwaite, C., Kinnersley, R., Freeman, R., Gregory, R., Adedoja, O., Gavini, S., Kioko, E., Kuhlmann, M., Ollerton, J., Ren, Z-X., and Newbold, T., Global effects of land-use intensity on local pollinator biodiversity (in revision, *Nature Communications*)
- Newbold, T., Adams, G.L., Robles, G.A., Boakes, E.H., Ferreira, G.B., Chapman, A.S., Etard, A., Gibb, R., **Millard, J.**, Outhwaite, C.L. and Williams, J.J., 2019. Climate and land-use change homogenise terrestrial biodiversity, with consequences for ecosystem functioning and human well-being. *Emerging Topics in Life Sciences*, 3(2), pp.207-219.
- August, Tom A, Affouard, Antoine, Fox, Nathan, Marlowe, Celia, **Millard, J.**, Sanderson, Roy, & Bonnet, Pierre. (2019). AI validated plant observations from social media: Flickr images from central London 2011-2019 (Version 1.0) [Data set]. Zenodo. <http://doi.org/10.5281/zenodo.3497314>

## Key skills, other information, and hobbies and interests

- **Core data science skills:** Data visualisation (shiny, ggplot2, plotly, gganimate, ggraph, DiagrammeR, patchwork); data handling (dplyr, reshape2, pandas, numpy); GIS (ggmap, raster, ArcGIS, rgdal, leaflet); database management (MySQL, SQLite); web development and web-scraping (css, html, jquery, WordPress, R Markdown, rvest); statistical programming (linear modelling, mixed-effects modelling); natural language processing (stringr, taxize, named-entity recognition, tf-idf, cosine similarity); geoparsing (CLIFF-CLAVIN); version control and reproducibility (checkpoint, pacman, git, GitHub, here::here)
- **Other information:** Full, clean driving license; Marine Mammal Medic; Outdoor Field Safety Certificate; RLSS Bronze Medallion; languages (English; Portuguese, pre-intermediate—A2)
- **Hobbies and interests:** Reading (anything by Kim-Stanley Robinson, sci-fi, popular science, and history); watching movies (sci-fi, thrillers, documentaries); travelling; outdoor activities (camping, walking, kayaking); eating out and socialising with family/friends/colleagues

## Talks and workshops given

- **BES Festival of Ecology 2020** –The Species Awareness Index: a Wikipedia derived conservation culturomics metric for public biodiversity awareness
- **SCAPE (Scandinavian Association for Pollination Ecology) 2020** – Global effects of land-use intensity on local scale pollinator biodiversity
- **Linnaean Student Conference 2019** – The importance and diversity of animal pollination (<https://www.youtube.com/watch?v=qn0sCxJiTdl&t=5854s>)
- **Planet PhD podcast** – Pollinator biodiversity in a changing world (<https://planetphd.podbean.com/e/joe-millard-pollinator-biodiversity-in-a-changing-world/>)
- **UCL seminar** – Text-analysis reveals taxonomic and geographic disparities in animal pollination literature (<https://www.ucl.ac.uk/biosciences/events/2019/jun/cber-research-talk-joe-millard-17-june-2019>)
- **SCCS (Student Conference on Conservation Science) Europe 2019** – Text-analysis reveals taxonomic and geographic disparities in animal pollination literature
- **RSPB Annual Science Meeting 2019** – Text-analysis reveals taxonomic and geographic disparities in animal pollination literature
- **Sao Paulo, Brazil 2019** – Postgraduate study in the UK: How, when and why?
- **Institute of Zoology 2018 & UCL CBER 2018-2021** – Hosted an IoZ GitHub R clinic and a CBER journal club (fortnightly in-person) and CBER data club (fortnightly online)

## Conferences and workshops attended

- **London R 2018** (Introduction to Data Science) & **Datacamp 2018** (Introduction to SQL)
- **H2O AI conference, London 2019**
- **Nature Research, London 2019** – 2 day manuscript writing workshop
- **Brexit BES early career workshop, London 2019** – 1 day workshop on adapting to Brexit
- **UCL Arena 1 teaching assistant training 2019 & UCL copyright training 2019**
- **UCL intermediate Python programming 2019**

## Prizes and awards

- **BES Quantitative Ecology Hackathon 2020** – awarded “best presentation” for our plant species identifying AI Shiny app (<https://methodsblog.com/2020/03/12/quantitative-sig-hackathon/>)
- **Leeds University Science Magazine Writing competition** – won first prize for my piece “The polar bear: a tale of two futures”
- **IEEMA Student Sustainability Conference** – won a committee place for my essay on sustainability issues
- **Linnaean Student Conference 2019** – won a place as speaker for my abstract on pollinator biology