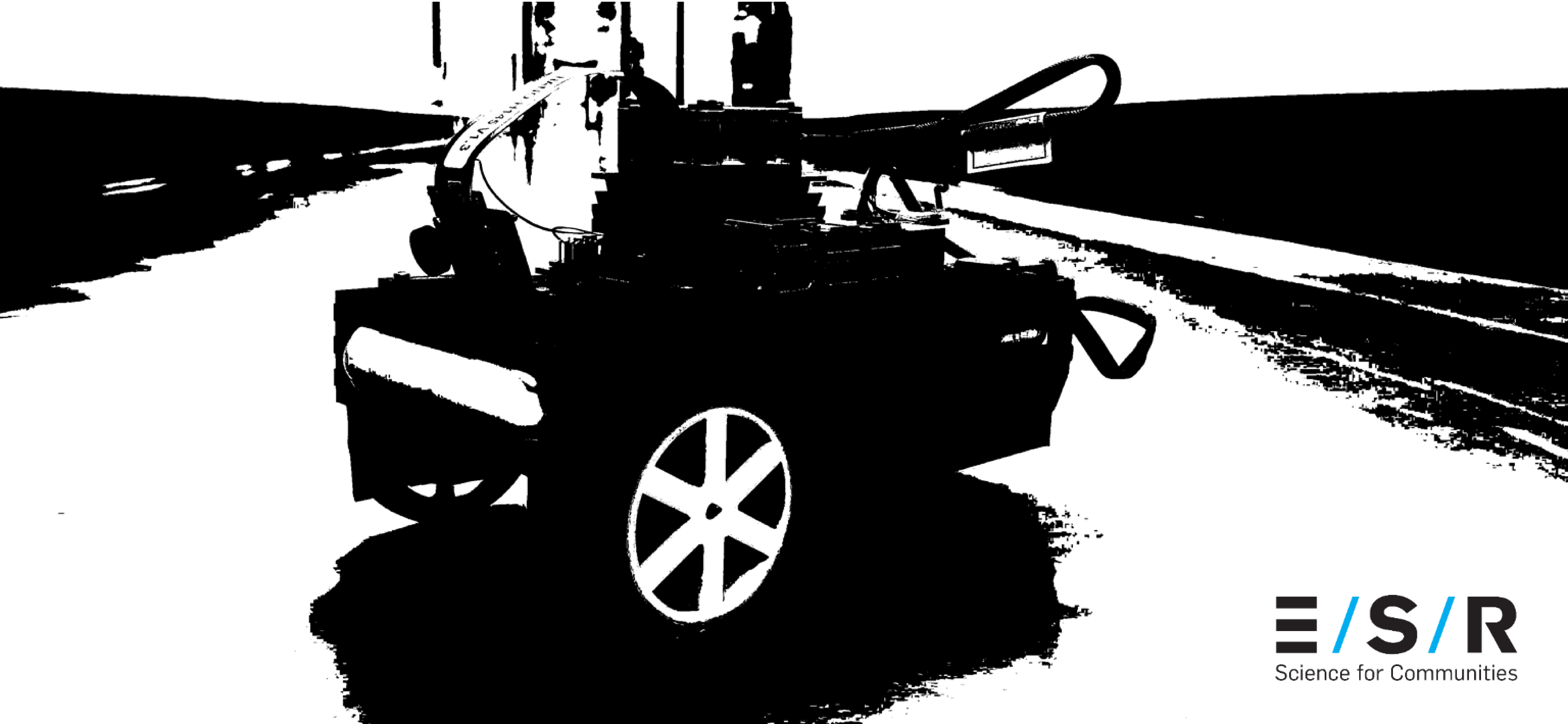
Accelerating data science

Richard Dean | Data Scientist | ESR



Science for Communities

Cheshire

Snowdon summit view by Denis Egan, on Flickr

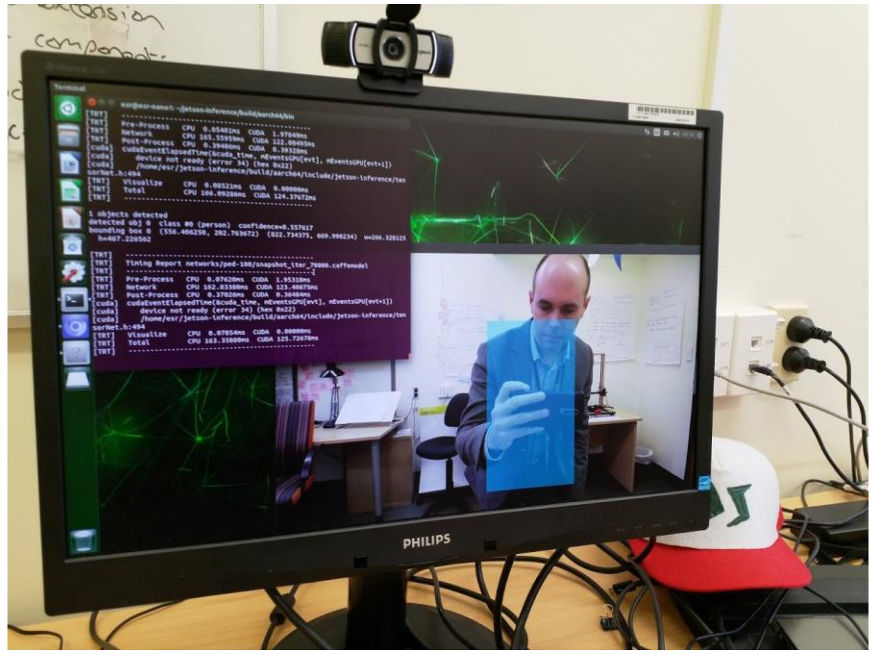
Science for Communities

River **Dee** at Llangollen by Richard

Science for Communities

Dean


HILUX
Science for Communities


PHILIPS


Christchurch
Science Centre

Richard

Science for Communities

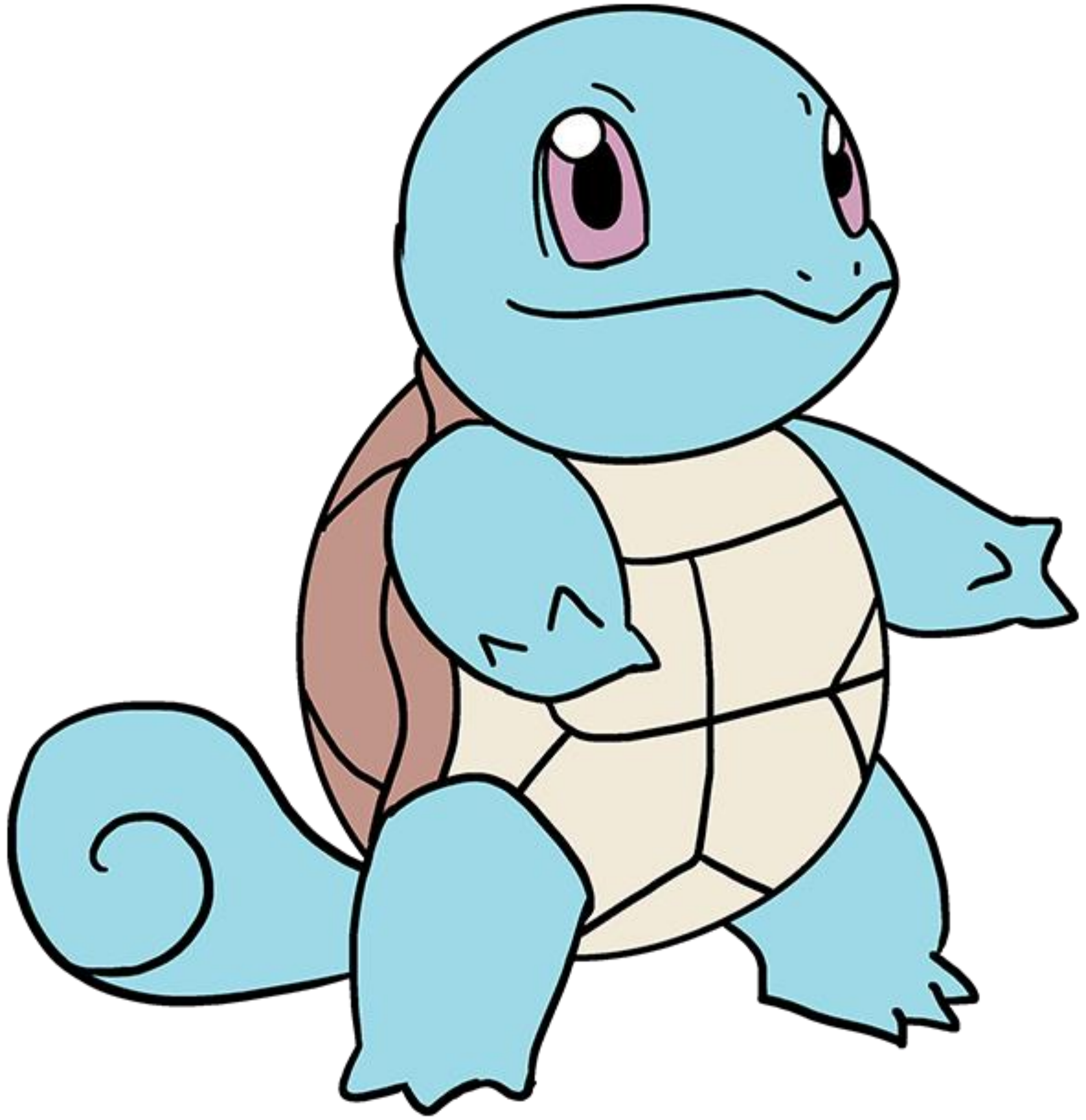
Hi everyone!

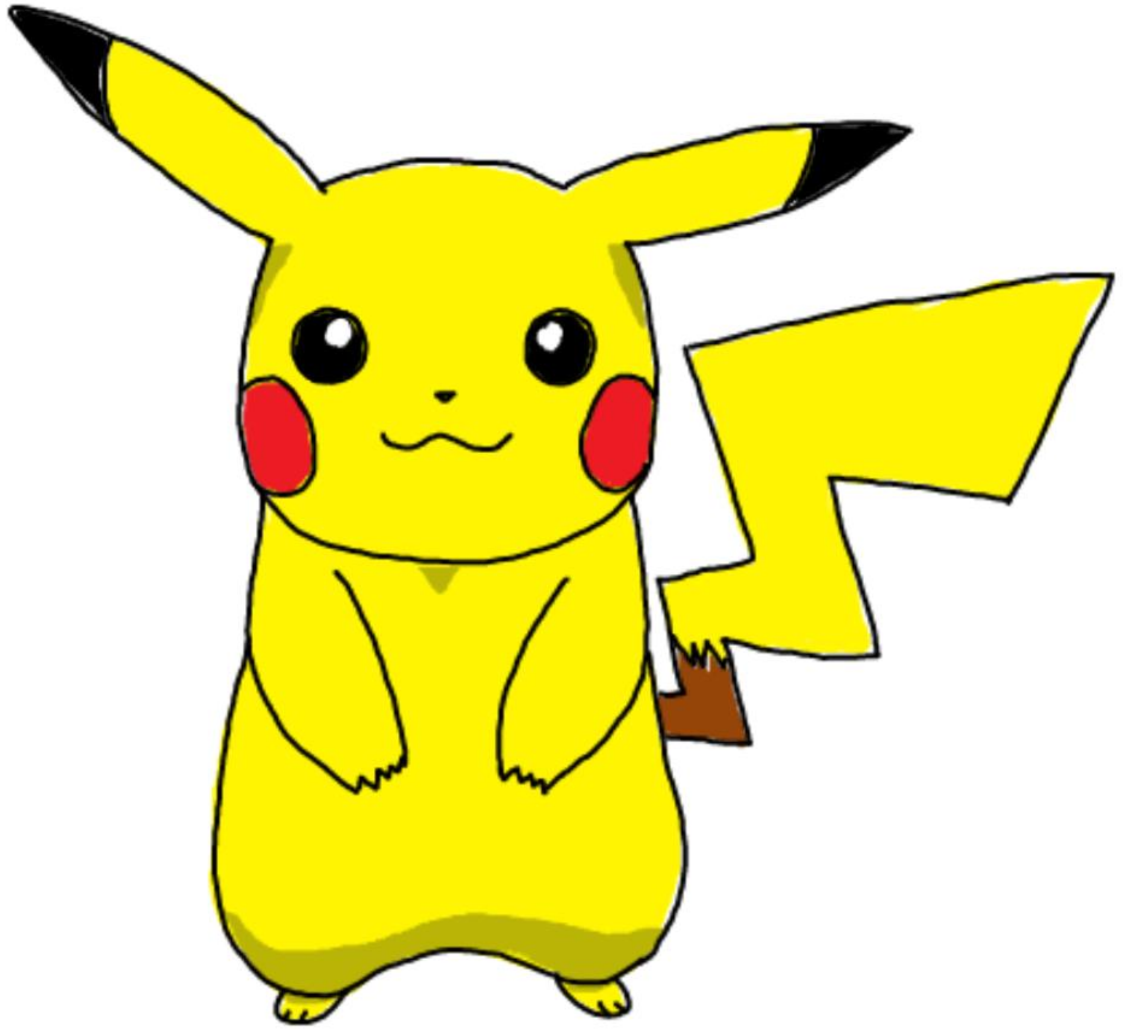
My mission today

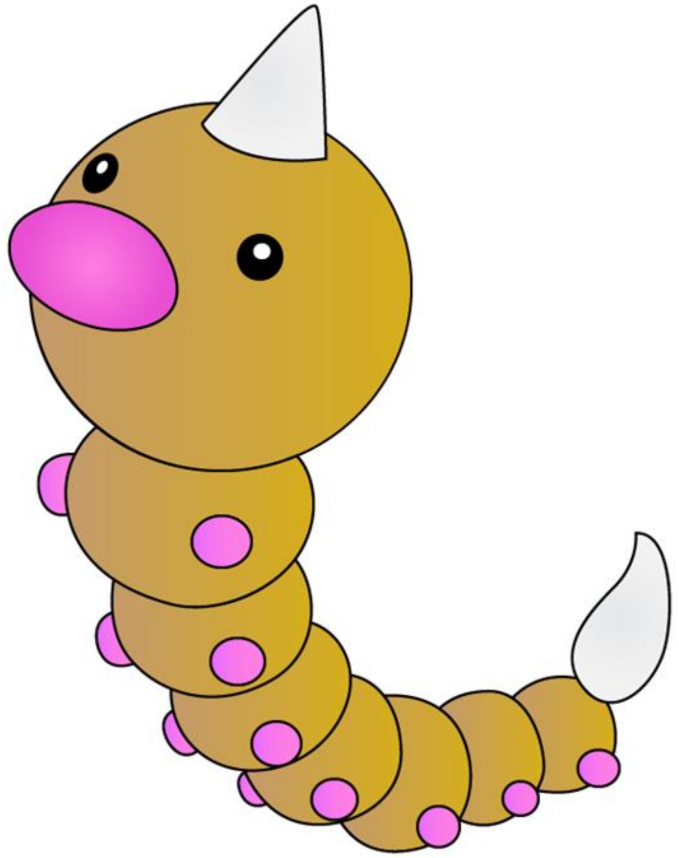
- talk a bit about data science
- what we’ve done to accelerate
- super powers
- evolve
- team up
- have some fun

Credit - How To Draw [redacted] - YourDrawingLessons.com

E/S/R Science for Communities





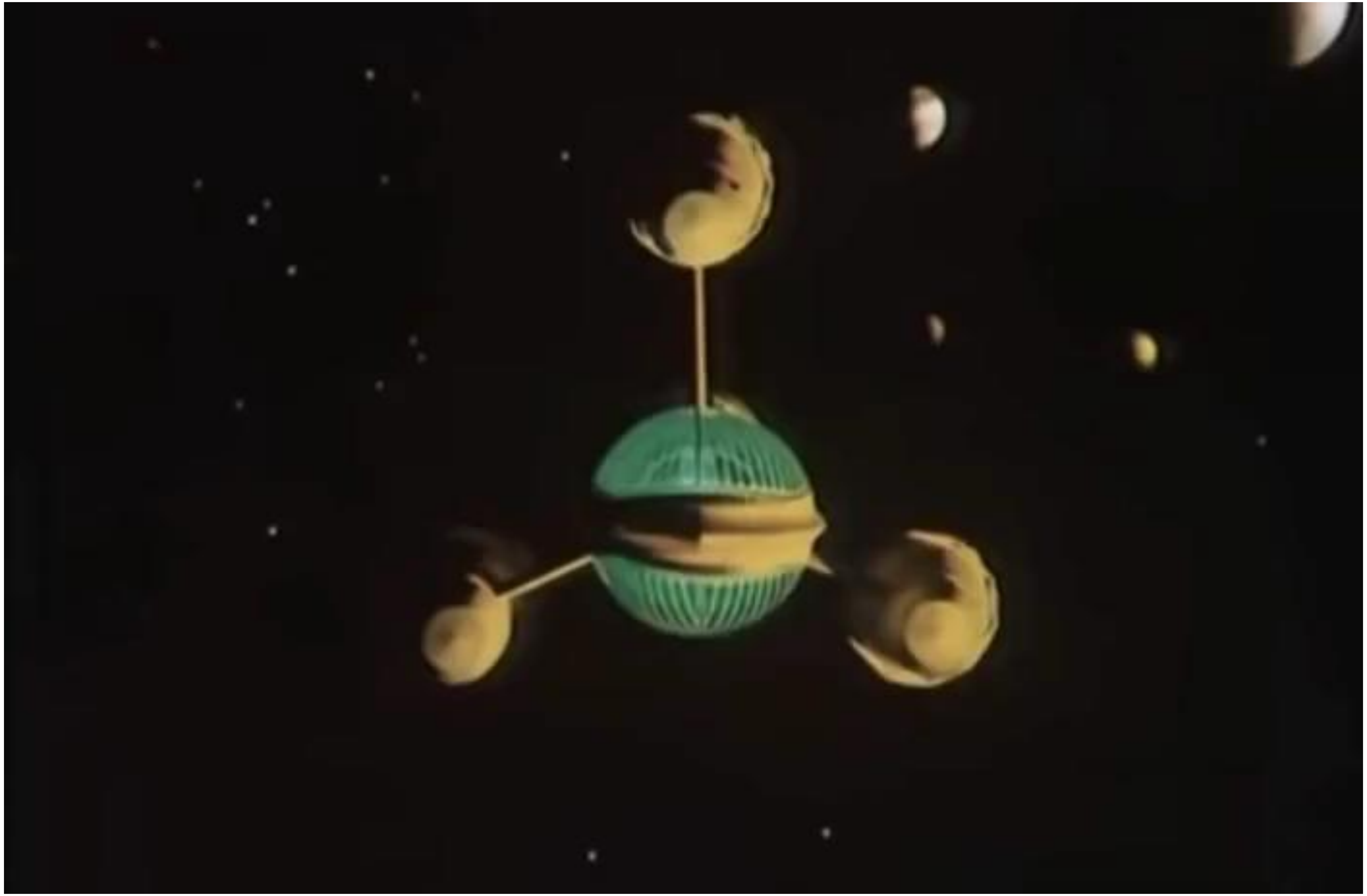


https://easydrawingguides.com/how-to-draw-squirtle-pokemon/

https://www.drawingtutorials101.com/how-to-draw-zubat-from-pokemon

https://www.drawingtutorials101.com/how-to-draw-weedle-from-pokemon

https://www.drawingnow.com/tutorials/15695/how-to-draw-pikachu/



ESR
Science for Communities

Ash is a Pokémon trainer

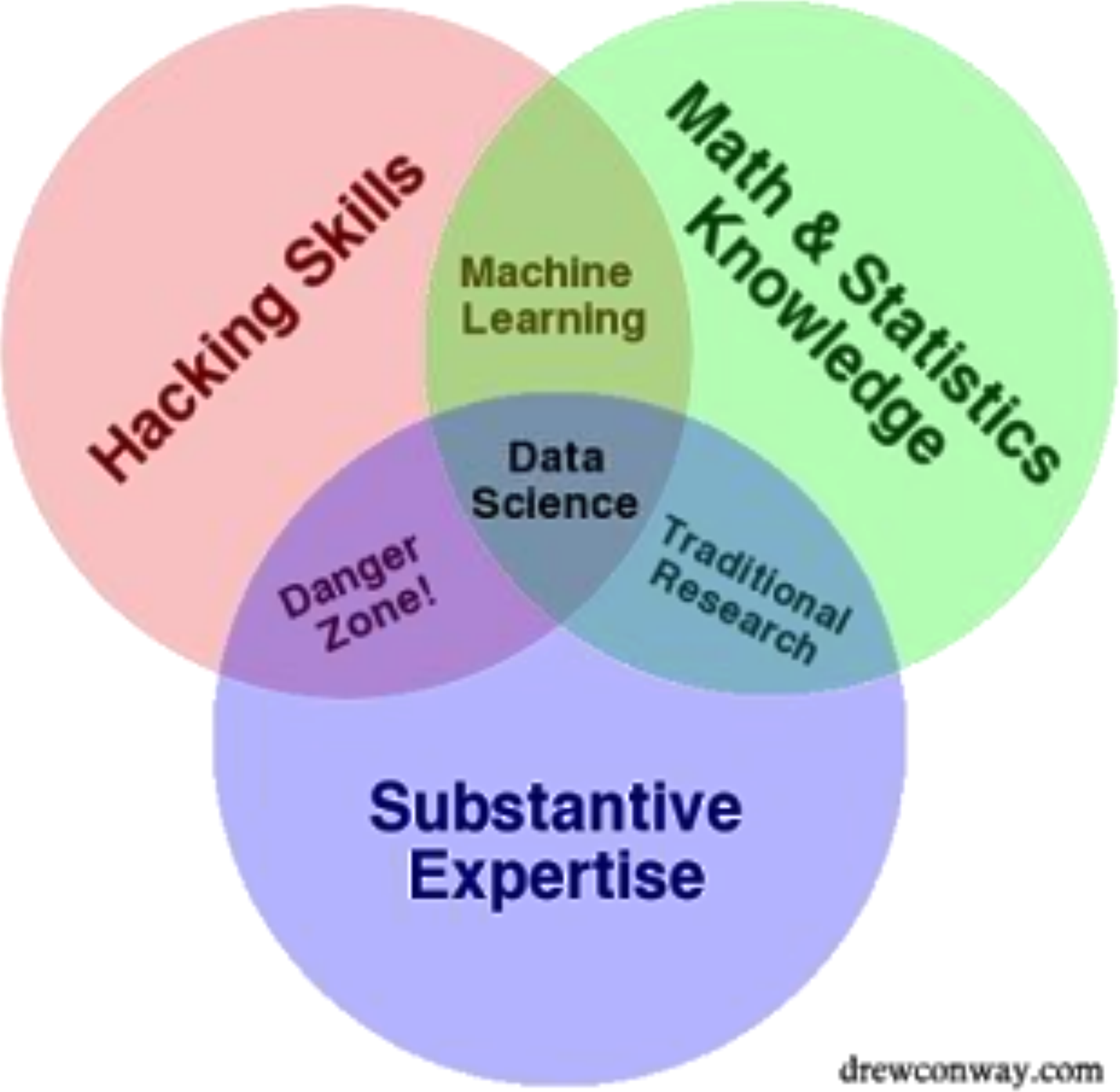
Credit - [How To Draw Ash Ketchum From Pokémon!](#) - YourDrawingLessons.com

Science for Communities

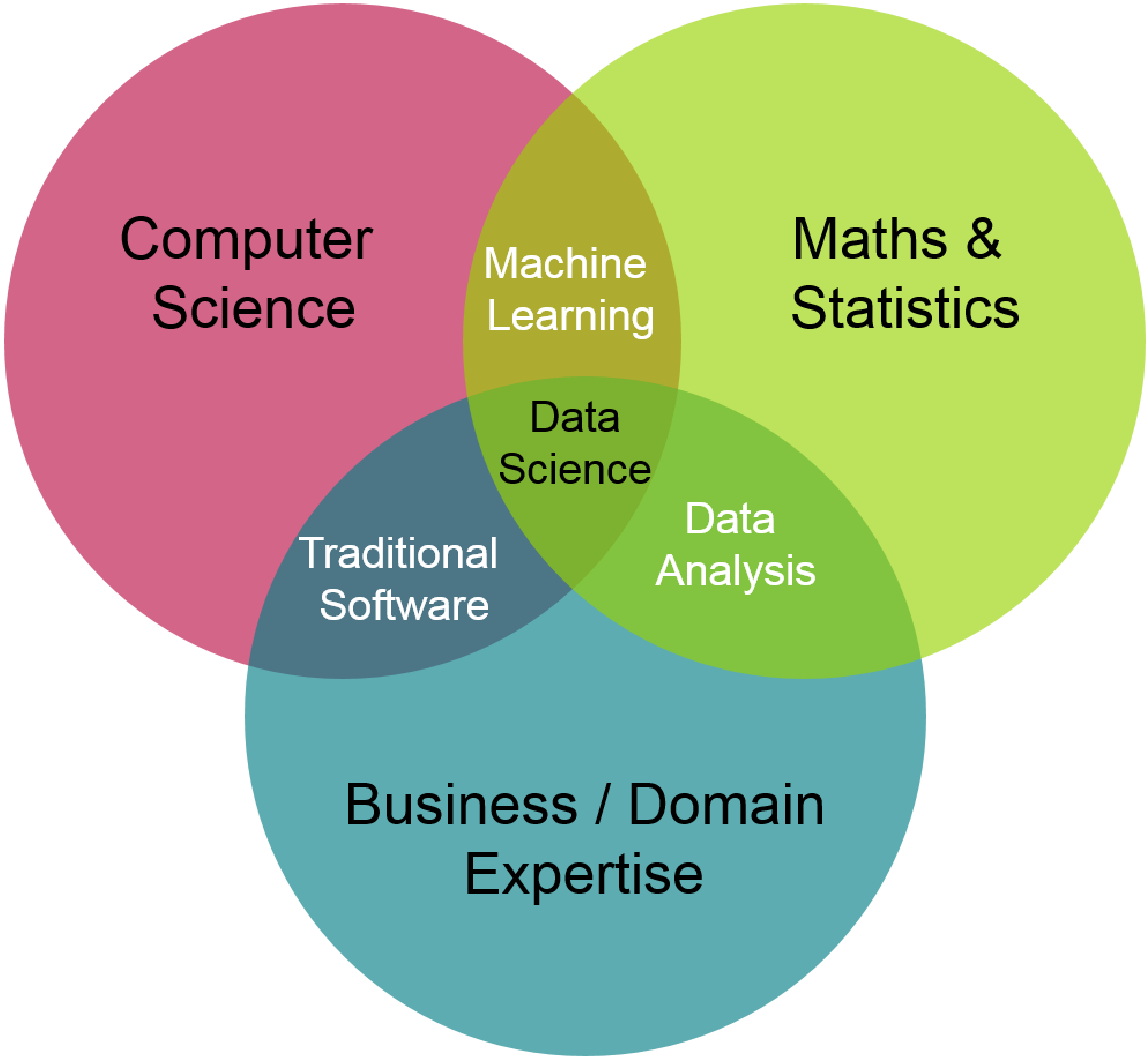
"A **data scientist** is someone who can **obtain**, **scrub**, **explore**, **model** and **interpret data**, blending **hacking**, **statistics** and **machine learning**. Data scientists not only are adept at working with data, but appreciate data itself as a first-class product."

