

CITIZEN SCIENCE WITH OPEN DATA AND OPEN SOURCE SOFTWARE: A CASE STUDY

CHRISTINE POERSCHKE

[GITHUB.COM/CPOERSCHKE/BEE-INFORMATICS](https://github.com/cpoerschke/bee-informatics)

BACKGROUND

THE CITIZEN

Hello,
My name is Christine and I am:
• A spare-time beekeeper in suburban London
• A full-time software developer at Bloomberg
• Studying for the British Beekeepers Association module examinations, obtaining the intermediate theory certificate recently
• A 2004 PhD graduate [2]
• An active and experienced open source software contributor

In their 2018 open access journal paper entitled "Tracking All Members of a Honey Bee Colony Over Their Lifetime Using Learned Models of Correspondence", Boenisch et. al. [4] present an in-depth description of a multi-step algorithm which produces motion paths of automatically tracked marked honey bees. Alongside their paper, the team at Freie Universität Berlin published the first trajectory dataset for all bees in a colony, extracted from 3 million images covering 3 days.

The sample dataset entitled "BeesBook Recording Season 2015 Sample Release" [5] is available online and the full dataset comprises 71 days of continuous positional data (3 Hz sample rate) and, in total, 2,775 bees in a one-frame observation hive were marked and recorded with four cameras.

THE SOFTWARE

The free availability of open source software is key to getting started.

Here's software I've used so far:

- Apache Solr
- Python, including:
 - Matplotlib
 - NumPy
 - QRCode
- Java
- Apache Maven

EXPLORING THE BEESBOOK 2015

DATASET SAMPLE WITH APACHE SOLR

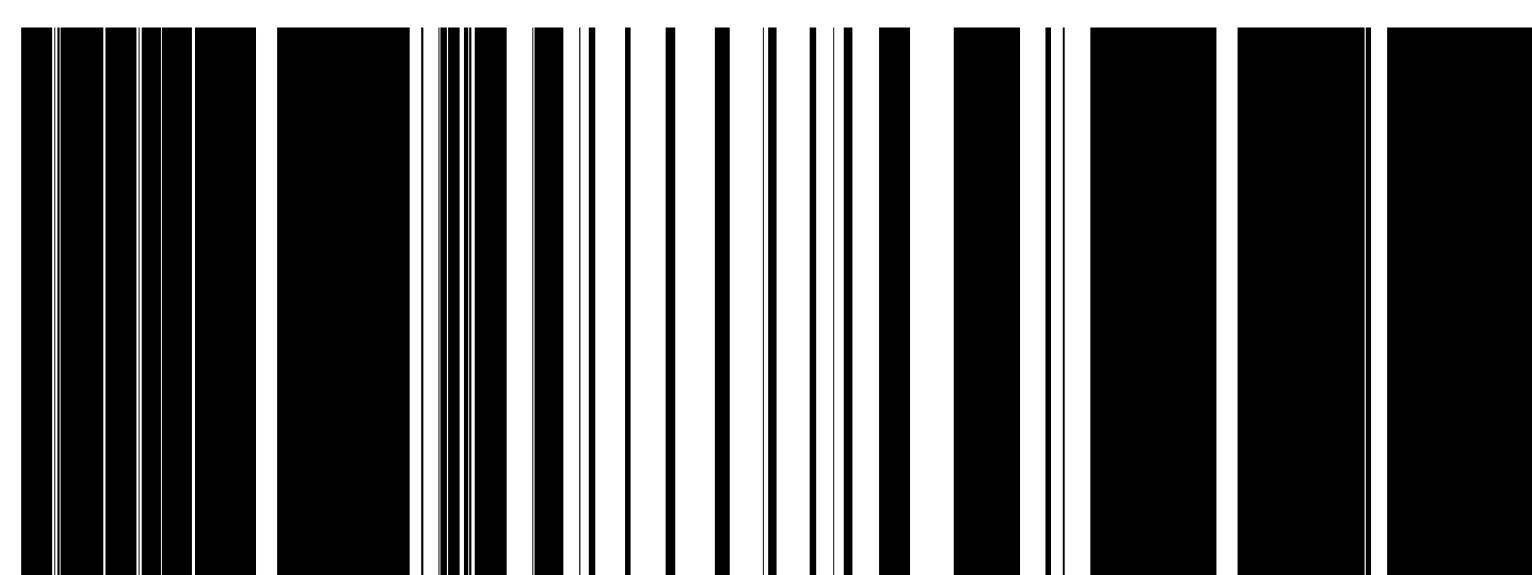
Self-service data access, free of charge, and clear terms of use are key to getting started.

The data lends itself to visualisation and the quantity of data points (ca. 200 million) can be motivating when exploring technical software features such as heatmaps and faceting.