Hillary Mason, Data Scientist, Accel, Scientist Emeritus, bitly, co-founder, HackNY

Science for Communities

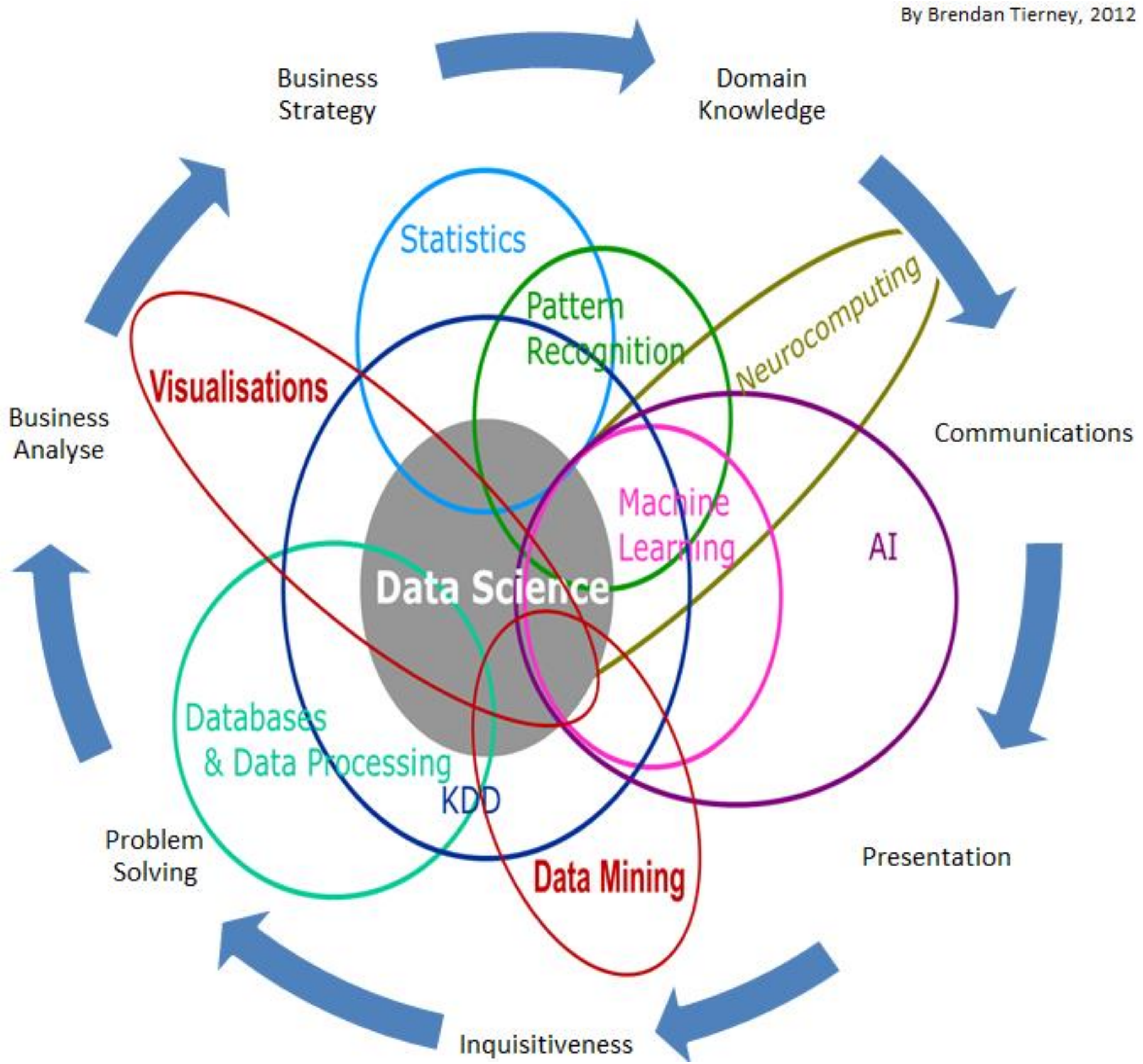


E/S/R
Science for Communities



Science for Communities

https://medium.com/@mike.dudkey78/wonderful-article-may-be-a-venn-diagram-could-help-i-found-this-one-f8f69a8ad1b3



https://brendantierneydatamining.blogspot.com/2012/06/data-science-is-multidisciplinary.html

Science for Communities

What do we mean by “Data Science”?

To improve **decision making**

by basing decisions on **insights**

extracted from **data**

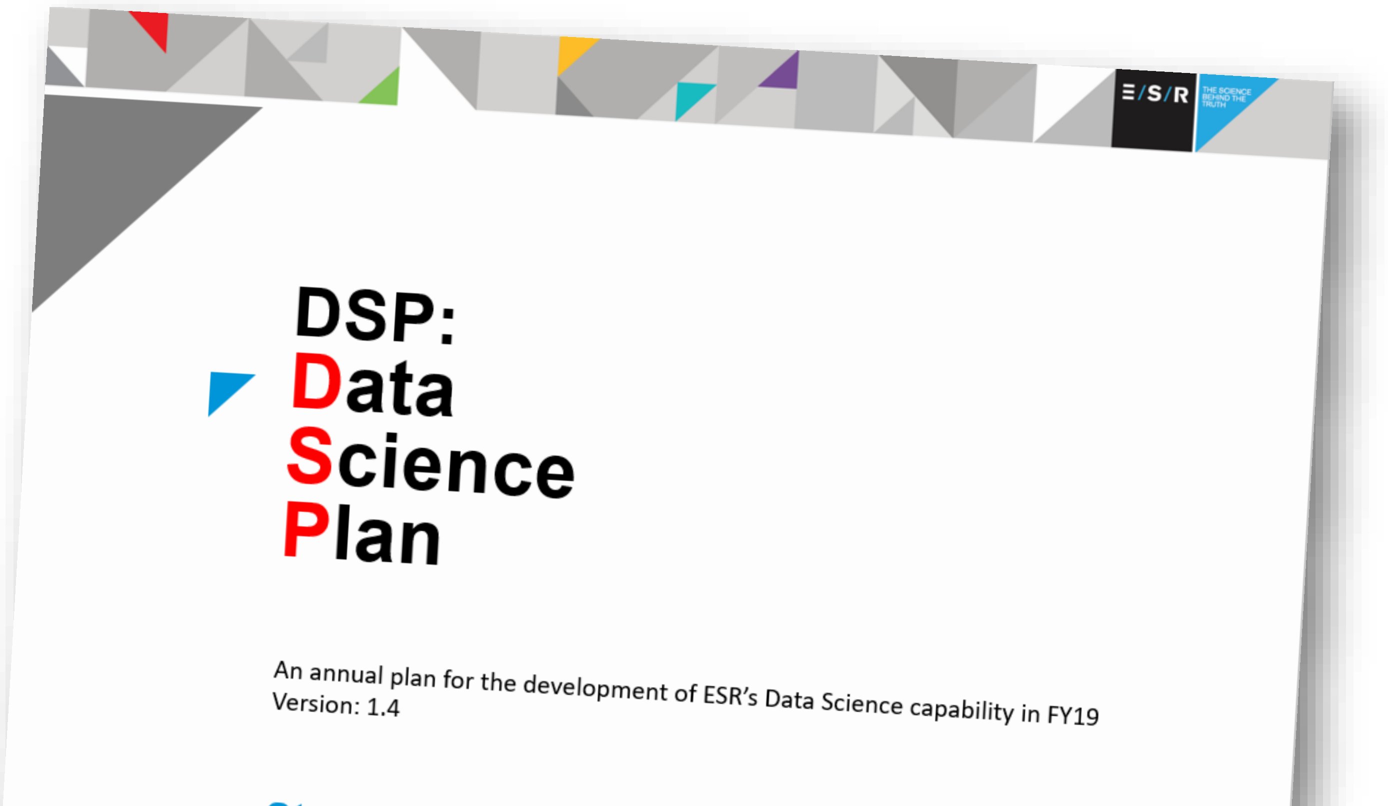
E/S/R
Science for Communities

Our Definition: Data Science

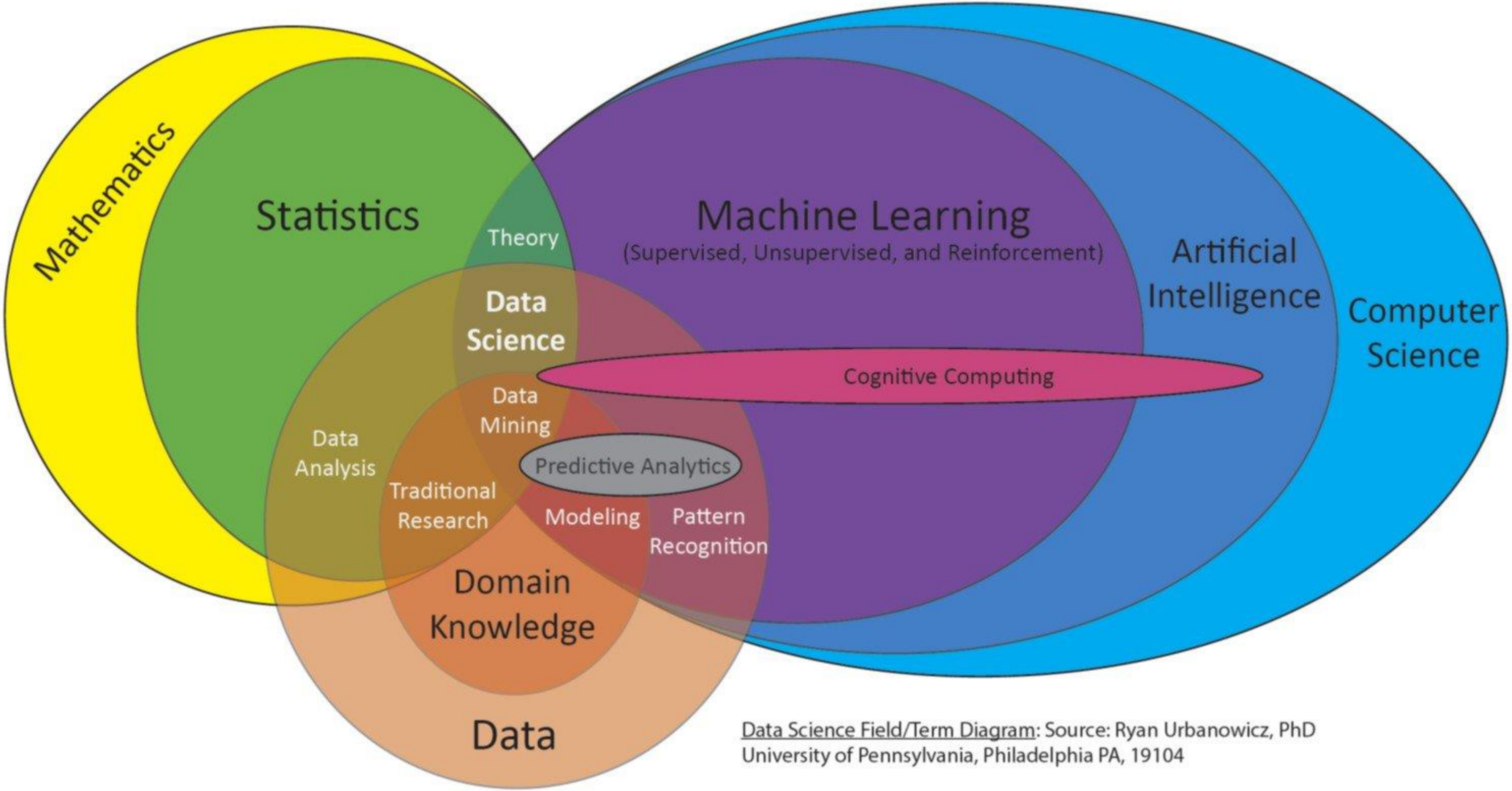
“The **interdisciplinary** field of inquiry in which **quantitative** and **analytical approaches, processes**, and **systems** are developed and used to **extract knowledge and insights** from increasingly large and/or complex **data sets**.”

United States National Institutes of Health

ESR Data Science Plan


DSP:
Data
Science
Plan

An annual plan for the development of ESR's Data Science capability in FY19
Version: 1.4

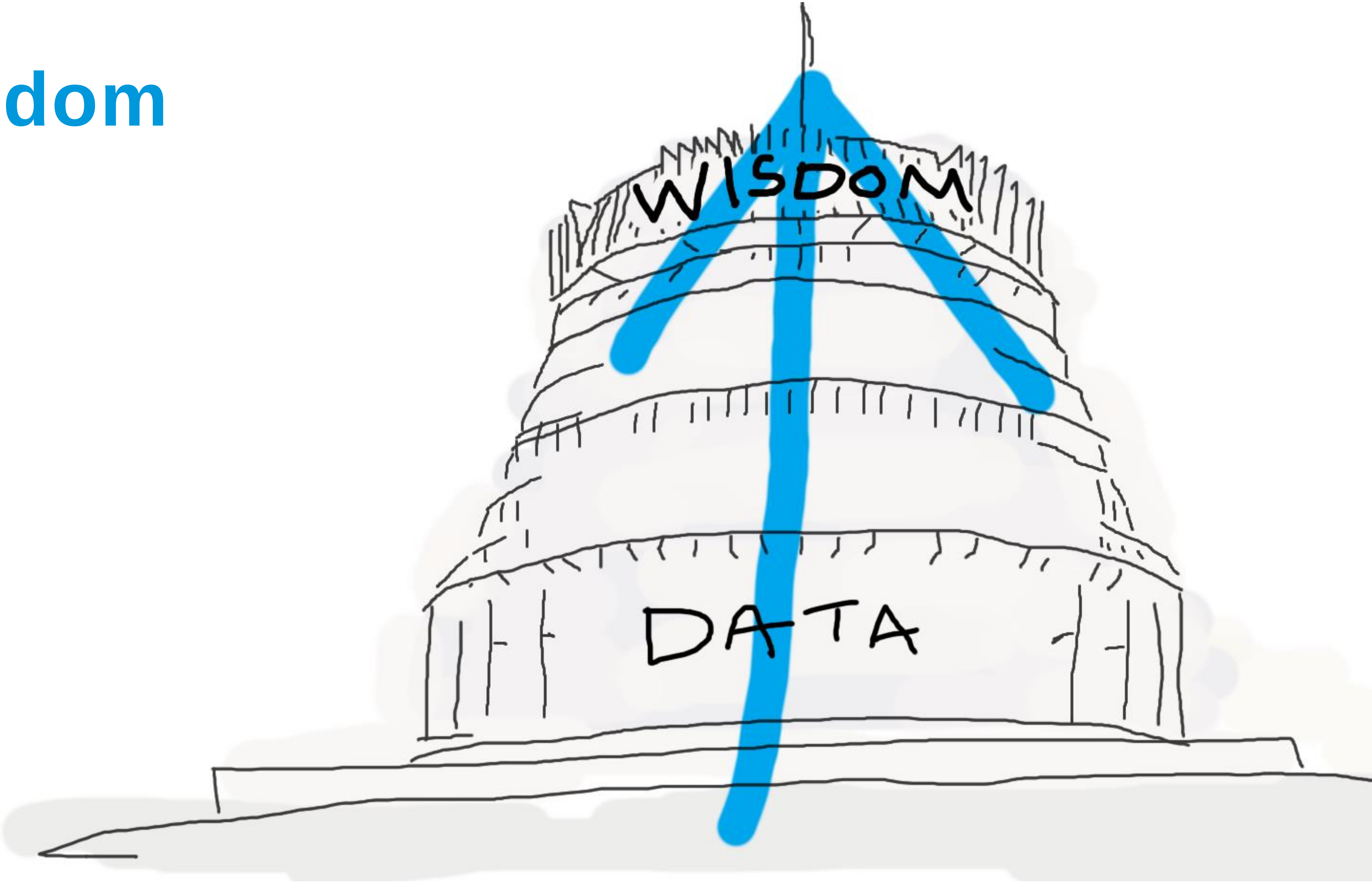


Data Science Field/Term Diagram: Source: Ryan Urbanowicz, PhD
University of Pennsylvania, Philadelphia PA, 19104

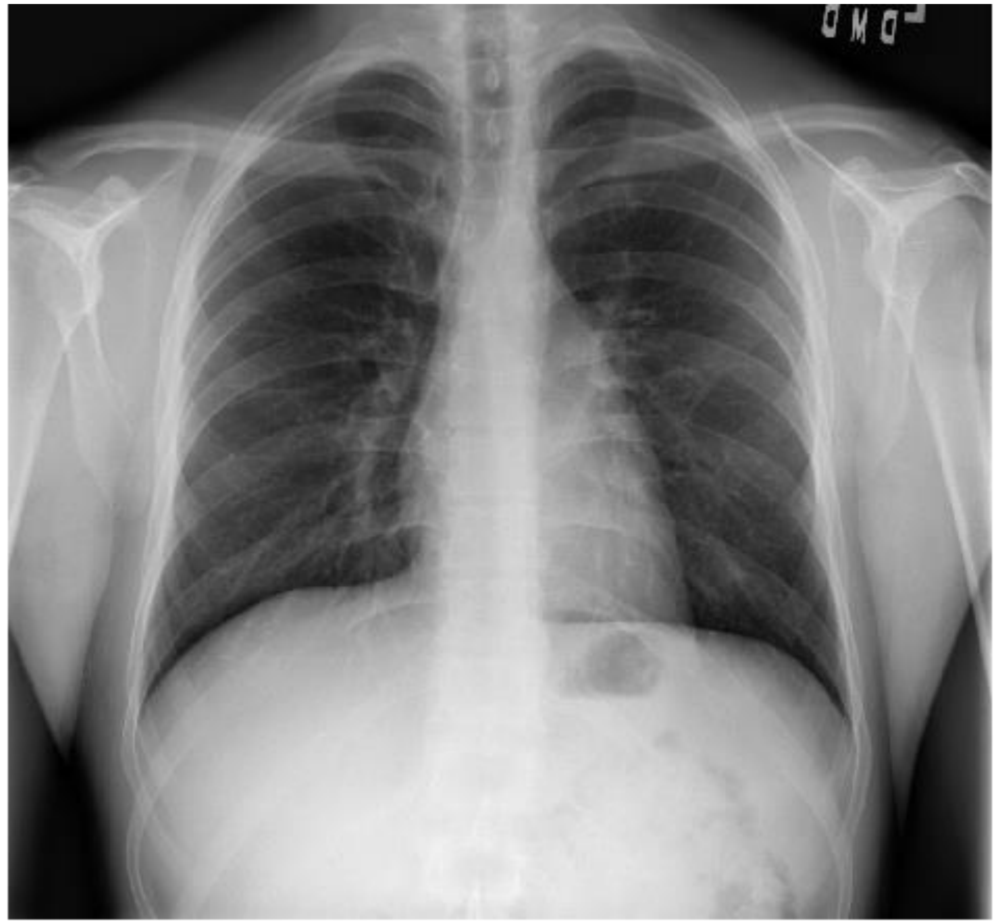
Science for Communities

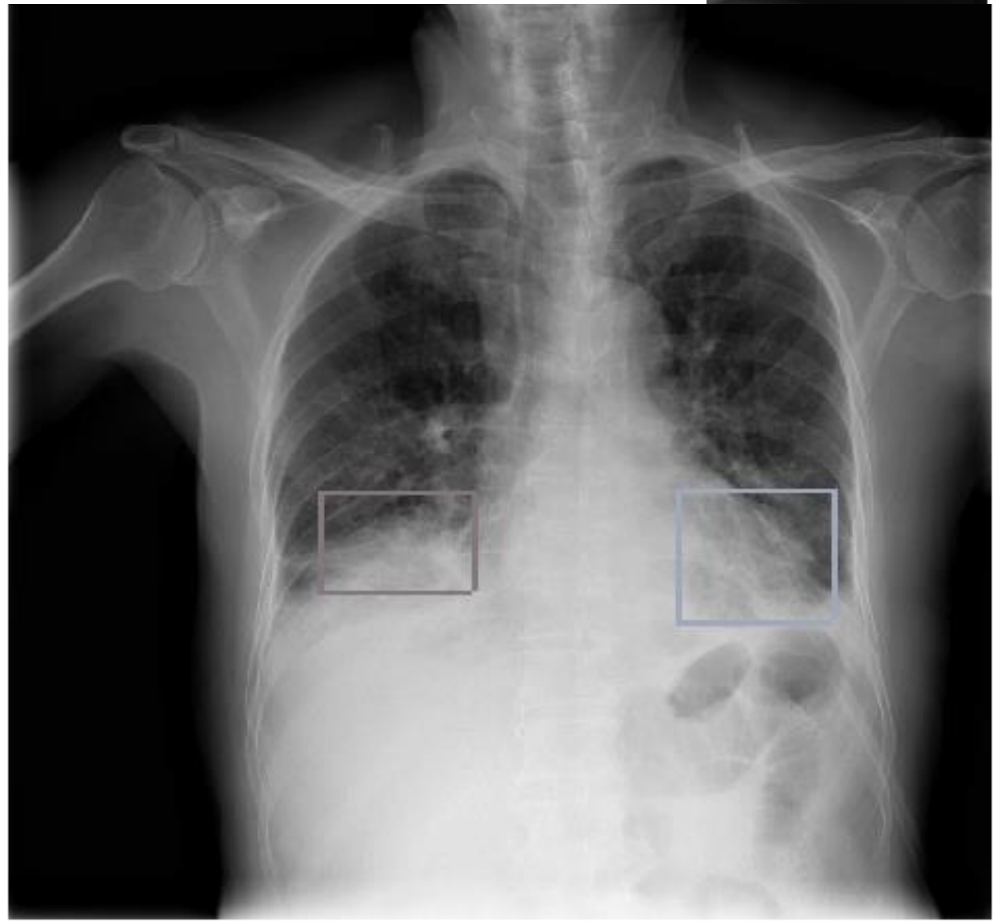
https://twitter.com/docurbs/status/1007375834347376642 | Ryan Urbanowicz

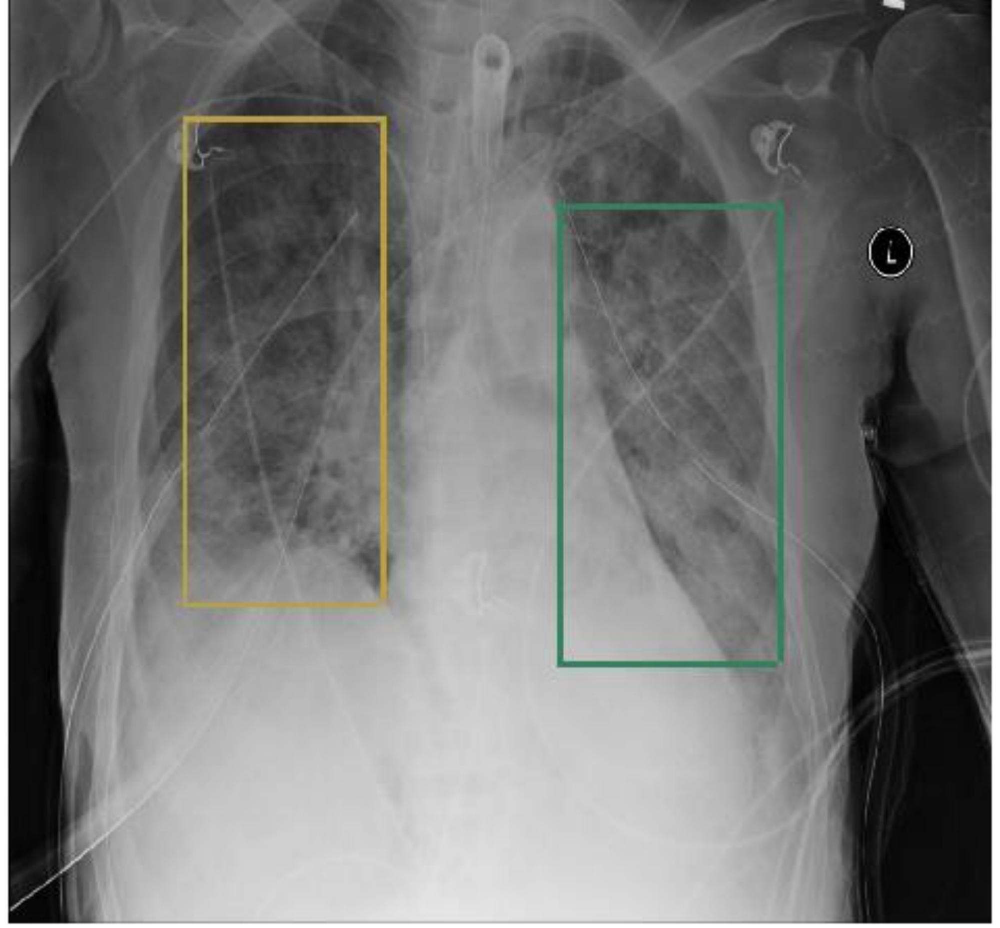
Data to wisdom

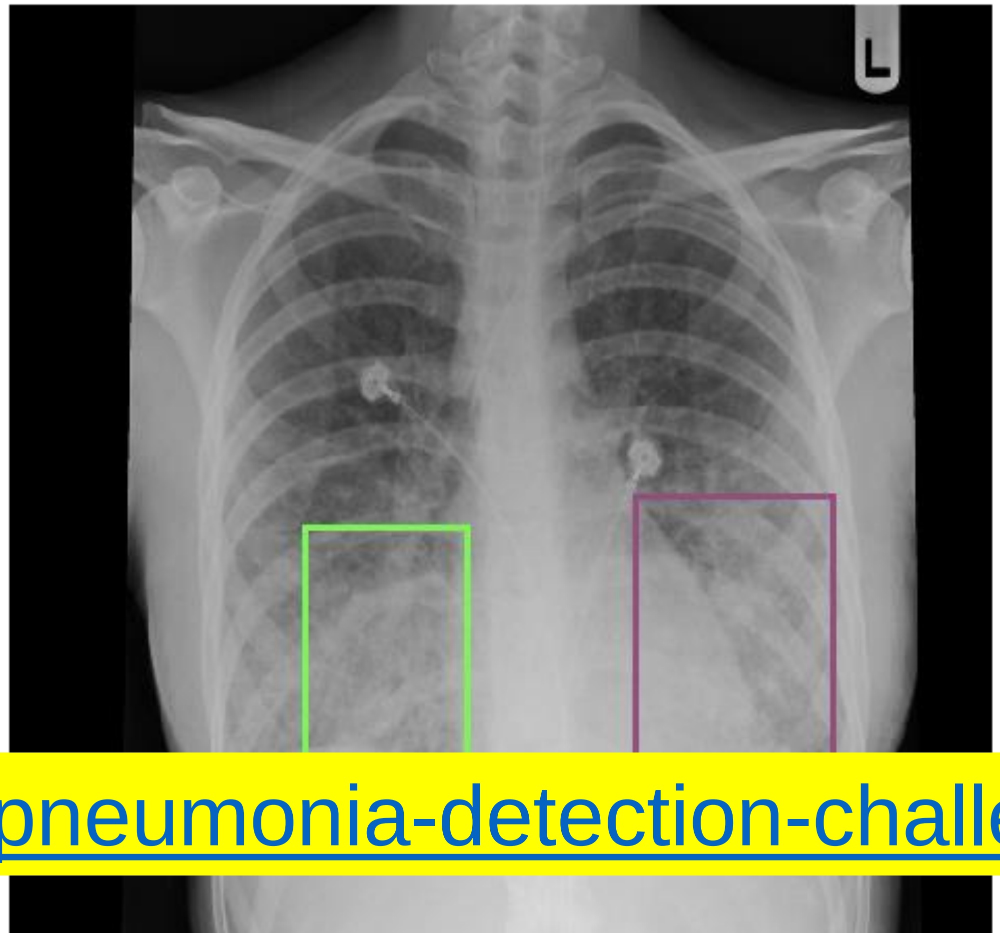


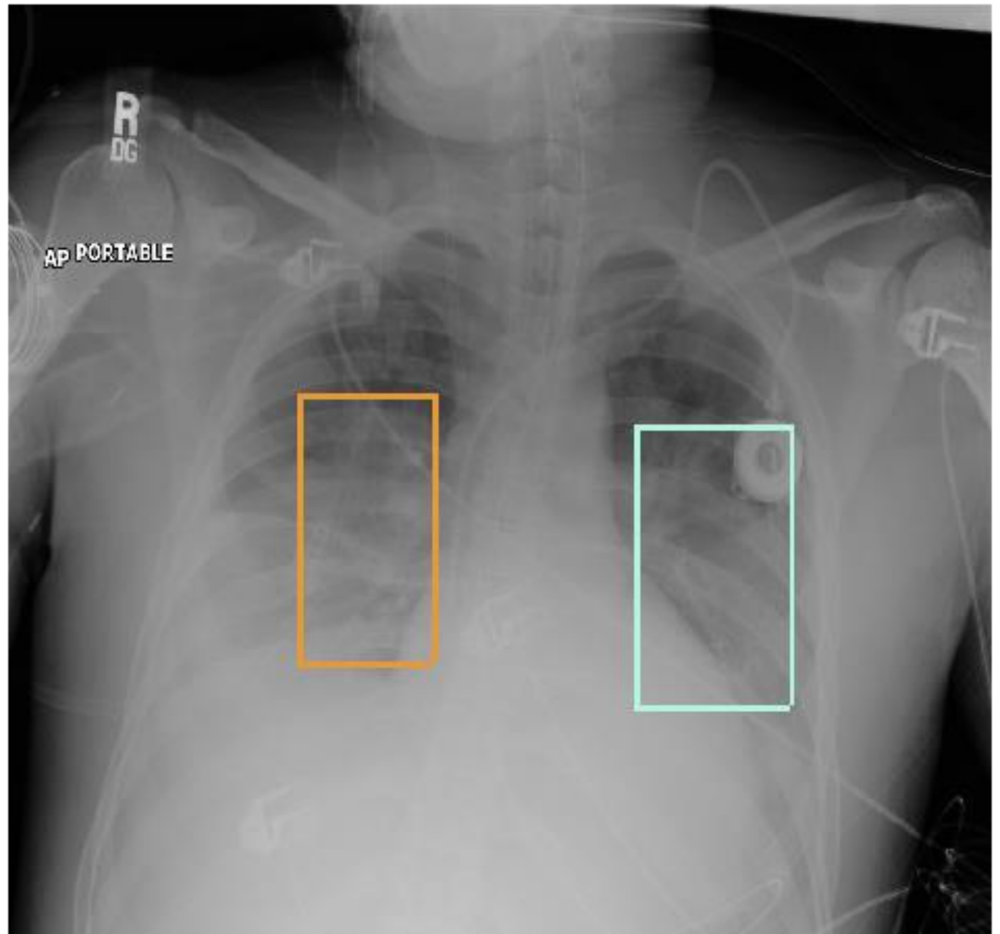
Science for Communities





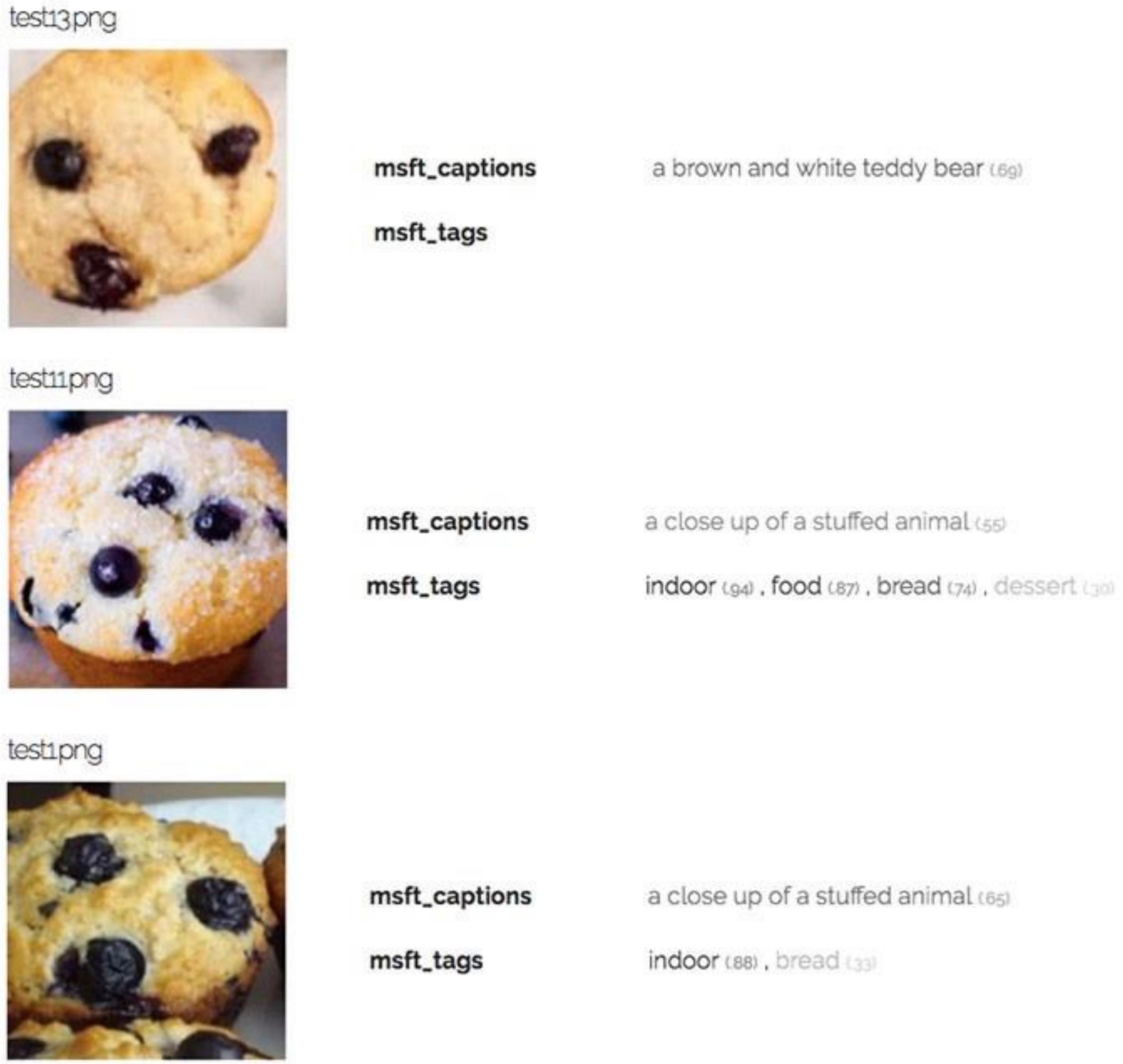






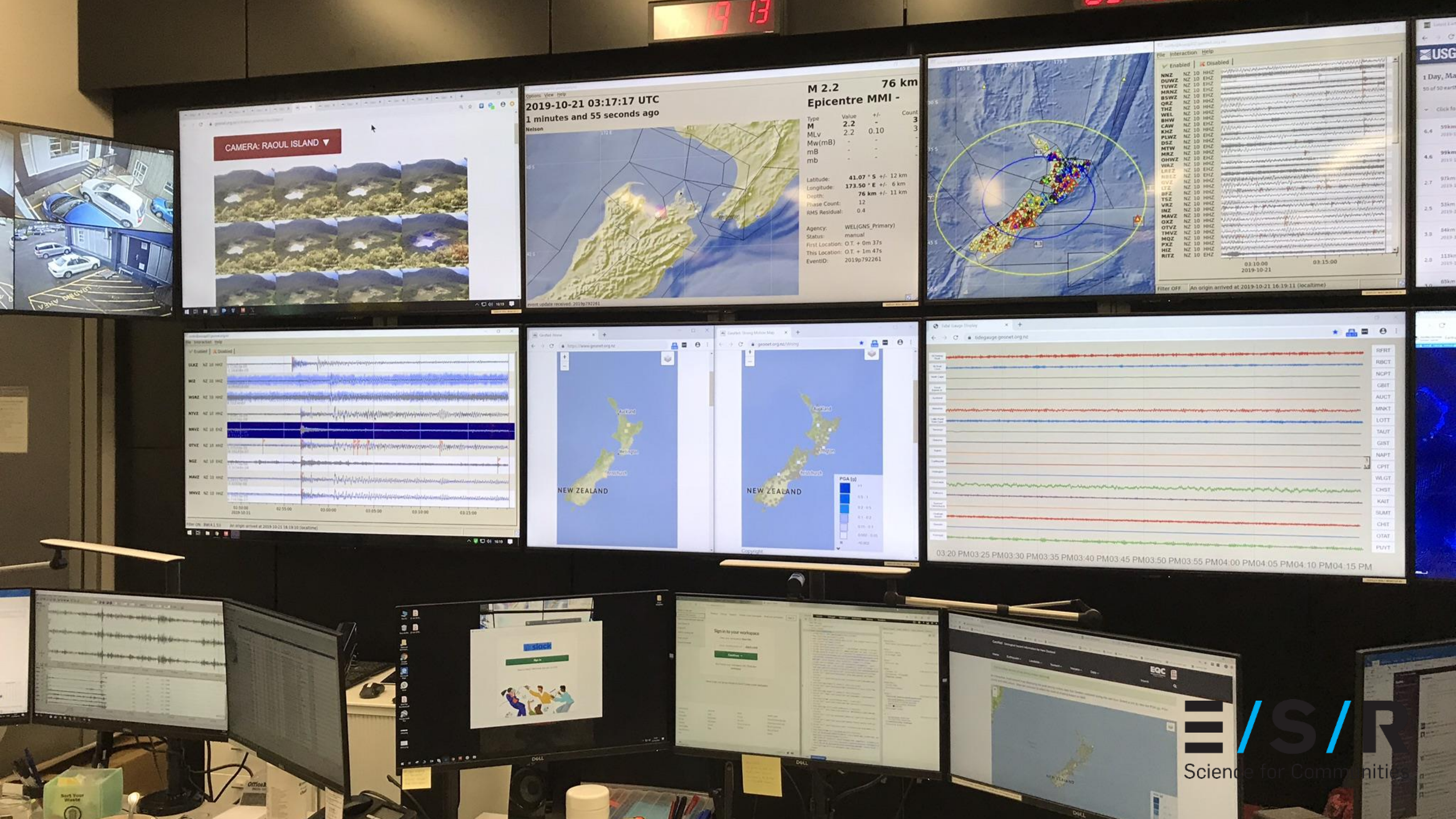
https://www.kaggle.com/c/rsna-pneumonia-detection-challenge

Predictions



https://www.freecodecamp.org/news/chihuahua-or-muffin-my-search-for-the-best-computer-vision-api-cbda4d6b425d/

E/S/R
Science for Communities



E/S/R
Science for Communities

Coronavirus COVID-19 Global Cases by Johns Hopkins CSSE

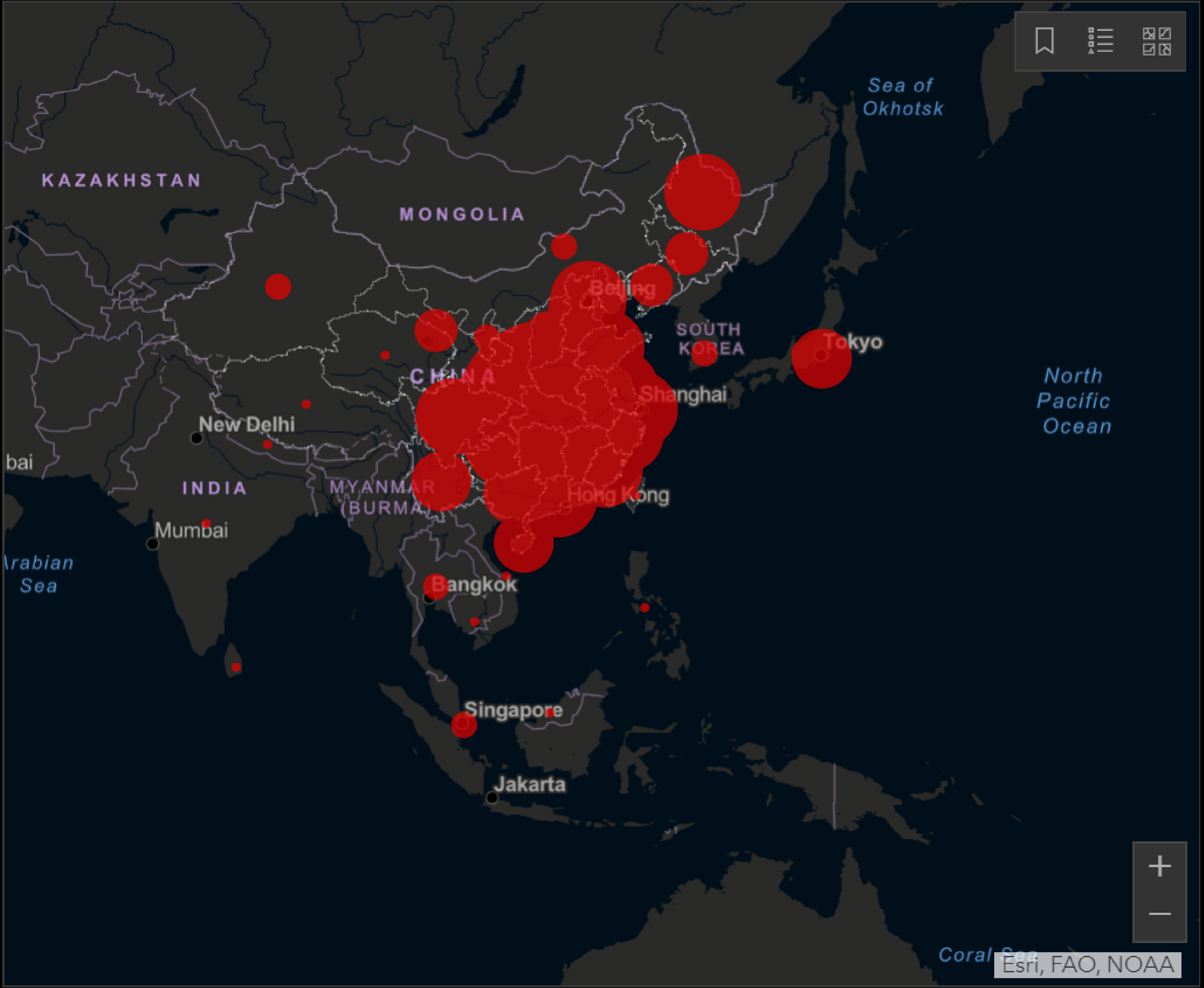
Total Confirmed
45,206

Confirmed Cases by Country/Region

44,687 Mainland China
175 Others
50 Hong Kong
47 Singapore
33 Thailand
28 South Korea
28 Japan
18 Malaysia
18 Taiwan
16 Germany
15 Australia
15 Vietnam
13 US

Country/Region

Last Updated at (M/D/YYYY)
2/13/2020 6:03:03 a.m.



Visualization: JHU CSSE. Automation Support: Esri Living Atlas team.
Data sources: WHO, CDC, ECDC, NHC and DXY. Read more in this blog. Contact US.
GitHub: Here. Google Sheet: Here. Time series table: Here. Feature layer: Here.
Point level: City level - US, Canada and Australia; Province level - China; Country level - other countries.