ILLUSTRATION:

- When was Bee 104 seen (black bars) and not seen (white bars)?

```
http://localhost:8983/solr/bee-hive/select?facet=true
&q=bee_id_i:104
&rows=0
&facet.mincount=0
&facet.range=timestep_dt
&facet.range.gap=%2B1MINUTE
&facet.range.start=2015-09-08T00:00:00.000Z
&facet.range.end=2015-09-08T23:59:59.999Z
```

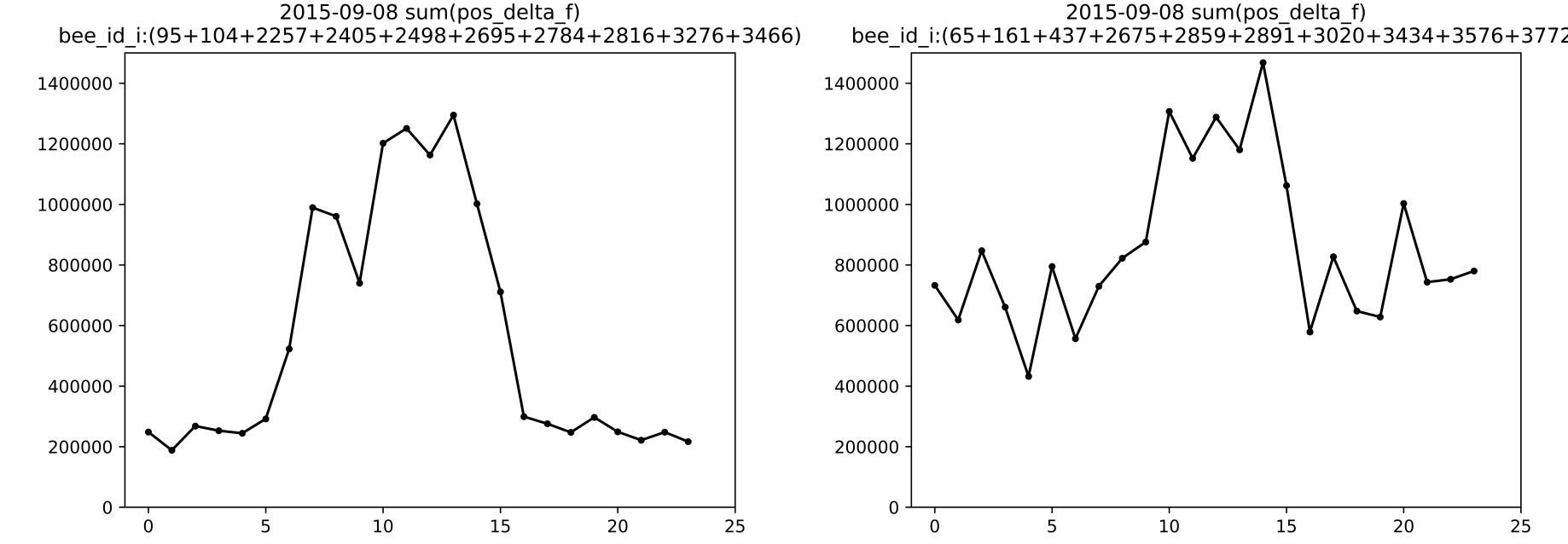


BEDTIME FOR BEES IS WHEN?

- Bee literature openly accessible via ResearchGate includes Crailsheim et. al. [1] on diurnal behavioural differences in forager and nurse honey bees, as well as Klein and Seeley [3] on foragers opportunistically napping during the day when forage is not available.
- Bee steps data can be number-crunched to identify groups of bees whose bee step metrics vary across a 24-hour day.

ILLUSTRATION:

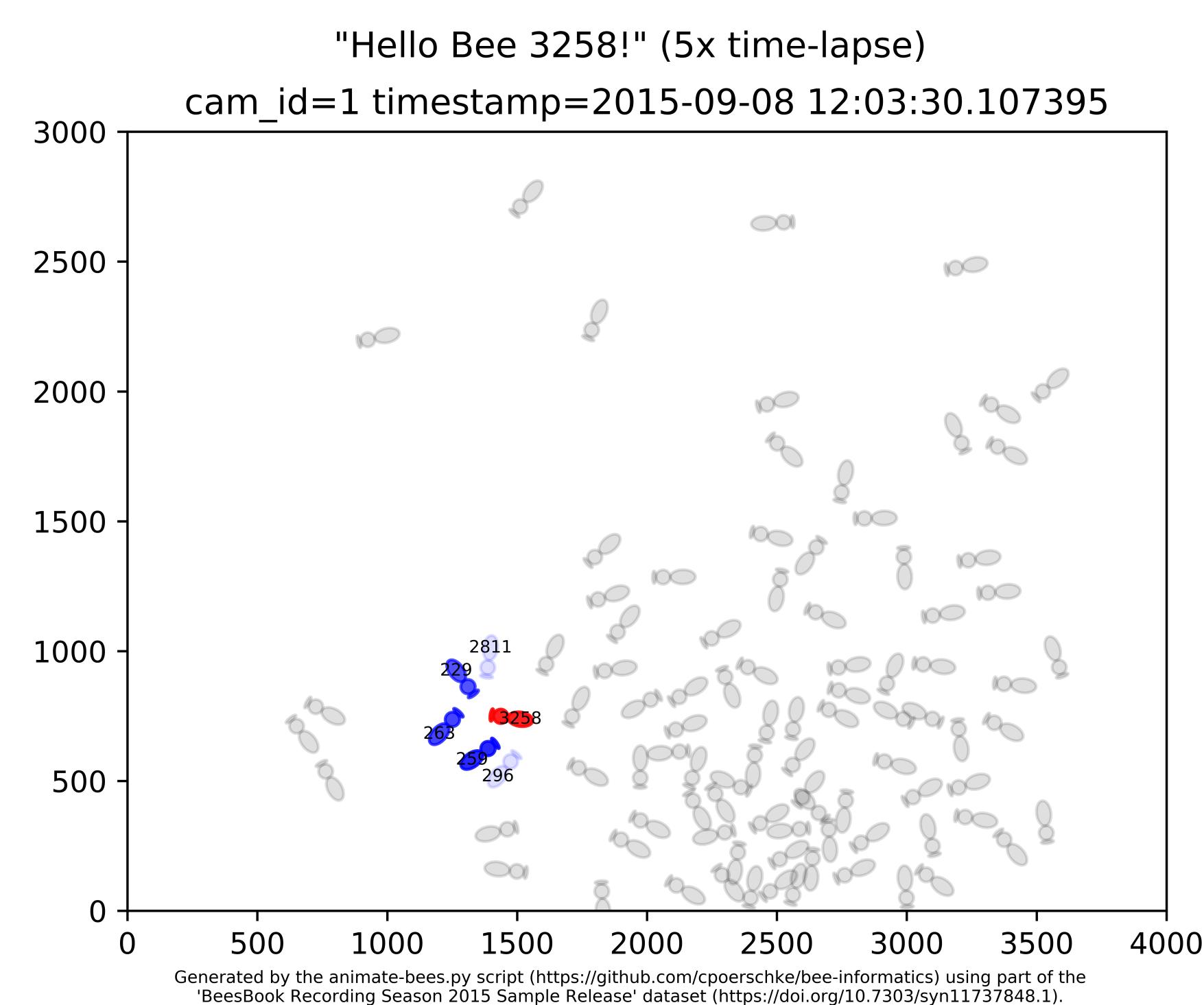
- Two groups of ten bees for whom at least 160,000 data points per bee were available and whose day vs. night behaviour varies.



ANIMATING BEES

Observation hive data can be visualised as labelled and color-coded cartoon-style bees: red for the focal bee and dark hues of blue for frequently met contacts and lighter hues for bees that only briefly met the focal bee.

Scan to watch the "Hello Bee 3258!" video:



FUTURE IDEAS

A BEE DATA HACKDAY

- A one day event to connect local bee and technology communities: bee researchers, bee enthusiasts, bee keepers, technology professionals, students and faculty.
- A collaborative space with Wi-Fi for 20-30 people and their laptops.
- An observation hive is nice-to-have, but optional.

SCHEDULE:

- Introductions: people, bees, data
- Brainstorming ideas:
 - What questions could be explored?
 - What technologies could be used?
- Group formation around shared ideas.
- Lunch.
- Work in groups.
- Informal presenting of the groups' work.

ONLINE BEESPACE

- Share your ideas on social media: [#APIMONDIA2019 #CITIZENSCIENCE](#)
- Subscribe to and/or comment on this 'Citizen Science with Open Data' ideas thread: [GITHUB.COM/CPOERSCHKE/BEE-INFORMATICS/ISSUES/1](https://github.com/cpoerschke/bee-informatics/issues/1)

REFERENCES

1. Crailsheim K, Hrassnigg N, Stabentheiner A (1996) Diurnal behavioural differences in forager and nurse honey bees (*Apis mellifera carnica* Pollm.). *Apidologie*. 27, 235-244. ([DOI.ORG/10.1051/APIDO:19960406](https://doi.org/10.1051/apido:19960406))
2. Poerschke C (2004) Development and evaluation of an intelligent handheld insulin dose advisor for patients with type 1 diabetes. Oxford Brookes University. ([ETHOS.BL.UK/ORDERDETAILS.DO?UIN=UK.BL.ETHOS.401010](https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.401010))
3. Klein BA, Seeley TD (2011) Work or sleep? Honeybee foragers opportunistically nap during the day when forage is not available. *Animal Behaviour*. 82, 77-83. ([DOI.ORG/10.1016/J.ANBEHAV.2011.03.026](https://doi.org/10.1016/j.anbehav.2011.03.026))
4. Boenisch F, Rosemann B, Wild B, Dormagen D, Wario F, Landgraf T (2018) Tracking all members of a honey bee colony over their lifetime using learned models of correspondence. *Front. Robot. AI* 5:35. ([DOI.ORG/10.3389/FROBT.2018.00035](https://doi.org/10.3389/FROBT.2018.00035))
5. Boenisch F, Rosemann B, Wild B, Wario F, Dormagen D, Landgraf T (2018) BeesBook Recording Season 2015 Sample. ([DOI.ORG/10.7303/SYN11737848.1](https://doi.org/10.7303/syn11737848.1))