Total Deaths
1,118

1,068 deaths
Hubei Mainland China

8 deaths
Heilongjiang Mainland China

8 deaths
Henan Mainland China

4 deaths
Anhui Mainland China

4 deaths
Hainan Mainland China

3 deaths
Beijing Mainland China

3 deaths
Chongqing Mainland

Total Recovered
5,123

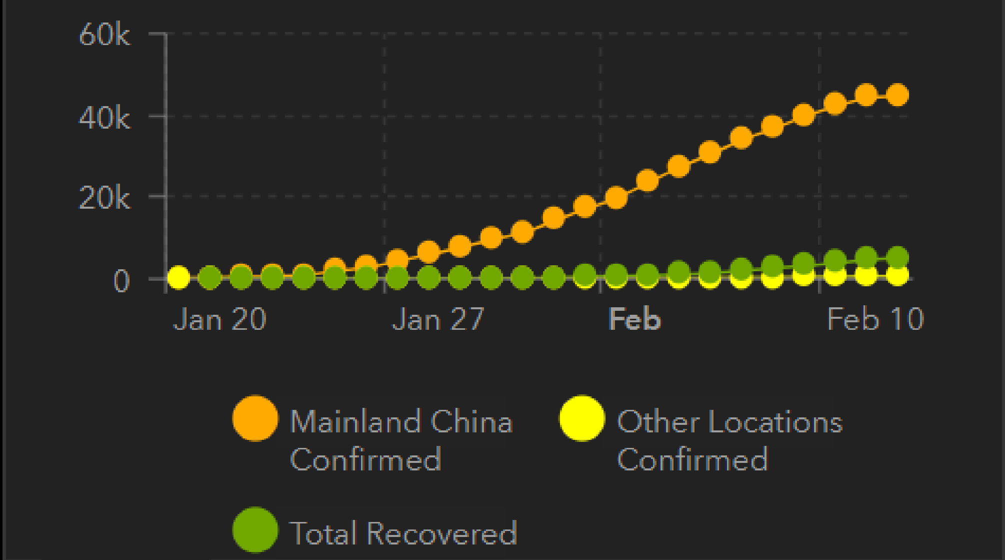
2,686 recovered
Hubei Mainland China

321 recovered
Zhejiang Mainland China

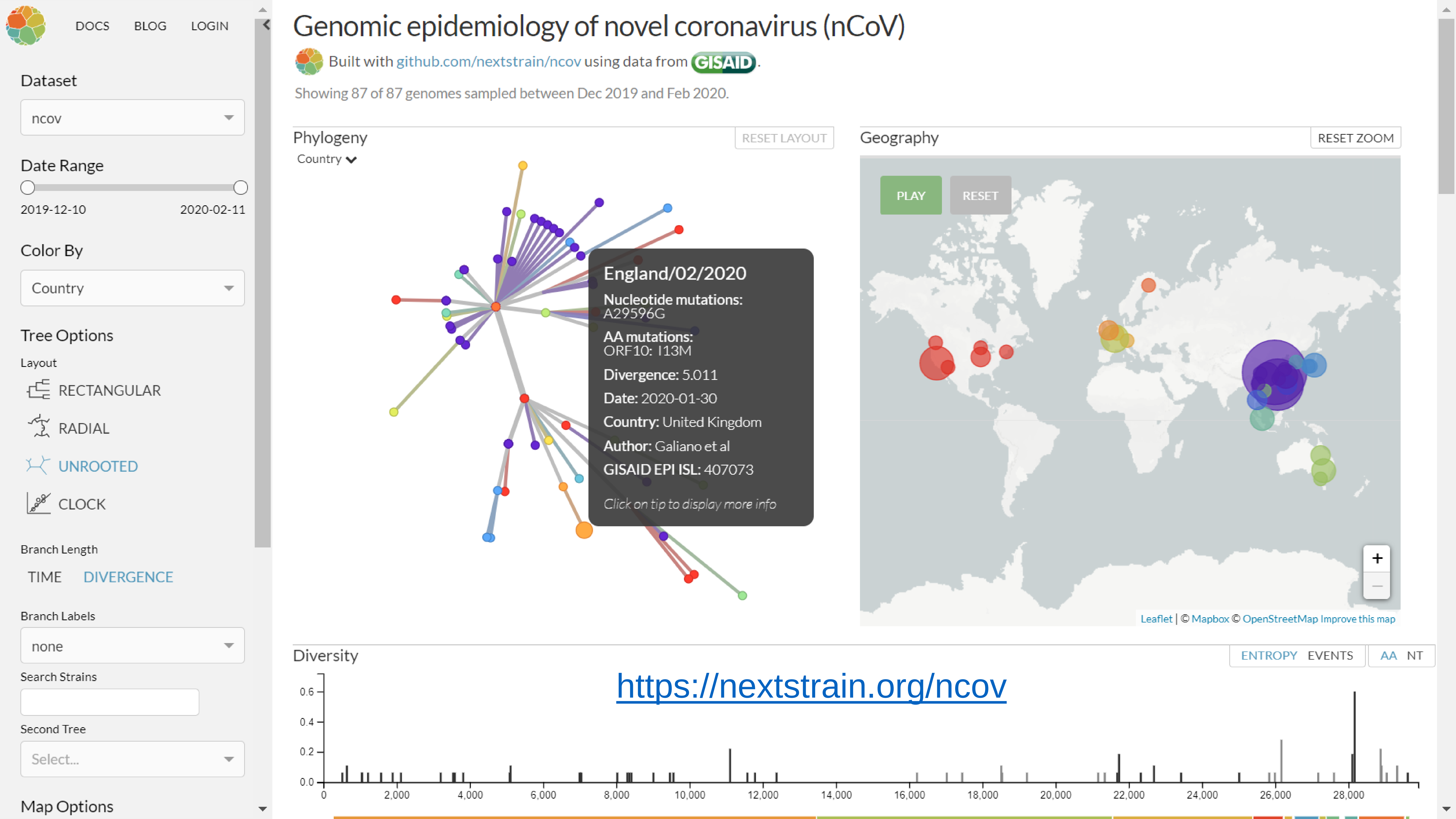
304 recovered
Hunan Mainland China

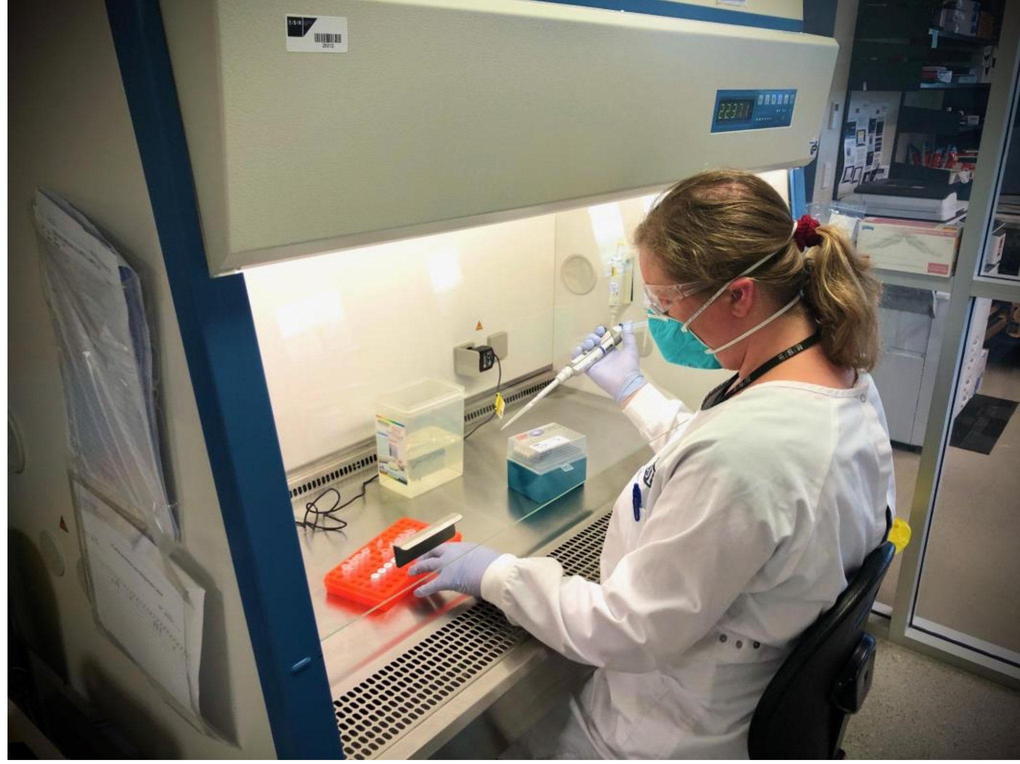
275 recovered
Guangdong Mainland China

246 recovered
Henan Mainland China

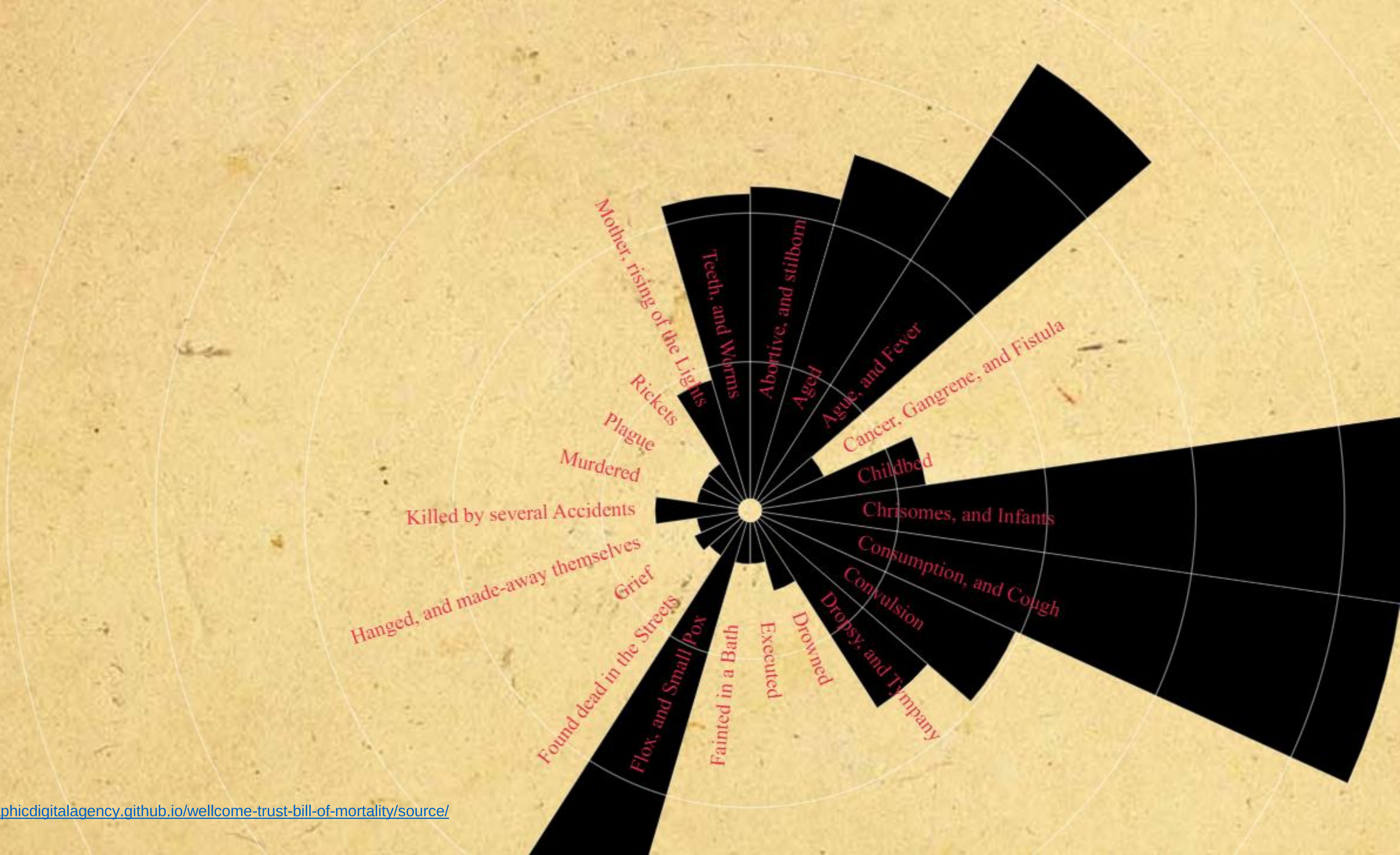


Actual
Logarithmic

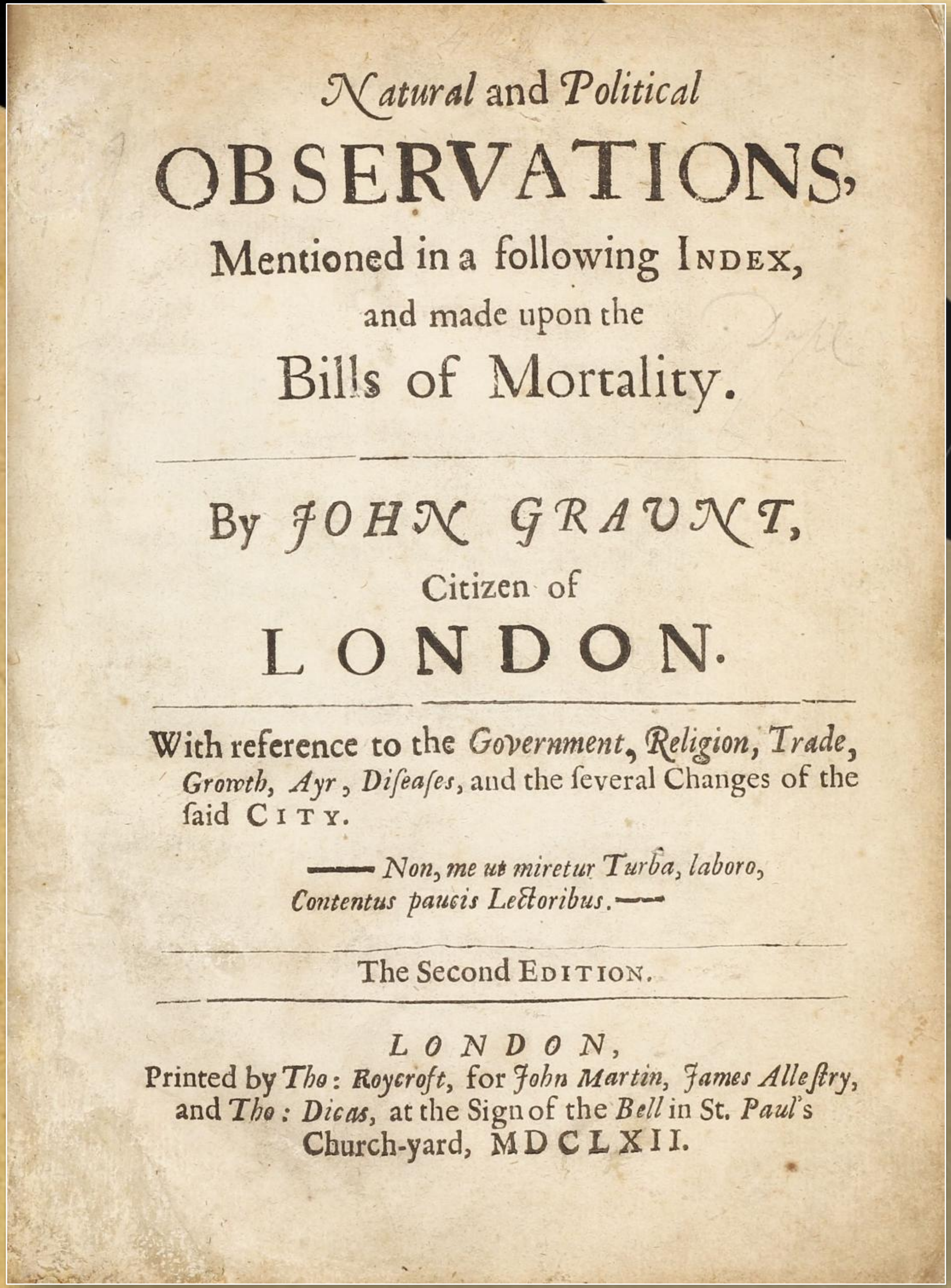

DOCS
BLOG
LOGIN
Dataset
ncov
Date Range
2019-12-10
2020-02-11
Color By
Country
Tree Options
Layout
RECTANGULAR
RADIAL
UNROOTED
CLOCK
Branch Length
TIME
DIVERGENCE
Branch Labels
none
Search Strains
Second Tree
Select...
Map Options
Genomic epidemiology of novel coronavirus (nCoV)
Built with github.com/nextstrain/ncov using data from GISAID.
Showing 87 of 87 genomes sampled between Dec 2019 and Feb 2020.
Phylogeny
RESET LAYOUT
Country
England/02/2020
Nucleotide mutations: A29596G
AA mutations: ORF10: I13M
Divergence: 5.011
Date: 2020-01-30
Country: United Kingdom
Author: Galiano et al
GISAID EPI ISL: 407073
Click on tip to display more info
Geography
RESET ZOOM
PLAY
RESET
Leaflet | © Mapbox © OpenStreetMap Improve this map
Diversity
ENTROPY
EVENTS
AA
NT
https://nextstrain.org/ncov
0.6
0.4
0.2
0.0
0
2,000
4,000
6,000
8,000
10,000
12,000
14,000
16,000
18,000
20,000
22,000
24,000
26,000
28,000




E/S/R
Science for Communities



http://graphicdigitalagency.github.io/wellcome-trust-bill-of-mortality/source/



Natural and Political
OBSERVATIONS,
Mentioned in a following INDEX,
and made upon the
Bills of Mortality.

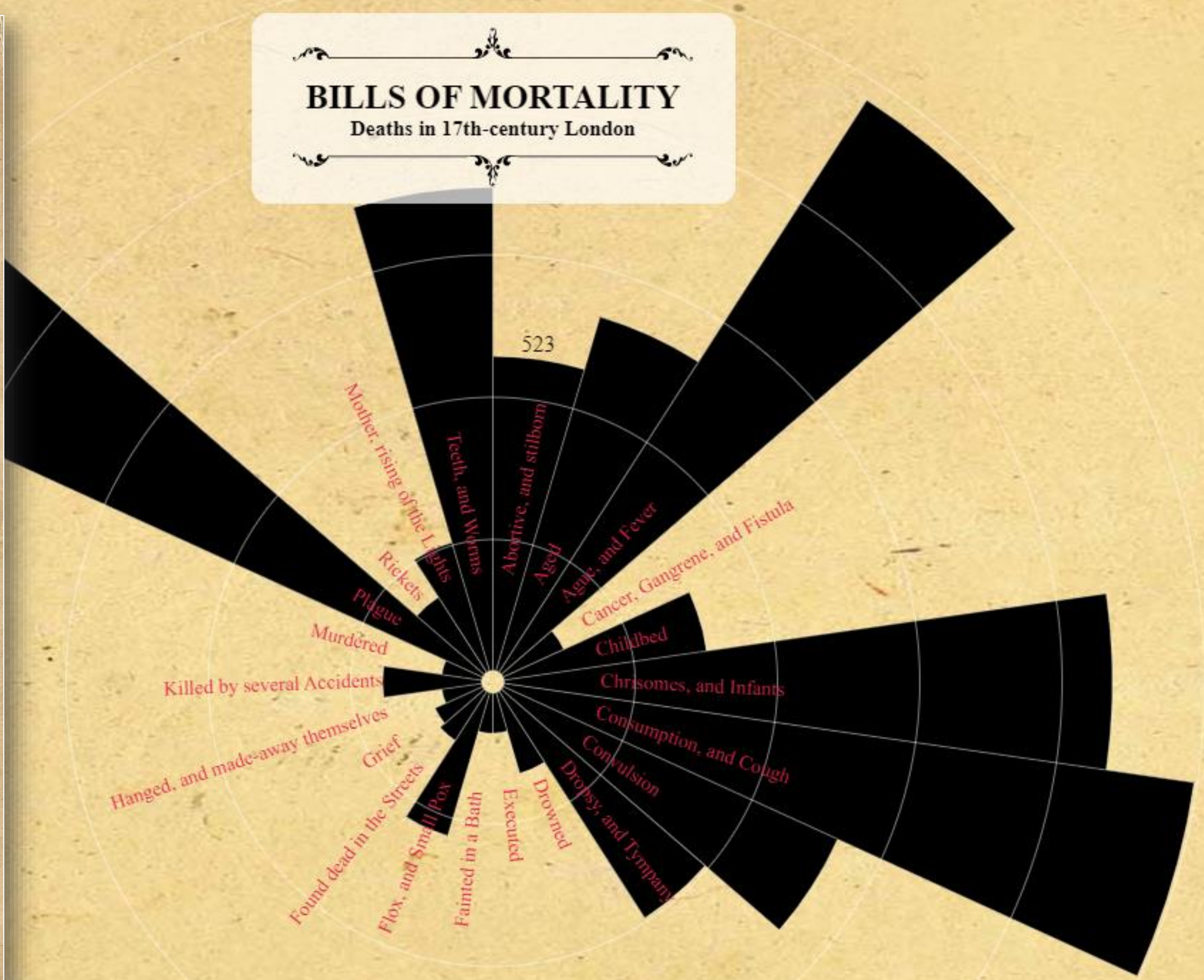
By JOHN GRAUNT,
Citizen of
LONDON.

With reference to the *Government*, *Religion*, *Trade*, *Growth*, *Ayr*, *Diseases*, and the several Changes of the said CITY.

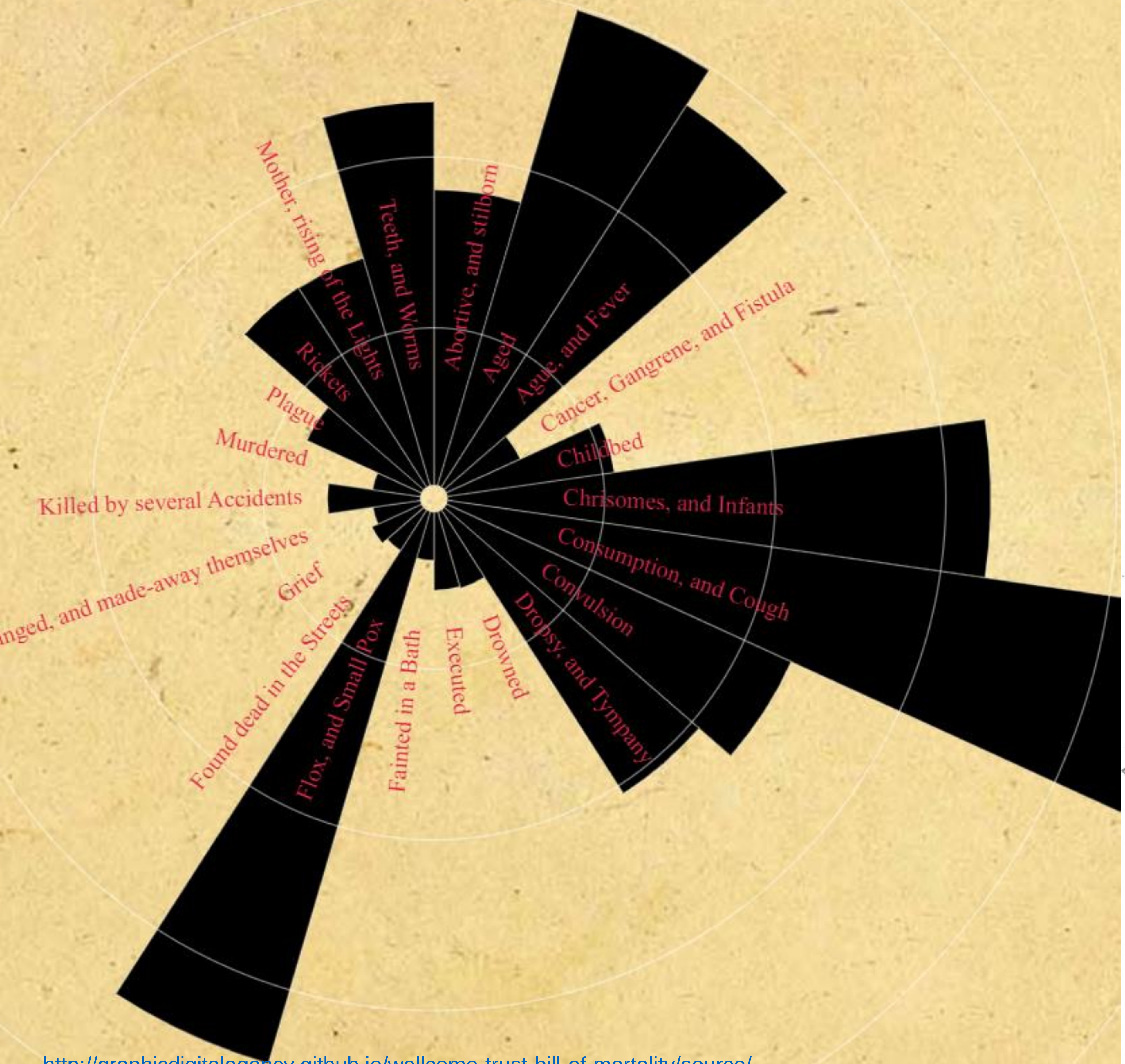
—— *Non, me ut miretur Turba, laboro,*
Contentus paucis Lectoribus. ——

The Second EDITION.

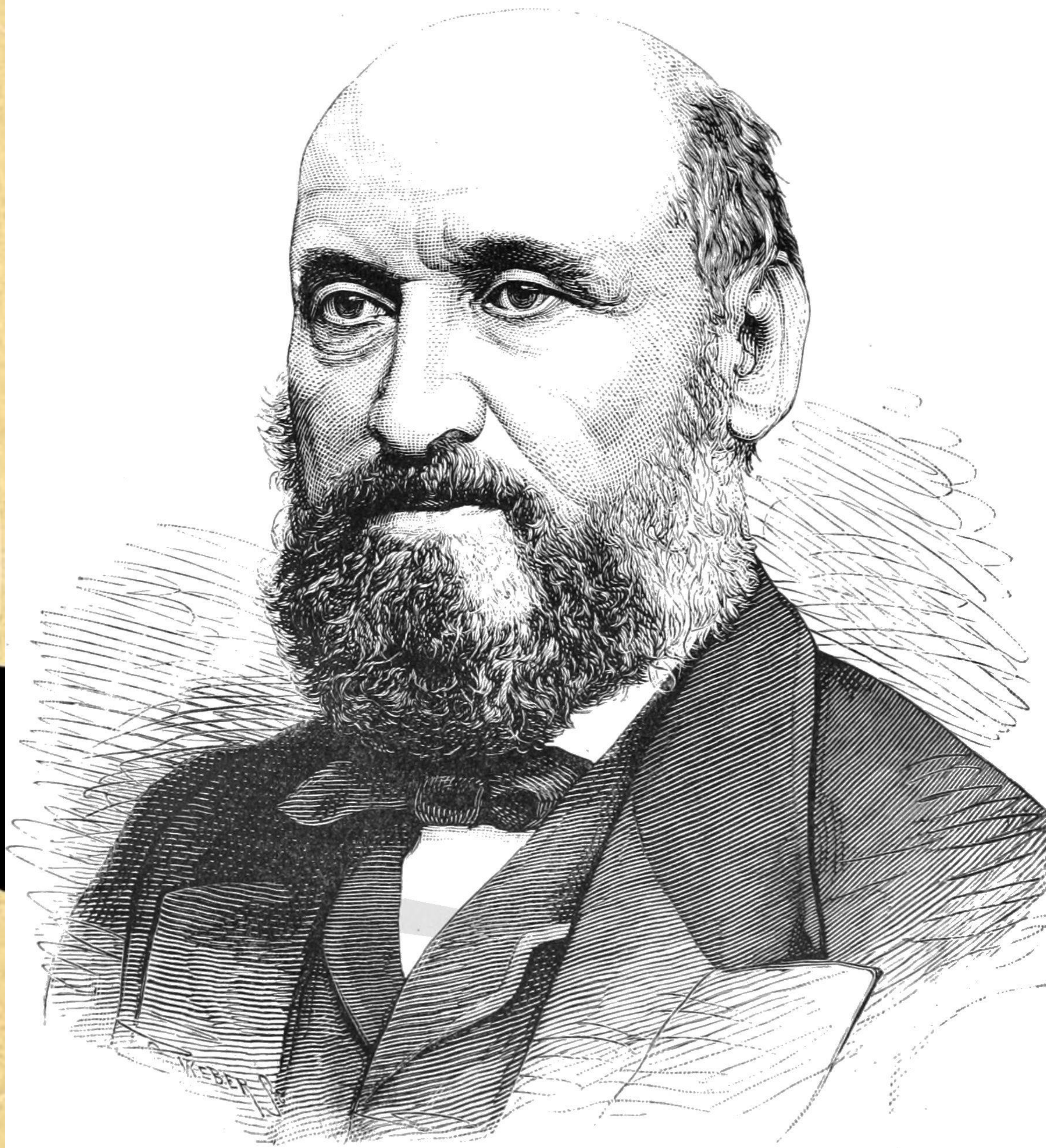
LONDON,
Printed by *Tho: Roycroft*, for *John Martin*, *James Allestry*, and *Tho: Dicas*, at the Sign of the *Bell* in St. *Paul's* Church-yard, MDCLXII.



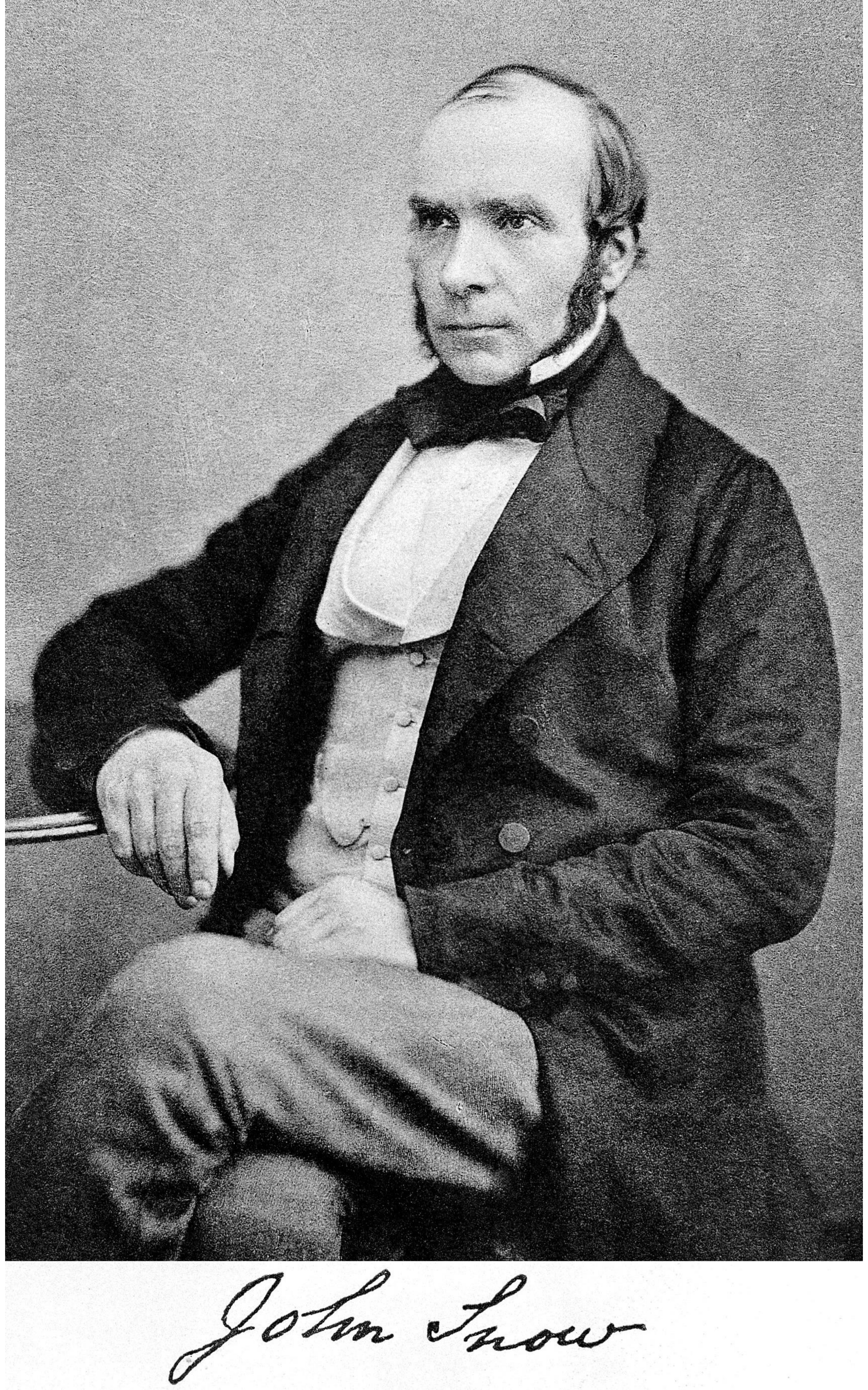
http://graphicdigitalagency.github.io/wellcome-trust-bill-of-mortality/source/



http://graphicdigitalagency.github.io/wellcome-trust-bill-of-mortality/source/



William Farr

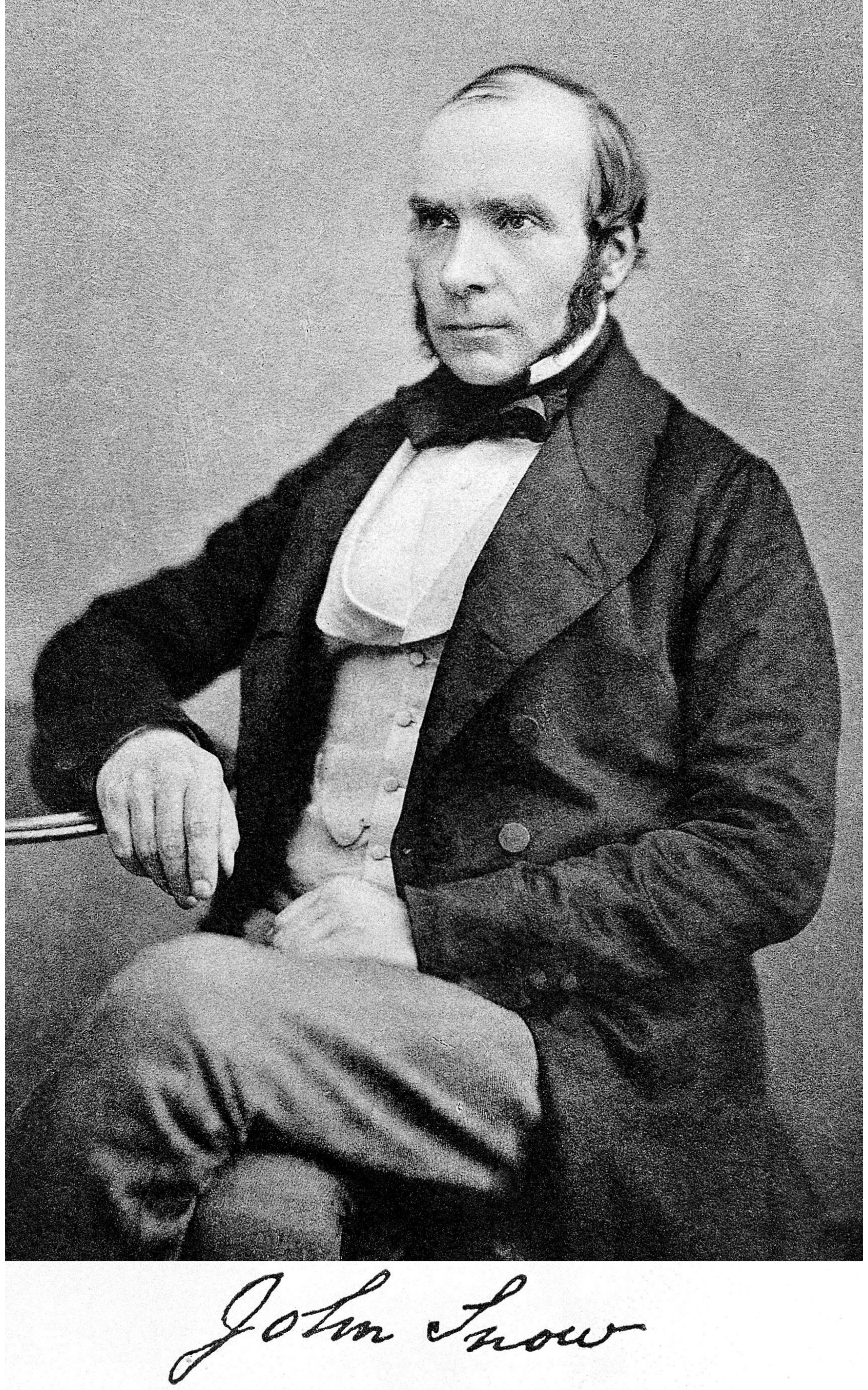

John Snow

MISTAKING CAUSE FOR EFFECT.

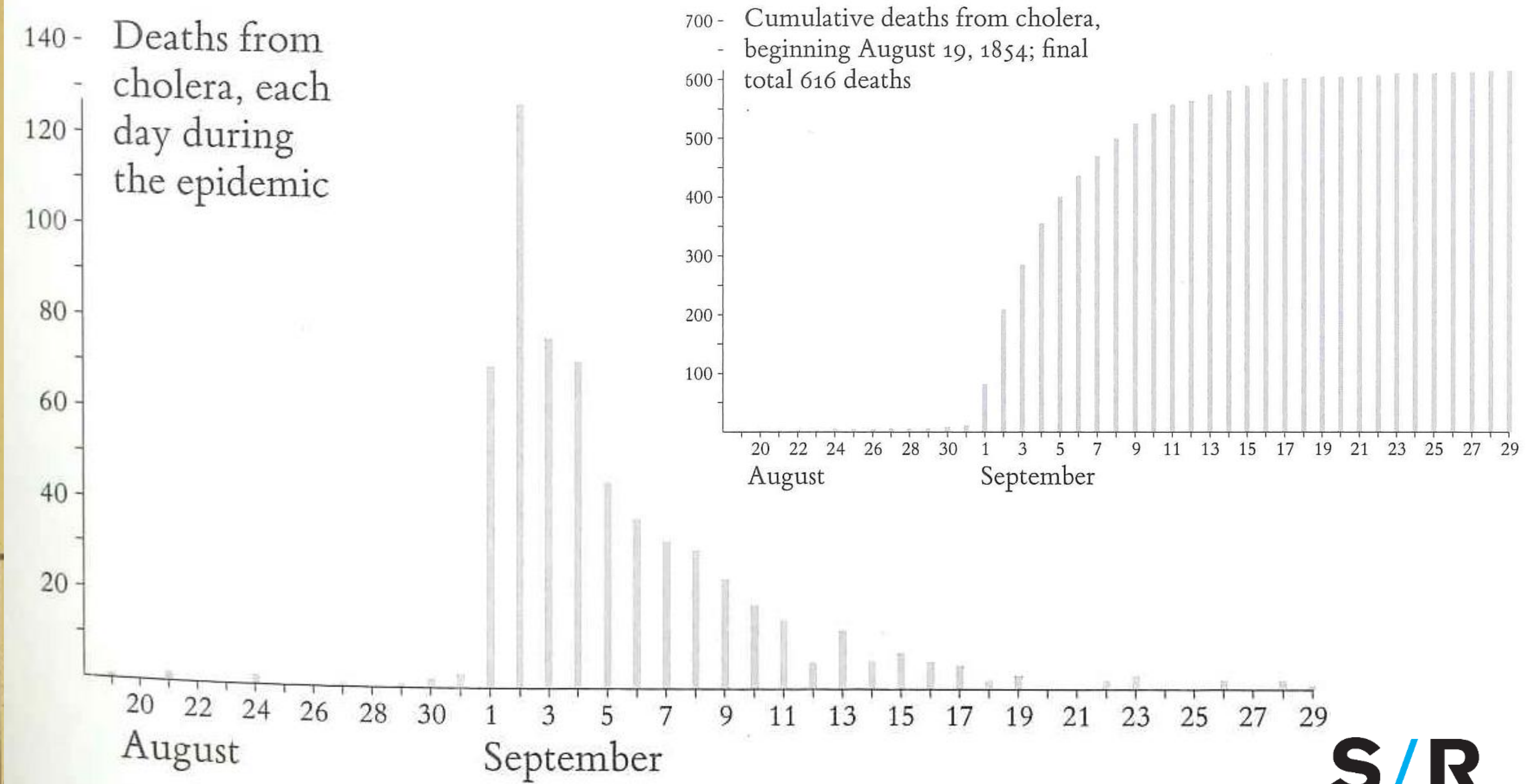
Boy. "I SAY, TOMMY, I'M BLOW'D IF THERE ISN'T A MAN A TURNING ON THE CHOLERA."

Caricature: Cholera.

E/S/R
Science for Communities



E/S/R
Science for Communities

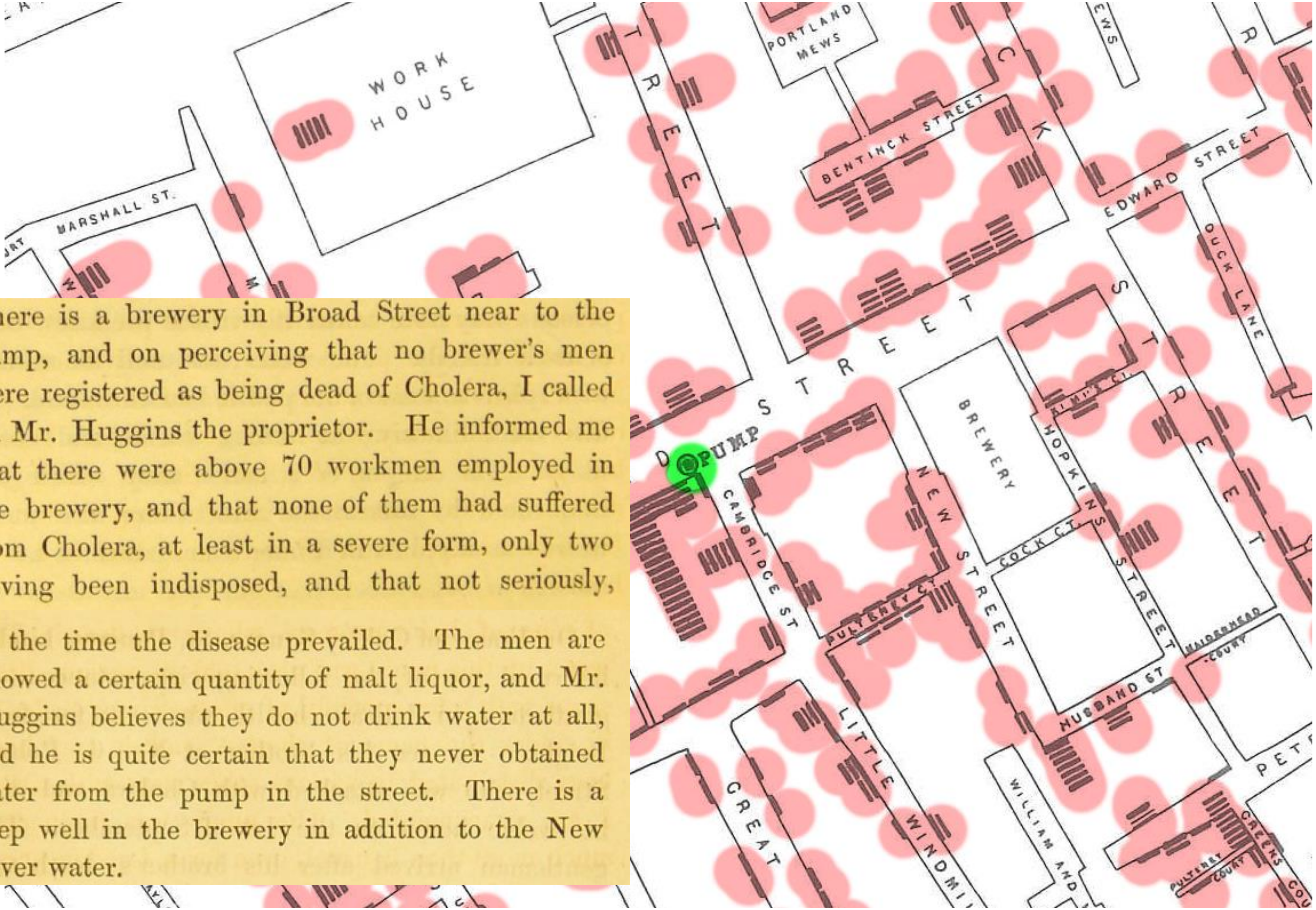


or Communities

Occupations.	Males. Adults.	Males. Sons.	Females. Spinsters, Wives, Widows.	Females. Daughters.	Total.
Postmaster (retired),	1	..	..	..	1
Government Clerk,	1	..	..	..	1
Police,	2	..	..	1	3
Fireman,	..	..	..	1	1
Chelsea Pensioner,	1	..	..	..	1
Solicitor,	..	1	1	..	2
Surgeon,					
Dentist,					
Druggist,					
Artist,					
Schoolmast					
Governess,					
Lodging Ho					
Eating and					
Domestic S					
Coachmen,					
Charwomen,	..	1	4	..	5
Nurse,	..	..	1	..	1
Laundress,	..	..	1	..	1
Hairdresser,	1	1	2	1	5
Hatter,	1	..	..	..	1
Tailor,	40	12	17	9	78
Shoemaker,	28	8	8	3	47

Ages	0—10.	10—20.	20—30.	30—40.	40—50.	50—60.	60—70.	70—80.	80—90.	0—90.
Males	79	32	48	50	47	16	19	4	2	297
Females	56	33	40	51	61	51	30	10	1	333
Total	135	65	88	101	108	67	49	14	3	630


OXFORD STREET
REGENT CIRCUS
PRINCES STREET
HANOVER STREET
ARGYLL STREET
LITTLE ARGYLL ST
GREAT MARLBOROUGH STREET
MARLBOROUGH MEWS
BLENHEIM MEWS
BLENHEIM STREET
QUEEN STREET
PUMP
POLAND STREET
NOEL STREET
BERWICK STREET
WARDOUR STREET
WARDOUR MEWS
LITTLE CHAPEL STREET
CARLISLE STREET
CROWN COURT
DEAN STREET
SOHO SQUARE
PORTLAND STREET
PORTLAND MEWS
WORK HOUSE
ARGYLL PLACE
LITTLE MARLBOROUGH ST
MARSHALL ST.
TYLER COURT
TYLER STREET
WEST STREET
MARSHALL STREET
DUFOURS PLACE
BENTINCK STREET
EDWARD STREET
DUCK LANE
ST ANNS COURT
RICHMOND BUILDINGS
RICHMOND MEWS
BREWERY YARD
BATEMANS BUILDINGS
FRITH STREET
QUEEN STREET
CREEK STREET
BROAD STREET
CAMBRIDGE ST
BREWERY
NEW STREET
HOPKINS STREET
COCK CT.
PULTENEY CT.
HUSBAND ST
PETER STREET
WILLIAM AND MARY YARD
GREENS COURT
WALKERS COURT
MEARDS COURT
MEARDS STREET
DEAN STREET
OLD COMPTON
CHURCH ST
CARNABY STREET
SOUTH ROW
MARLBOROUGH ROW
CROSS ST
CROSS STREET
KING STREET
POLLEN ST.
MASONS ARMS YARD
MADDOX STREET
GEORGE STREET
MILL STREET
MILL ST.
OLD BURLINGTON MEWS
CONDUIT STREET
REGENT STREET
CHAPEL PLACE
NAYLORS YARD
SILVER STREET
UP. JAMES ST
UP. JOHN ST
BRIDLE STREET
GREAT PULTENEY STREET
LITTLE WINDMILL STREET
LITTLE PULTENEY STREET
GREAT CROWN CT
CT. CROWN CT.
UP. RUPERT ST
QUEENS HEAD COURT
GOLDEN SQUARE
LOW. JAMES ST
BEAK ST.
WARWICK STREET
NEW BURLINGTON STREET
NEW BURLINGTON MEWS
BOYLE ST.
HEDDON COURT
SAVILE
ARCHER STREET
PLOUGH YARD
RUPERT STREET
RICHMOND ST
GEORGE YD
KING STREET
PRINCES STREET
GERRARD STREET
MACCLESFIELD STREET
SMITHS COURT
HAM YARD
GREAT WINDMILL



There is a brewery in Broad Street near to the pump, and on perceiving that no brewer's men were registered as being dead of Cholera, I called on Mr. Huggins the proprietor. He informed me that there were above 70 workmen employed in the brewery, and that none of them had suffered from Cholera, at least in a severe form, only two having been indisposed, and that not seriously, at the time the disease prevailed. The men are allowed a certain quantity of malt liquor, and Mr. Huggins believes they do not drink water at all, and he is quite certain that they never obtained water from the pump in the street. There is a deep well in the brewery in addition to the New River water.


Broadwick Street Pump | B.Weber

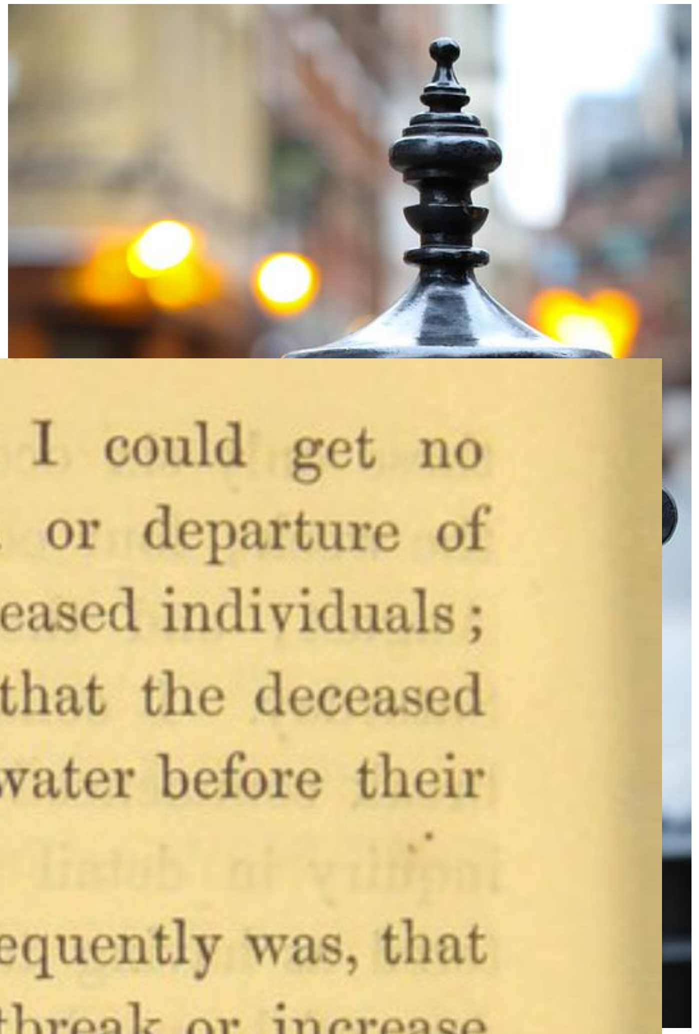
E/S/R
Science for Communities

On proceeding to the spot, I found that nearly all the deaths had taken place within a short distance of the pump in Broad Street. There were only ten deaths in houses situated decidedly nearer to another street-pump. In five of these cases, the families of the deceased persons told me that they always sent to the pump in Broad Street, as they preferred the water to that of the pump which was nearer. In three other cases, the deceased were children who went to school near the pump in Broad Street. Two of them were known to have drunk the water, and the parents of the third think it probable that it did so. The other two deaths, beyond the district which this pump supplies, represent only the amount of mortality from Cholera that was occurring before the eruption took place.

With regard to the 73 deaths occurring in the locality belonging as it were to the pump, there were 61 instances in which I was informed that the deceased persons used to drink the water from the pump in Broad Street, either constantly or occasionally. In six instances I could get no information, owing to the death or departure of every one connected with the deceased individuals; and in six cases I was informed that the deceased persons did not drink the pump water before their illness.

The result of the inquiry consequently was, that there had been no particular outbreak or increase of Cholera, in this part of London, except among the persons who were in the habit of drinking the water of the above-mentioned pump well.

I had an interview with the Board of Guardians of St. James's parish on the evening of Thursday, 7th September, and represented the above circumstances to them. In consequence of what I said, the handle of the pump was removed on the following day.



What do we mean by “Data Science”?

Occupations.	Males. Adults.	Males. Sons.	Females. Spinsters, Wives, Widows.	Females. Daughters.	Total.
Postmaster (retired),	1	..	..	..	1
Government Clerk,	1	..	..	..	1
Police,	2	..	..	1	3
Fireman,	..	..	..	1	1
Chelsea Pensioner,	1	..	..	..	1
Solicitor,	..	1	1	..	2
Surgeon,	1	..	..	..	1
Dentist,	1	..	..	..	1
Druggist,	..	1	..	..	1
Artist,	1	..	1	..	2
Schoolmaster,	..	..	1	..	1
Governess,	..	..	1	..	1
Lodging House Keeper,	..	..	2	..	2
Eating and Coffee House Keeper,	1	..	..	1	2
Domestic Servants,	2	..	28	2	32
Coachmen,	1	1	1	1	4
Charwomen,	..	1	4	..	5
Nurse,	..	..	1	..	1
Laundress,	..	..	1	..	1
Hairdresser,	1	1	2	1	5
Hatter,	1	..	..	..	1
Tailor,	40	12	17	9	78
Shoemaker,	28	8	8	3	47
Undertaker,	1	1	1	..	3
Dressmakers, including Staymakers and Waistcoat Makers,	..	..	15	..	15
Straw Hat Maker,	..	..	1	..	1
Commercial Traveller,	..	..	1	..	1
Pawnbroker,	2	..	..	..	2
Marine Store Dealer,	..	..	1	..	1
Livery Stable Keeper,	2	..	..	..	2
Carman,	2	..	1	..	3
Warehouseman,	..	1	..	..	1
Shopman and Shopwoman,	1	..	1	..	2
Messengers and Porters,	15	6	2	5	28
Errand Boy,	..	1	..	..	1
Printer,	2	1	..	..	3
Compositor,	1	..	..	..	1
Bookbinder,	2	..	..	..	2
Stationer,	2	..	..	..	2
Piano-forte Maker,	3	1	..	1	5
Picture Dealer,	..	..	1	..	1
Engravers and Chasers,	4	..	1	..	5
Artificial Flower Makers,	..	..	2	..	2
Feather Manufacturers,	..	..	1	2	3
Dyer,	..	..	..	1	1
Draper,	..	..	1	..	1
Mattress Maker,	..	..	1	2	3
Brush Maker,	..	..	1	..	1
Carried forward,	119	36	99	30	284

Ages	0—10.	10—20.	20—30.	30—40.	40—50.	50—60.	60—70.	70—80.	80—90.	0—90.
Males	79	32	48	50	47	16	19	4	2	297
Females	56	33	40	51	61	51	30	10	1	333
Total	135	65	88	101	108	67	49	14	3	630

To improve **decision making**

by basing decisions on **insights**

extracted from **data**

E/S/R Science for Communities

2015 – PHE Public Health Data Science

Differences in the data

Previously

- Demography
- Closed observational and interventional epidemiology datasets
- Health surveys
- Surveillance
- Geographical and environmental
- Health Services

Public health data science *

- Electronic healthcare records
- Social media
- Open data
- ‘omics data
- Wearables and internet of things
- Mobile apps
- Citizen driven

* Public Health Data Science uses all those data on the left hand side.

Credit - Julian Flowers (PHE), Rob Aldridge (UCL)

ESR Science for Communities

Differences in ways of working

Previously

- Collation and description
- Excel and stats packages
- Static reports
- Manual processing
- Waterfall project
- User feedback
- Epidemiology + statistics
- Structured / small data
- Slow
- Costly

Public health data science *

- Prediction and prescription
- R / Python / PowerBI / Tableau
- Interactive reporting
- Automated processing
- Agile
- User need
- Epidemiology + system models + machine learning + programming
- Structured + unstructured + big data
- Faster
- Cheaper

Science for Communities

Credit - Rob Aldridge (UCL)

A working definition of

Public Health Data Science

- *“**Using data science tools and methods** to harness data to prevent disease, prolong life and promote human health through organized efforts and informed choices of society, organizations, public and private, communities and individuals”*

Credit - Rob Aldridge (UCL)

E/S/R Science for Communities

Get the data

Download indicator data and definitions

Profile: Local Authority Health Profiles

Data for County & UA (pre 4/19)

Data for County & UA (pre 4/19) in North East region

Indicator definitions

Domain: Health protection

Data for County & UA (pre 4/19)

Data for County & UA (pre 4/19) in North East region

Indicator definitions

Indicator: Excess winter deaths index

Data for County & UA (pre 4/19)

Data for County & UA (pre 4/19) in North East region

Indicator definition

Get the data with R

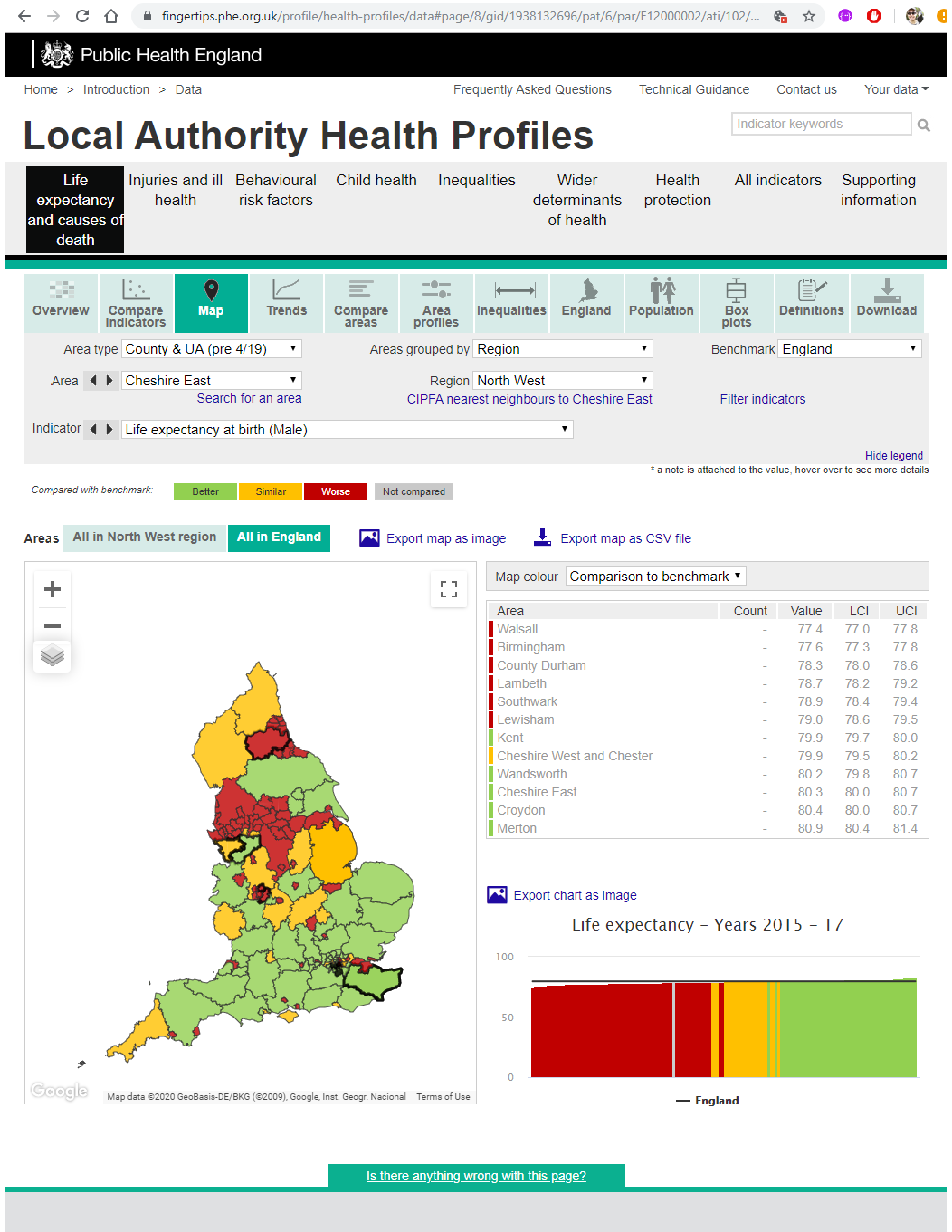
The fingertipsR package allows you to download public health data using R

Get the data with Python

The fingertips_py package allows you to download public health data using Python

Public Health Data API

The Fingertips API (Chrome or Firefox only) allows public health data to be retrieved in either JSON or CSV formats



Area profile

Download a detailed report of the data for

Report year 2017

County Durham

Findable

Accessible

Interoperable

Reusable

Science for Communities

2016

E/S/R
Science for Communities

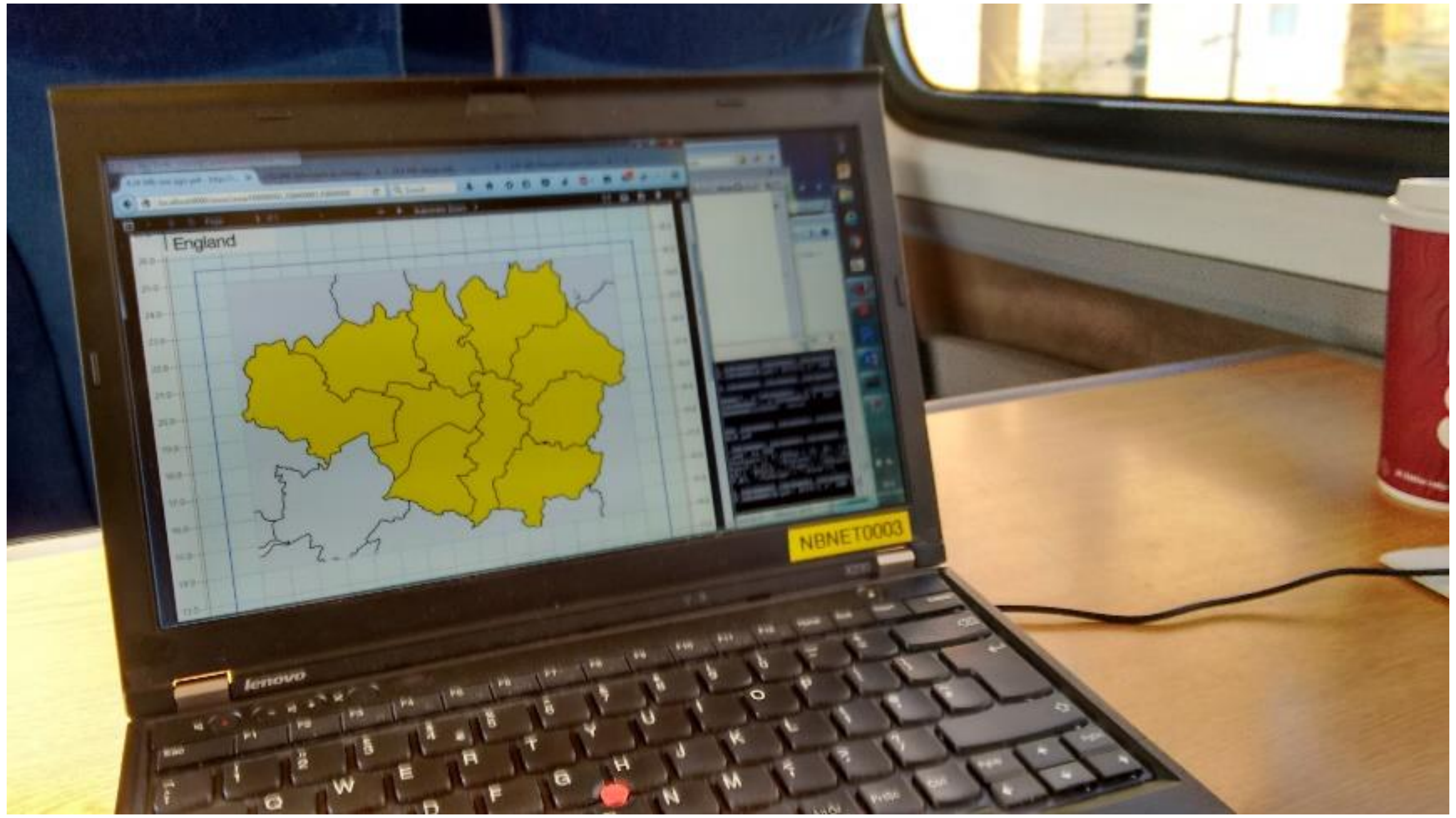
http://www.henry4school.fr/UK/history/brexit.htm

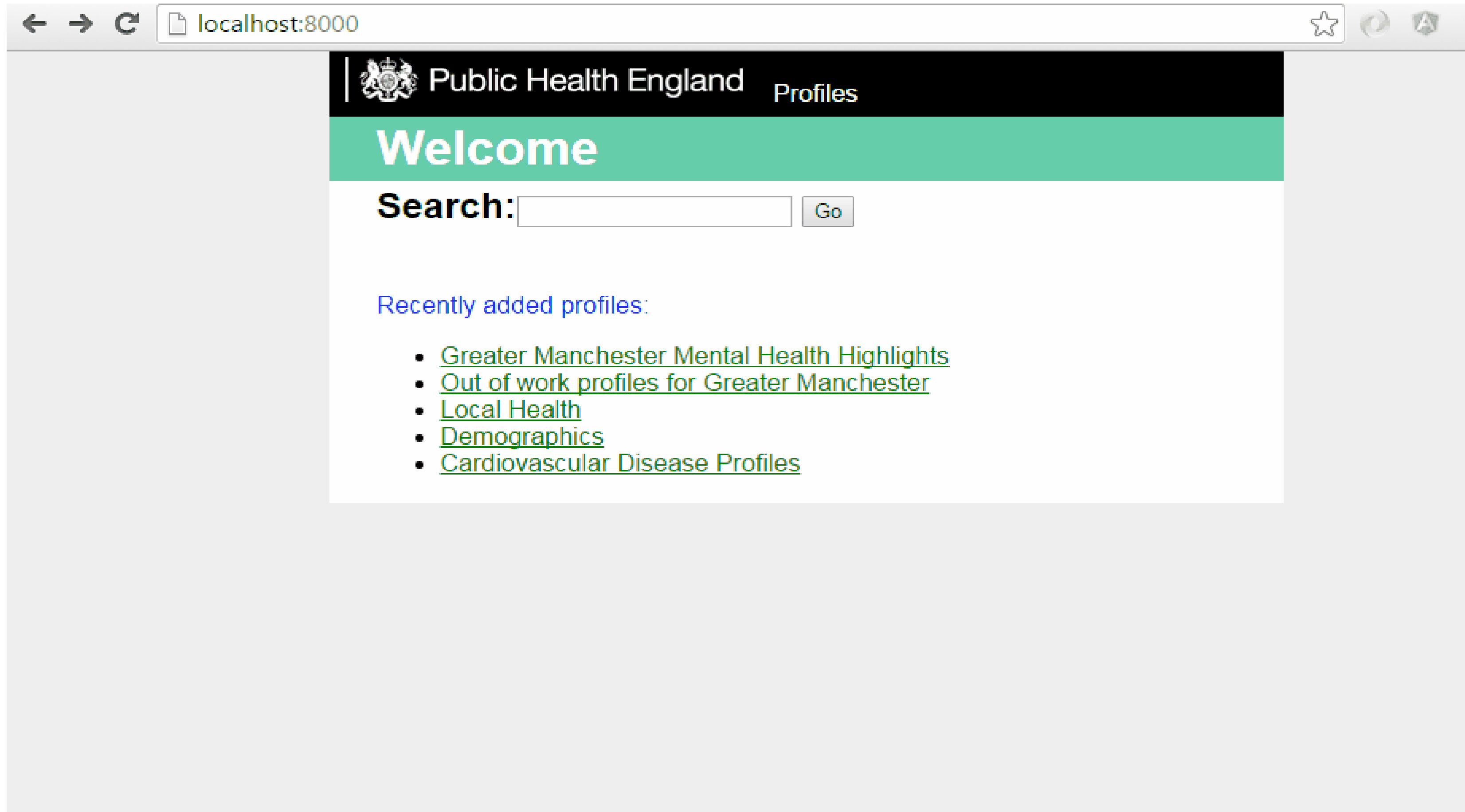
What's a data science accelerator?

- A data science project that tackles a business problem
- Access to required data
- Participant able to commit 1 day per week [3 months | 15 weeks] to the project
- Support from line manager and senior manager
- Coding experience is useful but not essential.

ESR Science for Communities

My data science accelerator experience





Science for Communities

Shared learning

Alan Lewis (Greater London Authority)

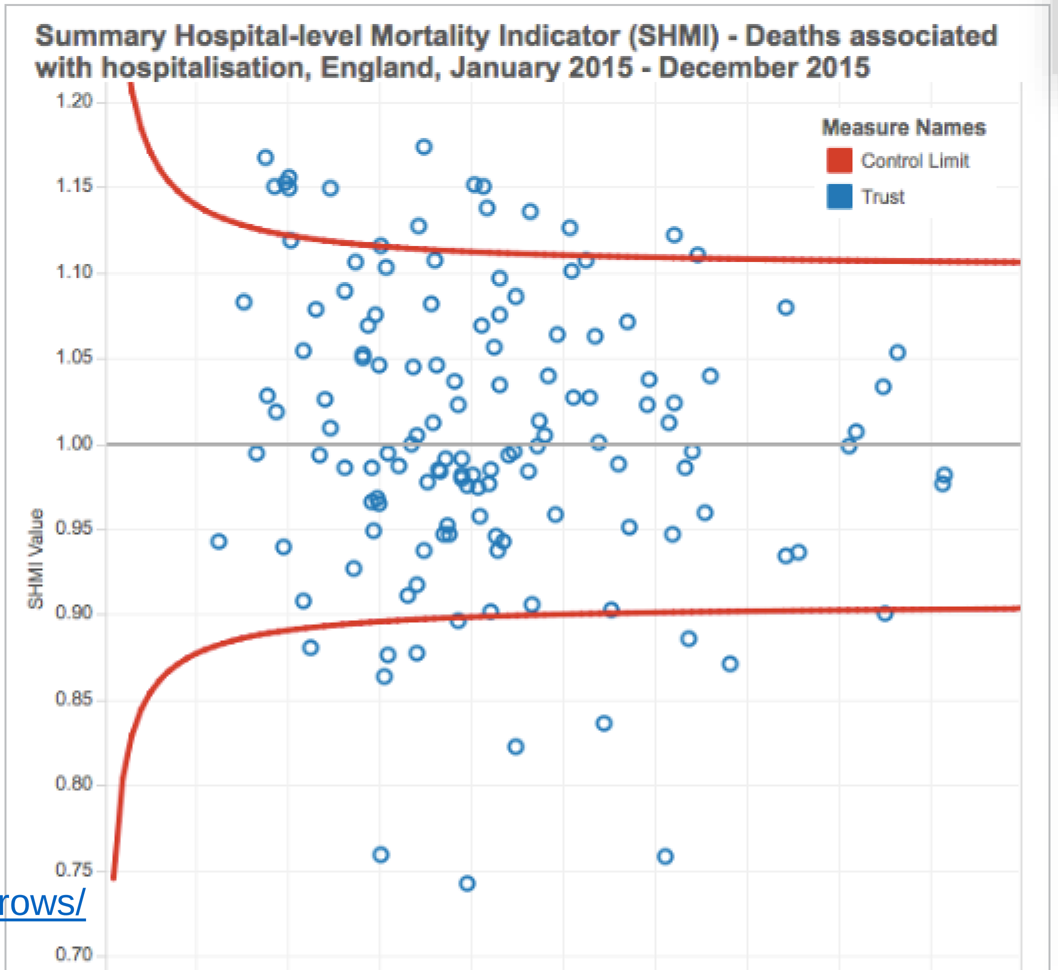
Alan wanted to identify where housing benefit claimants live, and w
pattern is changing, in order to improve local authority budget fore
used de-identified DWP monthly summaries of housing benefit clai
location (Lower Layer Super Output Areas), and clustered 17 million
points about claimant characteristics to produce this map:

As the map shows, London is a patchwork - although there is some
emerging on areas that are consistently popular with working famil
went on to apply a pattern recognition technique (k Nearest Neigh
investigate how the clusters have changed and to explore any corre
their distribution.

https://dataingovernment.blog.gov.uk/2016/08/12/the-accelerator-grows/

Sarah Culkin (Department of Health)

How do we know if too many patients died in a hospital? Explainin
concept of 'expected patient deaths' statistics to non-statisticiar
challenge that Department of Health analysts regularly encounte
clinicians and senior managers understand this concept and mak
of these statistics, Sarah developed an interactive tool to visualis
Summary Hospital-Level Mortality Indicator (SHMI) formula work
Python's Flask library to produce the app, which is deployed on H
SVG to produce the tool's animations. The tool is currently being
before being rolled out to NHS managers.



Richard Boland (Department for Education)

Richard built a web app using Python's Django library to help his colleagues at the Education Funding Agency prioritise resources. Using machine learning, the app predicts financial and governance risks associated with an Academy Trust, so that resource can be prioritised towards the cases where it is most useful. It applies Latent Dirichlet Allocation to the action plan text from Academy Trusts' Financial Management and Governance Self-Assessments (FMGS). In order to bring the data to life and put it in context with the history of the submitting Academy Trust, Richard also developed a D3 force-directed graph to represent the journey of individual academies from Trust to Trust.

The final product is not publicly available, but since the end of the project Richard has put his new skills to use to build another web app to inform his work on sampling strategies.

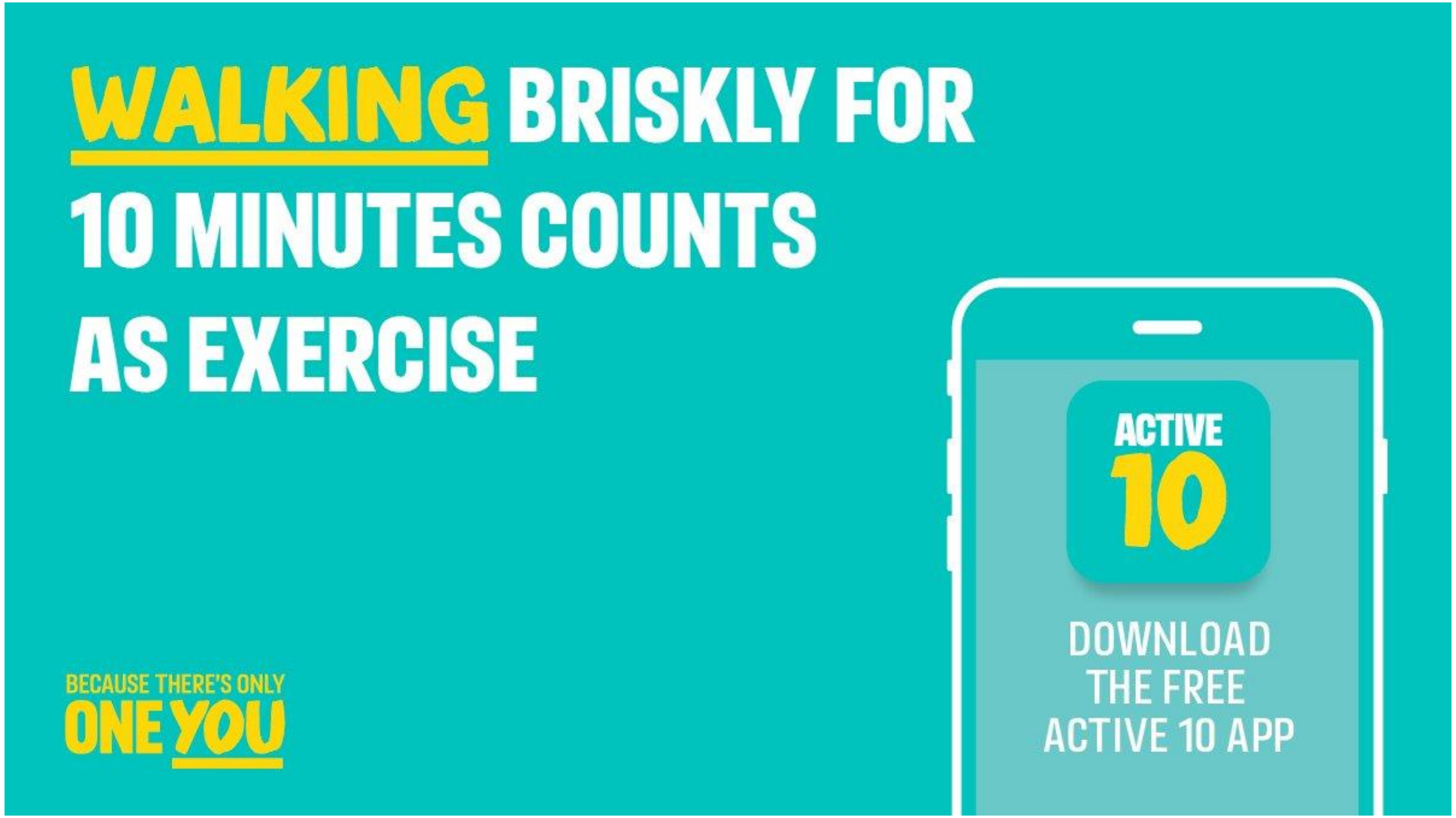
Adam Bray (Education Funding Agency)

Adam worked on forecasting construction inflation to predict the cost of building works at schools and therefore allocate budget more efficiently.

$$\left(1-\sum_{i=1}^{p}\phi_i L^i\right)(1-L)^d X_t = \delta + \left(1+\sum_{i=1}^{q}\theta_i L^i\right)\varepsilon_t$$

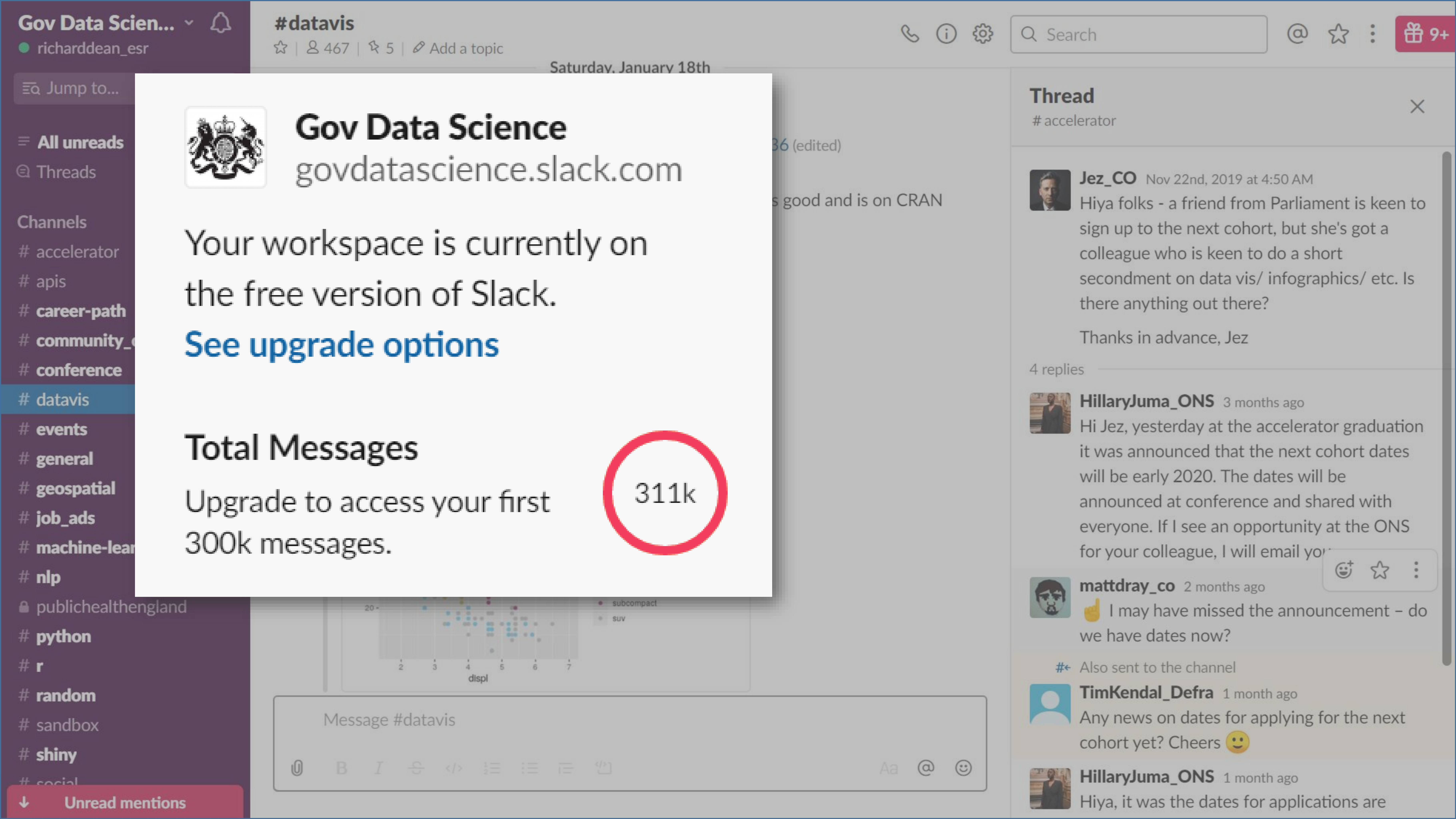
(Generalised ARIMA formula - L represents lag)

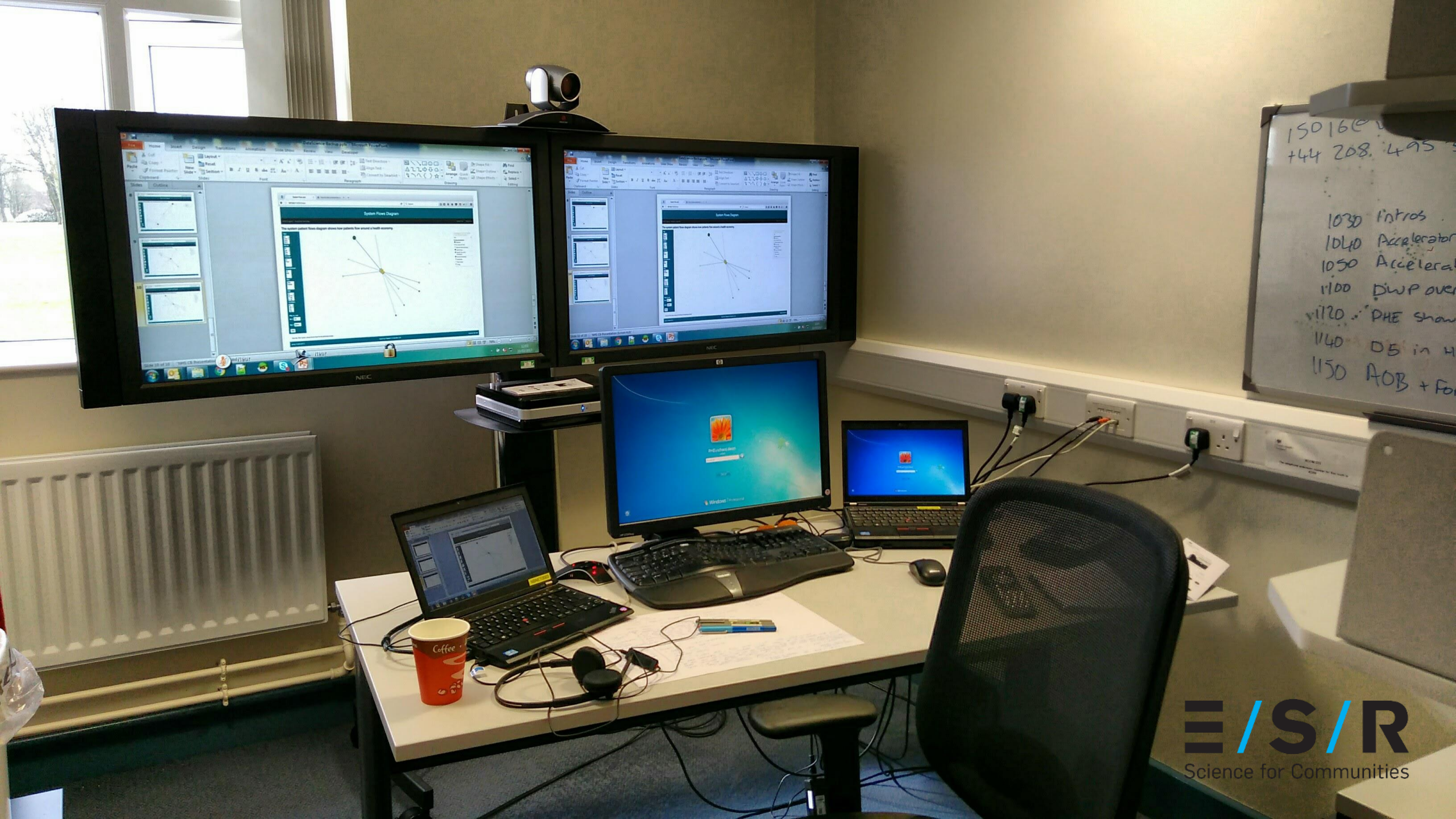
Adam used Python to implement both a Grey Model and an Autoregressive Integrated Moving Average Model (ARIMA), and found that it was possible to forecast inflation to within an error of around 3% for short time horizons. This model will be used in construction market intelligence for school building.



One in every six UK deaths can be attributed to inactivity, which ultimately costs the NHS over £900 million each year.

E/S/R
Science for Communities


Gov Data Scien...
richarddean_esr
Jump to...
All unreads
Threads
Channels
accelerator
apis
career-path
community_
conference
datavis
events
general
geospatial
job_ads
machine-lear
nlp
publichealthengland
python
r
random
sandbox
shiny
Unread mentions
#datavis
467 | 5 | Add a topic
Saturday, January 18th
Search
Gov Data Science
govdatascience.slack.com
Your workspace is currently on the free version of Slack.
See upgrade options
Total Messages
Upgrade to access your first 300k messages.
311k
Message #datavis
Thread
accelerator
Jez_CO Nov 22nd, 2019 at 4:50 AM
Hiya folks - a friend from Parliament is keen to sign up to the next cohort, but she's got a colleague who is keen to do a short secondment on data vis/ infographics/ etc. Is there anything out there?
Thanks in advance, Jez
4 replies
HillaryJuma_ONS 3 months ago
Hi Jez, yesterday at the accelerator graduation it was announced that the next cohort dates will be early 2020. The dates will be announced at conference and shared with everyone. If I see an opportunity at the ONS for your colleague, I will email yo
mattdray_co 2 months ago
I may have missed the announcement – do we have dates now?
Also sent to the channel
TimKendal_Defra 1 month ago
Any news on dates for applying for the next cohort yet? Cheers
HillaryJuma_ONS 1 month ago
Hiya, it was the dates for applications are



E/S/R
Science for Communities

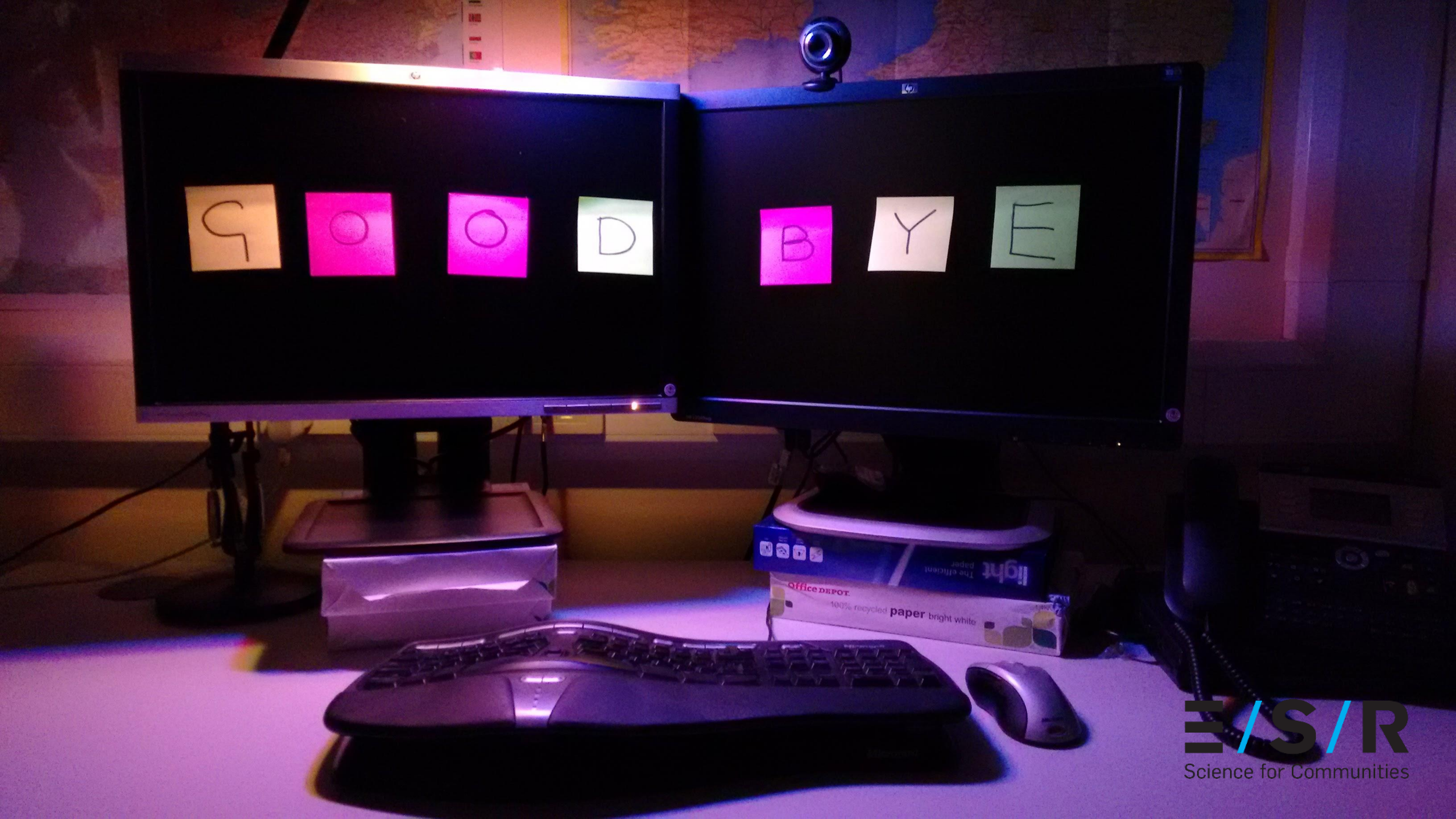
:(

Your PC ran into a problem and needs to restart. We're just collecting some error info, and then we'll restart for you.

0% complete

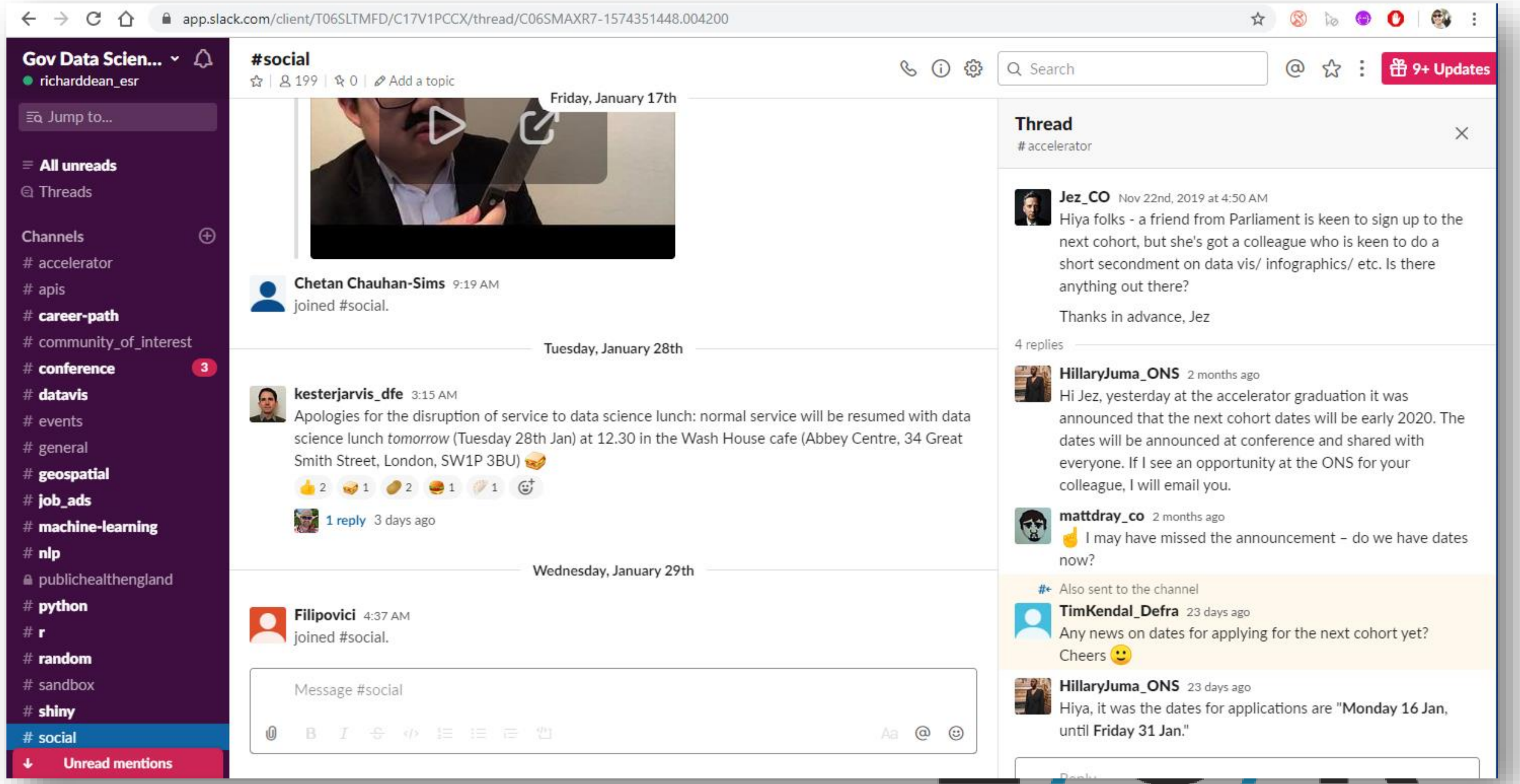
For more information about this issue and possible fixes, visit https://www.windows.com/stopcode

If you call a support person, give them this info:

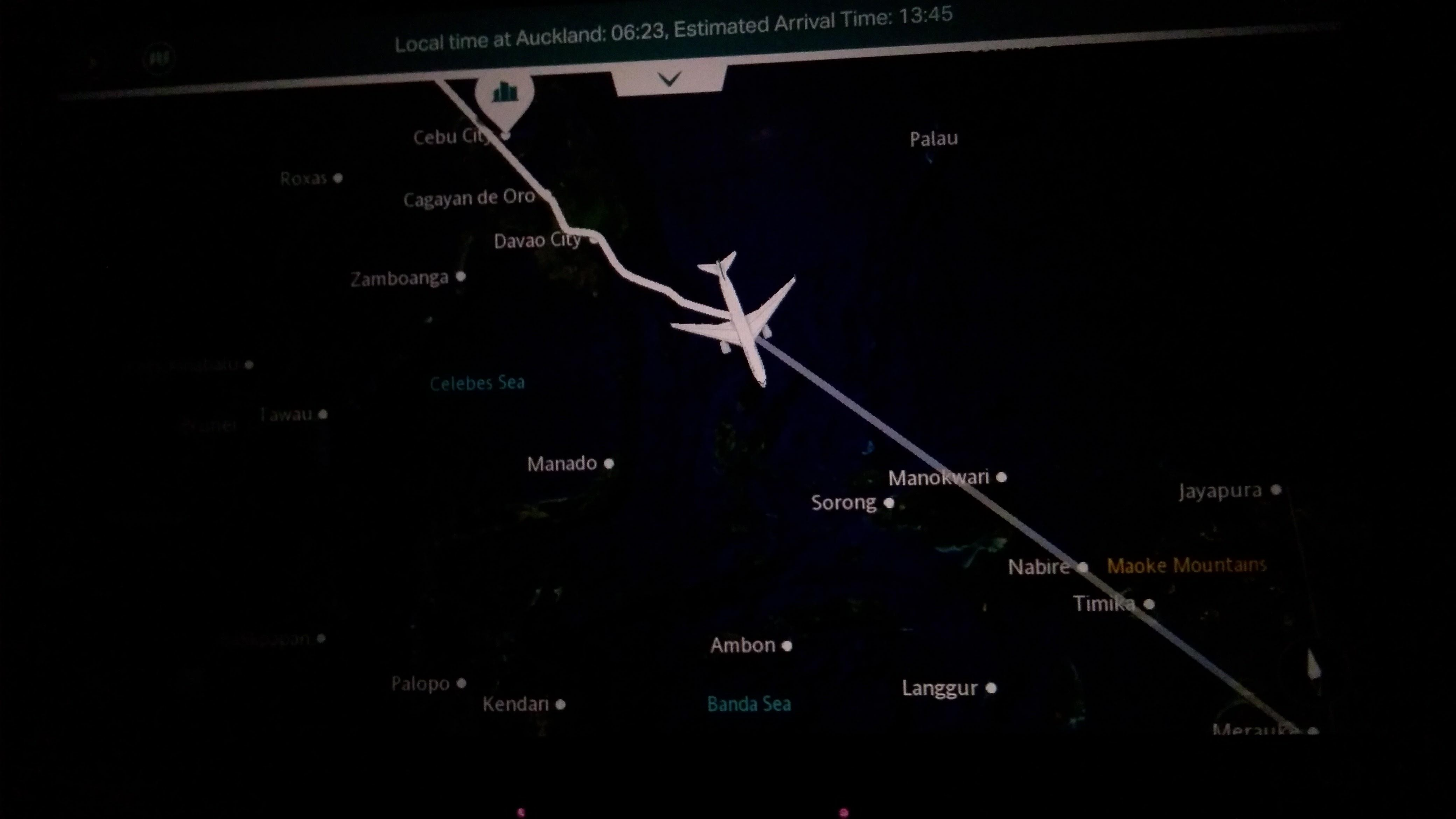

G O O D B Y E
Office DEPOT
100% recycled paper bright white
E/S/R
Science for Communities

Moving to NZ

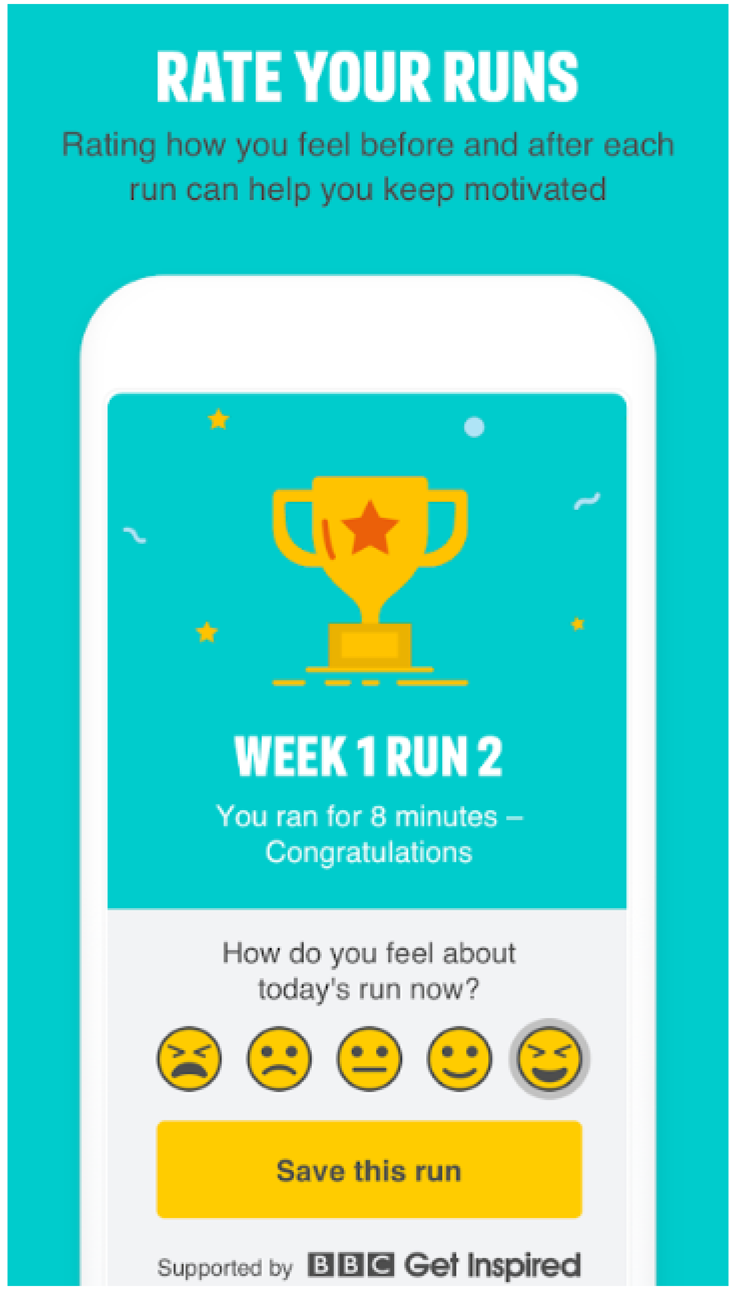
- Enjoying PHE, active 10, variety in UK, things going well
- Data science community established, data beers / biscuits, / day / conferences / accelerators.
- Slack
- What would NZ provide – bit of a leap?
- Saw from UK that the NZ data science meetup community was very active



Science for Communities



Local time at Auckland: 06:23, Estimated Arrival Time: 13:45
Cebu City
Palau
Roxas
Cagayan de Oro
Davao City
Zamboanga
Celebes Sea
Tawau
Manado
Manokwari
Sorong
Jayapura
Nabire
Maoke Mountains
Timika
Ambon
Langgur
Palopo
Kendari
Banda Sea

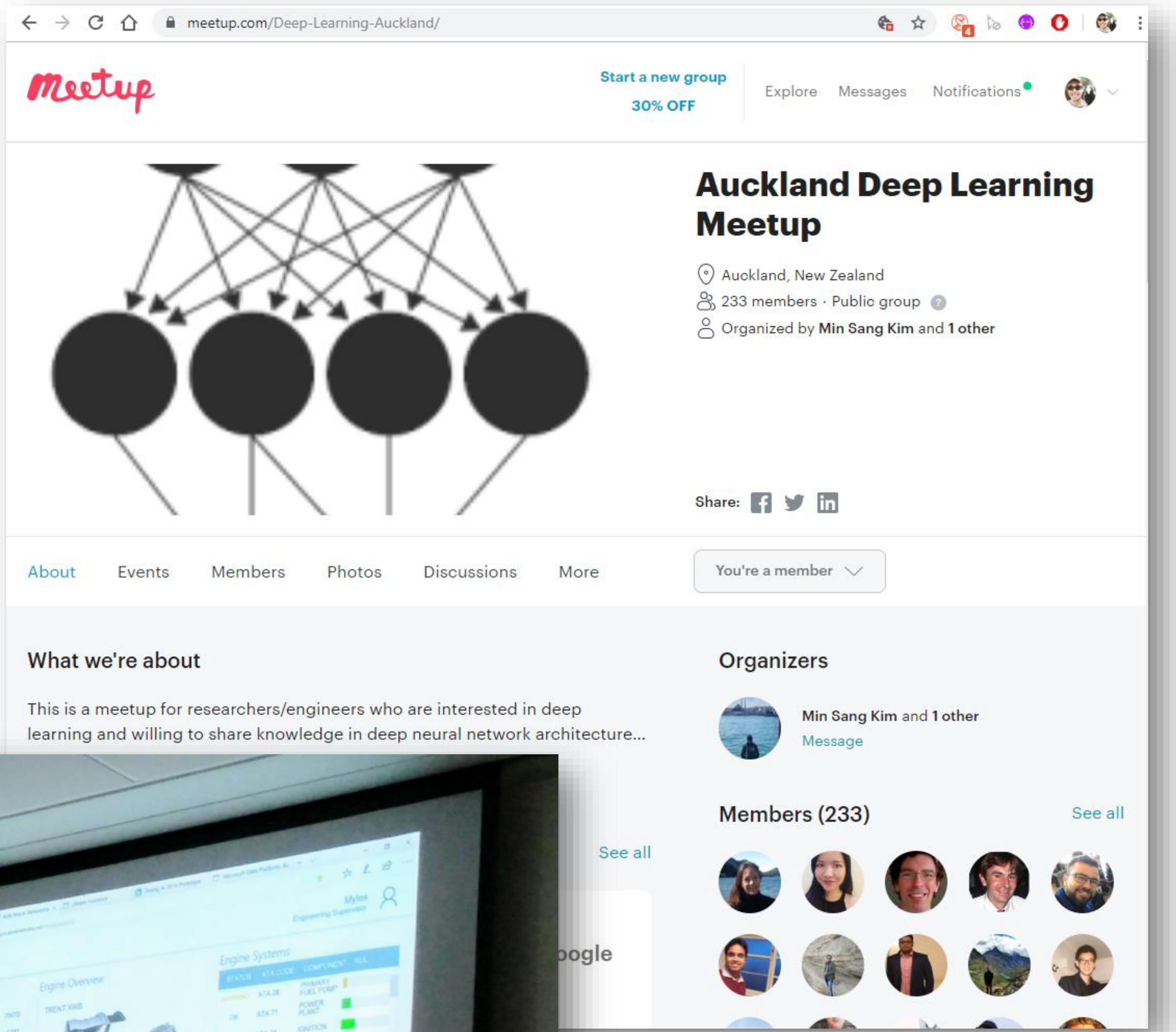


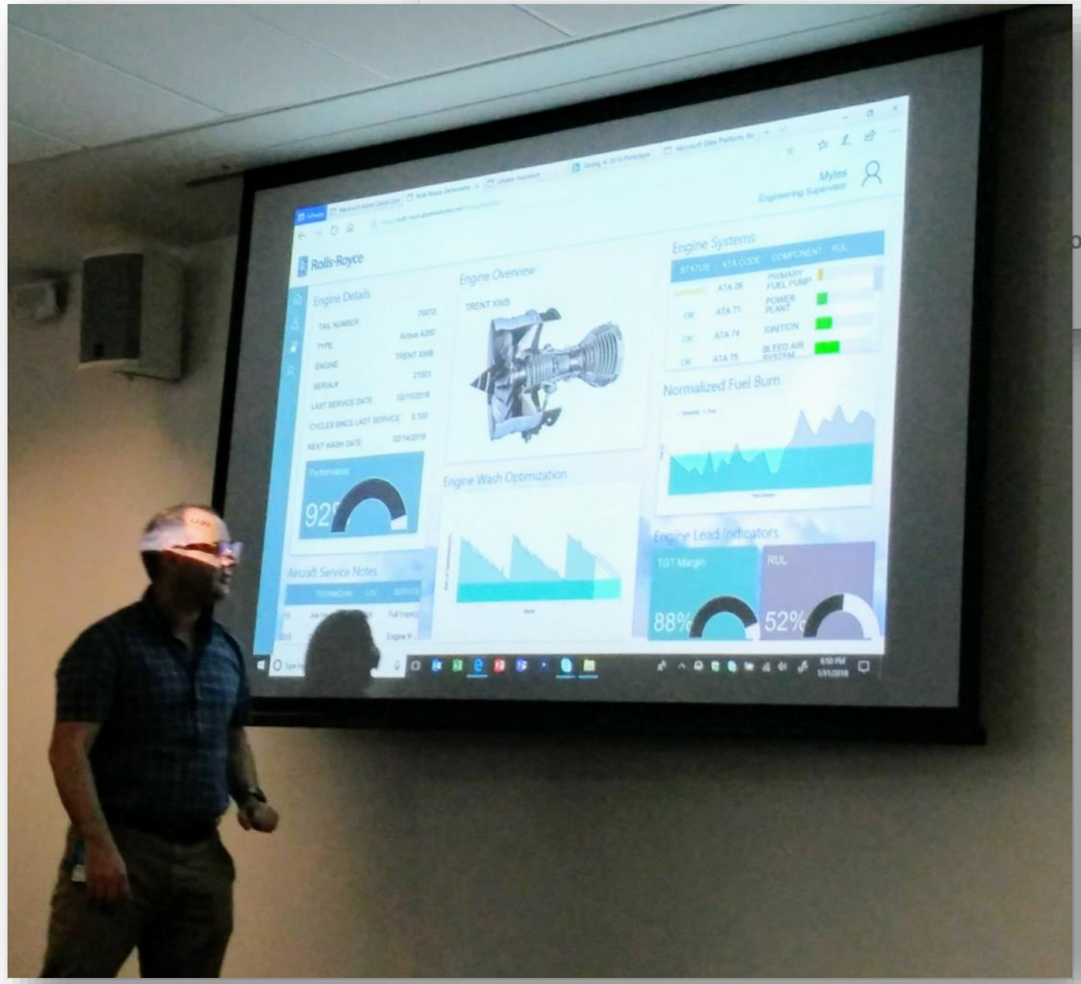
E/S/R
Science for Communities


1 min pitch
Name of Team
Idea
How many people you have now
What needing
- Role
- Skills
Declarations

Data Science Communities

- Meetups
- Deep learning
- Machine learning + AI (in various guises)
- Tech forum
- Gov tech
- Uni seminars
- Gov analytics network
- Haven't quite found critical mass



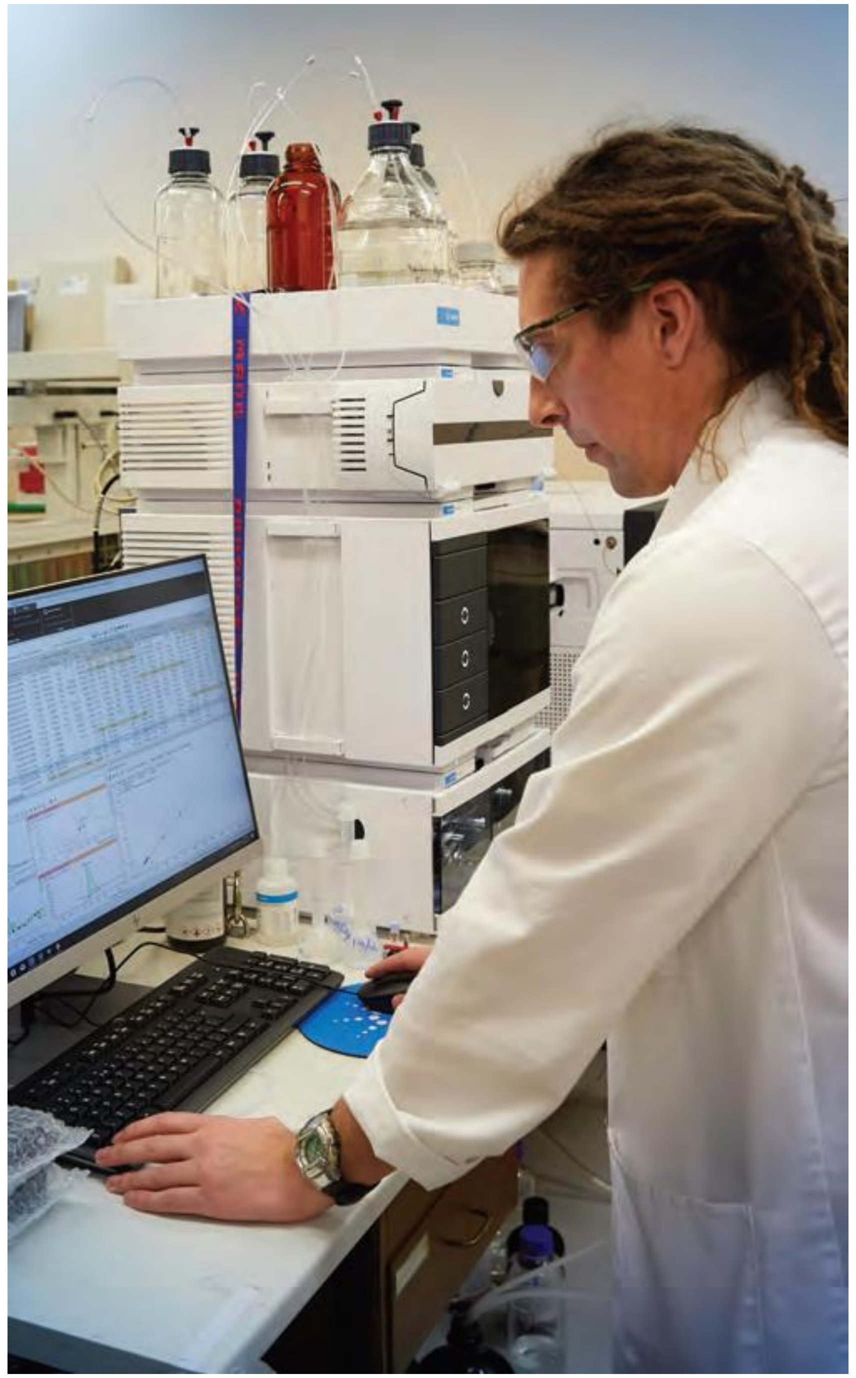


E/S/R
Science for Communities

Remote working

Where to do data science in NZ?

What is ESR?

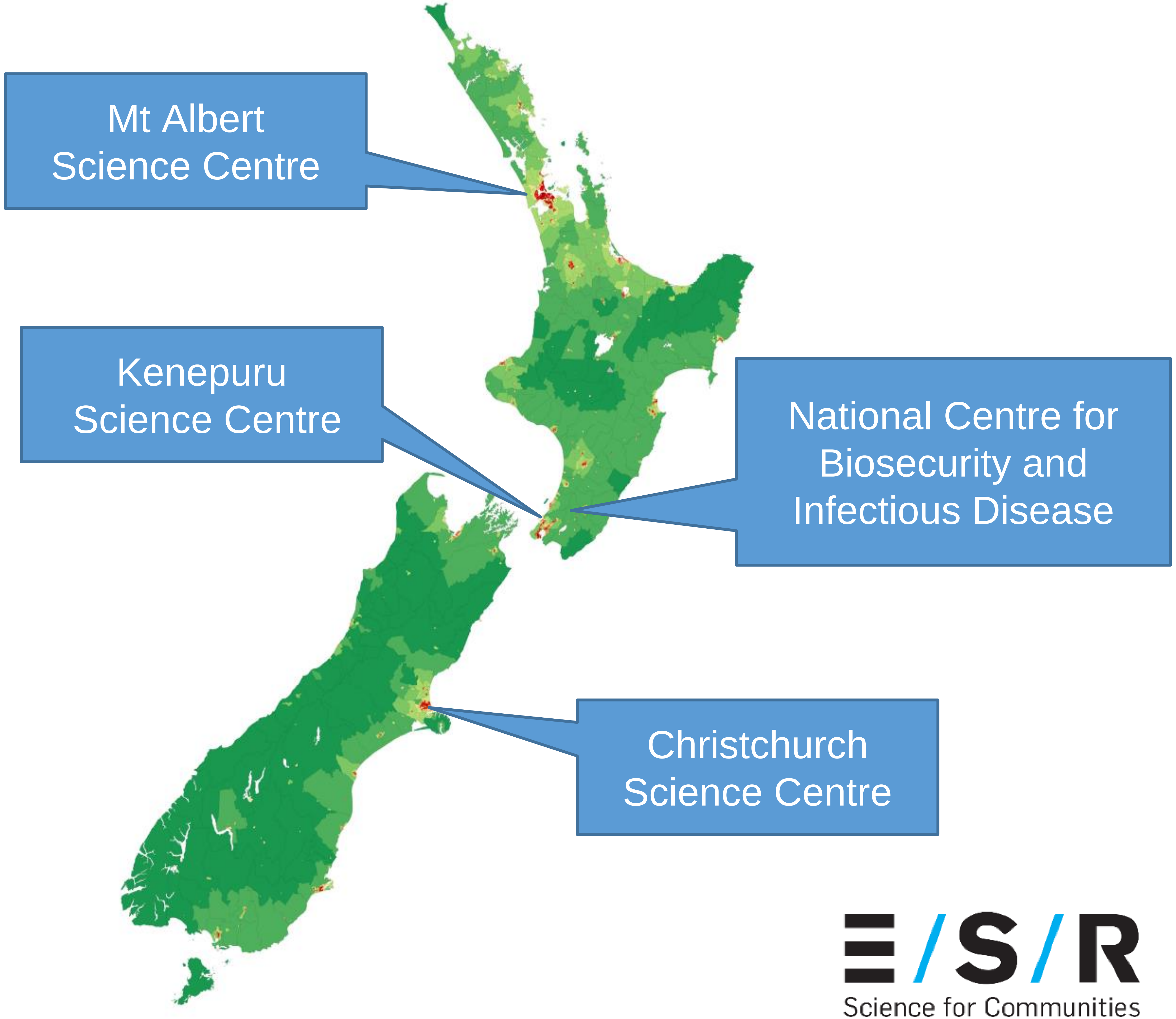


Science for Communities

Science for Communities

Who are we?

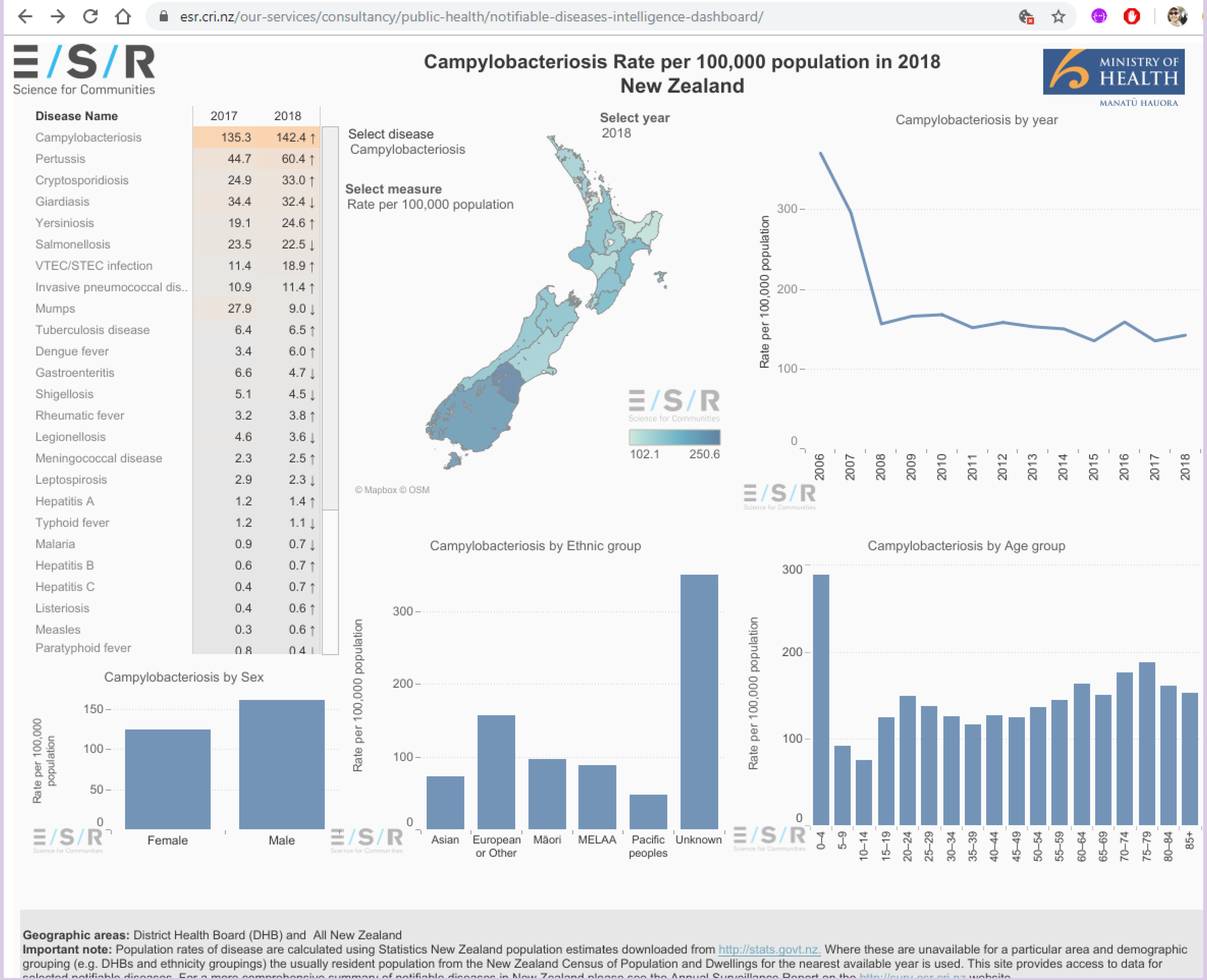
- Crown Research Institute
- Approx. 420 staff
- 4 locations across NZ



E/S/R
Science for Communities

PUBLIC HEALTH

Safeguard the health of New Zealanders through improvements in the management of biosecurity and threats to public health



Disease Name	2017	2018
Campylobacteriosis	135.3	142.4 ↑
Pertussis	44.7	60.4 ↑
Cryptosporidiosis	24.9	33.0 ↑
Giardiasis	34.4	32.4 ↓
Yersiniosis	19.1	24.6 ↑
Salmonellosis	23.5	22.5 ↓
VTEC/STEC infection	11.4	18.9 ↑
Invasive pneumococcal dis..	10.9	11.4 ↑
Mumps	27.9	9.0 ↓
Tuberculosis disease	6.4	6.5 ↑
Dengue fever	3.4	6.0 ↑
Gastroenteritis	6.6	4.7 ↓
Shigellosis	5.1	4.5 ↓
Rheumatic fever	3.2	3.8 ↑
Legionellosis	4.6	3.6 ↓
Meningococcal disease	2.3	2.5 ↑
Leptospirosis	2.9	2.3 ↓
Hepatitis A	1.2	1.4 ↑
Typhoid fever	1.2	1.1 ↓
Malaria	0.9	0.7 ↓
Hepatitis B	0.6	0.7 ↑
Hepatitis C	0.4	0.7 ↑
Listeriosis	0.4	0.6 ↑
Measles	0.3	0.6 ↑
Paratyphoid fever	0.8	0.4 ↓

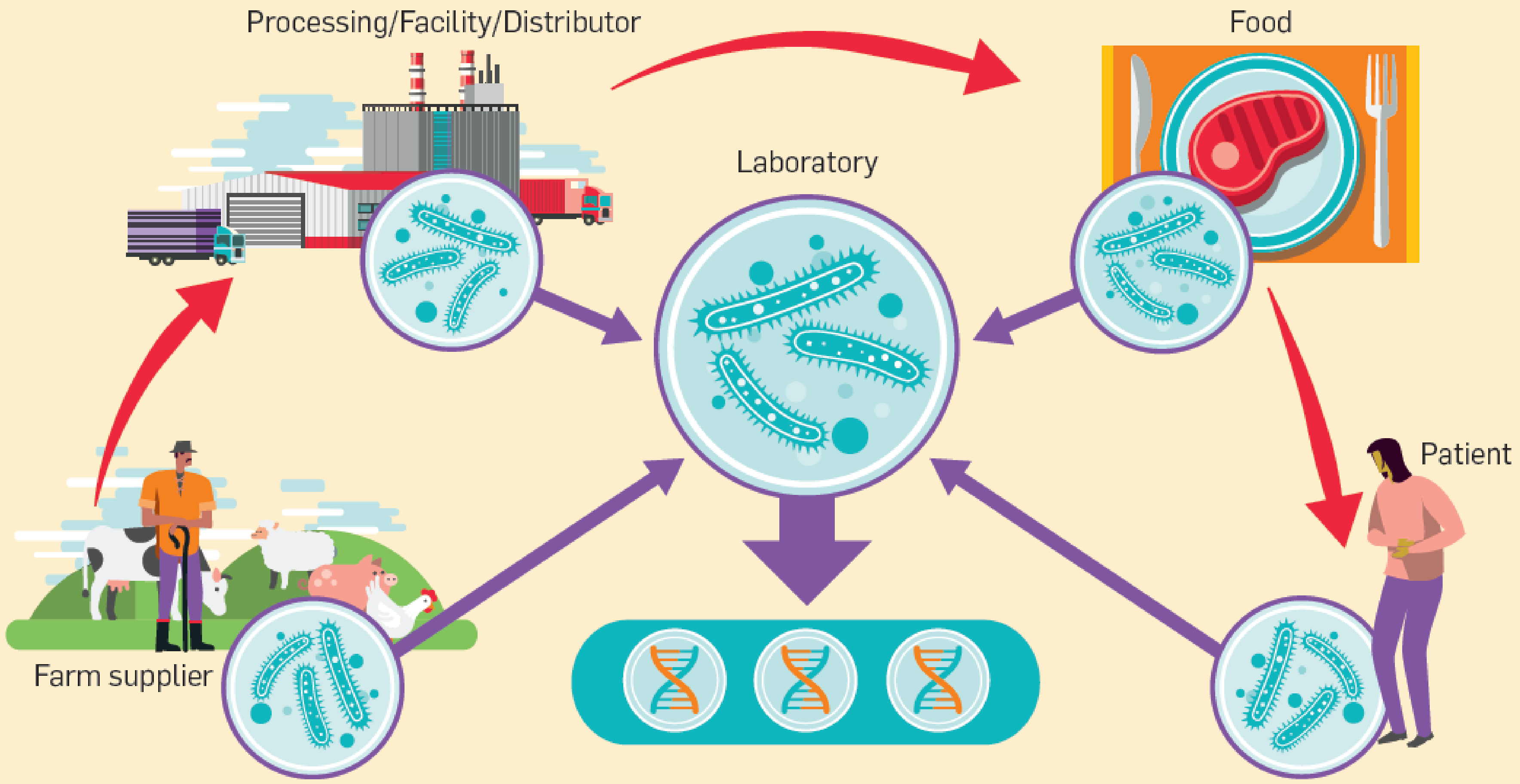
FORENSICS

Increase the effectiveness of forensic science services applied to safety, security and justice investigations and processes

E/S/R
Science for Communities

FOOD SAFETY

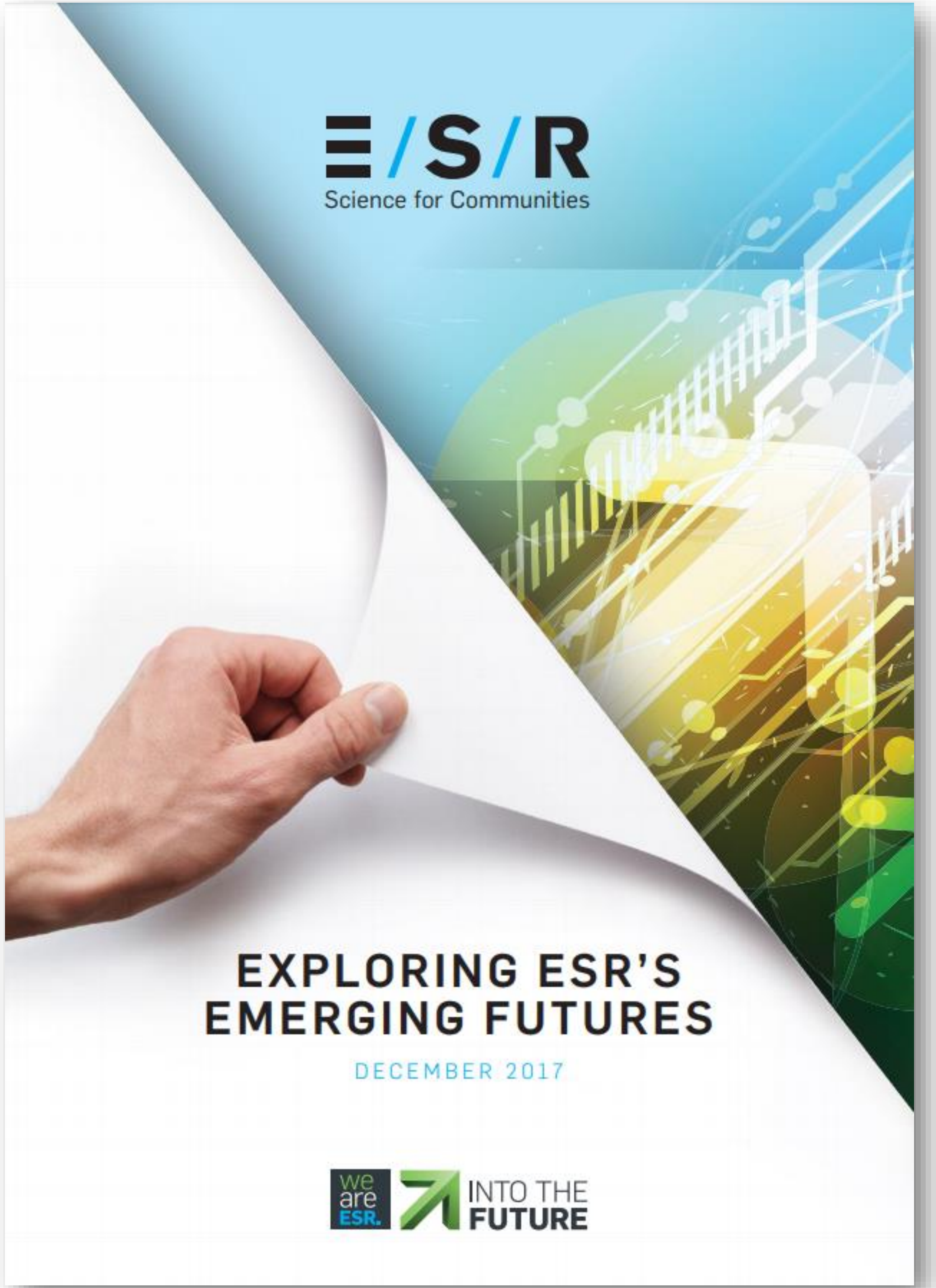
Enhance protection of New Zealand's food based economy through the management of food safety risks associated with traded goods.



Science for Communities

WATER AND THE ENVIRONMENT

Improve the safety of freshwater and groundwater resources for human use and the safer use of biowastes



EXTERNAL TRENDS

A number of potentially disruptive global trends and sub-themes were identified which will likely influence the way that ESR will operate in the future and these are highlighted below.

- → Rapid technological change or revolution
  - Automation
  - Miniaturisation
  - Increasing mobility
  - Artificial intelligence and machine learning
  - Big data and data analytics
  - Advanced digital capability incl. social media, digital crime, block chain certification, virtuality, the Internet of Things
- → Rapid contemporaneous and inter-linked development of core sciences
  - Remote monitoring
  - Data sciences
  - Rapidly emergent new applications based on emerging technologies e.g. bacterial data storage.
- → Data sharing
  - Transparency
  - Open datasets
- → Changing economic models
  - Subscription models such as My Food Bag and Netflix
  - The Giga-economy
  - Competition through delivery to emergent needs
  - Social investment models
- → Social trends and changing norms
  - Diversity
  - Changing expectations and changing ideas of 'community'
  - The growth of collaborative approaches
  - Terrorism
- → New methods of teaching and learning

Through analysis and grouping of these external trends and discussion of the implications each has on the future operations of ESR, several megatrends emerged:

- → **Mutualism and alliance:** we need to work even more closely with our clients and consumers
- → **Adaptable and responsive – cutting to the chase:** we must be highly focussed on the most highly prioritised requirements of our clients and identify and respond rapidly to emergent needs
- → **Power to the people – and other clients:** we need to recognise and address the increasing engagement of consumers in every aspect of our work and the opportunity to work collaboratively with consumers more broadly to best serve our clients
- → **Capability revolution – science and beyond:** whilst the trends that will impact us drive rapid scientific capability change, our non-science capabilities must also evolve quickly
- → **Culture – highly empowered, highly accountable:** our staff will be working more directly with clients and consumers and must therefore be both empowered and accountable in terms of alignment with ESR's strategic direction.

https://www.esr.cri.nz/home/about-esr/corporate-publications/exploring-esrs-emerging-futures/

E/S/R
Science for Communities

Data science groundwork

Future State

ESR will have

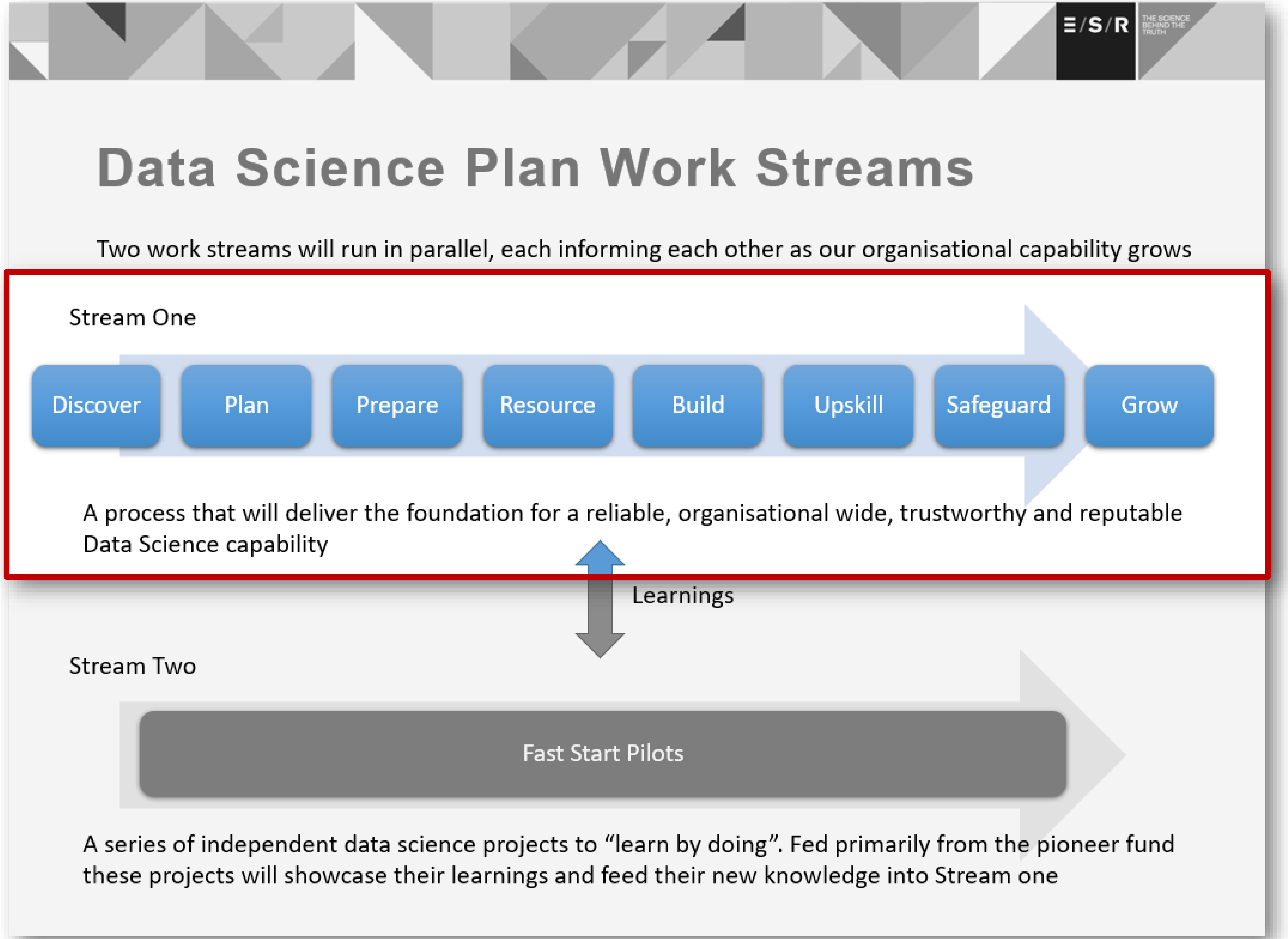
- Ready access to a Data Platfo
  both cloud and on premise ✓
- A customer focus on our client
- Strategically led, Coordinated,
- A Data Science group structure
  and accessible to all ESR scie
- Privileged access to unique da
  access to public datasets
- Ready access to ESR Data La
- A well understood and practice
  consistency and conclusions t
- A baseline of core Data Scienc
- Data Science that is a key cont
- A Data Science Capability that
- Our resources have a more ou
  Science community to better u

Future State pathway

We will

- Accelerate our understanding of what is possible by discovering and observing what leading edge organisations are doing and benchmarking ourselves against the best. ✓
- Drive actions from our Data Science Plan ✓
- Rapidly grow and expand our capability through upskilling and recruitment ✓
- Establish a new centralised Data Science Team that is led by an expert and equipped with the skills and technology that are available to all ESR ✓
- Establish and nurture customer relationships to gain access to rare datasets and to understand their requirements
- Take our partners and stakeholders with us on our data science journey
- Undertake a Data Discovery Audit of our datasets and design data management practices that protect our citizens, communities, client and our organisation
- Establish leading Governance, privacy and security practices for working with data

✓ = Achieved


E/S/R
THE SCIENCE BEHIND THE TRUTH
Data Science Plan Work Streams
Two work streams will run in parallel, each informing each other as our organisational capability grows
Stream One
Discover
Plan
Prepare
Resource
Build
Upskill
Safeguard
Grow
A process that will deliver the foundation for a reliable, organisational wide, trustworthy and reputable Data Science capability
Learnings
Stream Two
Fast Start Pilots
A series of independent data science projects to "learn by doing". Fed primarily from the pioneer fund these projects will showcase their learnings and feed their new knowledge into Stream one

Machine learning

Visualisation

Random forests

Cloud computing

Artificial intelligence

Classification

Ethics

Communicating

Geospatial

Volume

Velocity

Variety

Wrangling

Automation

Making predictions

Statistics

Gradient boosting

Neural networks

Regression

Clustering

Governance

Explainable algorithms

Insight

Cleaning

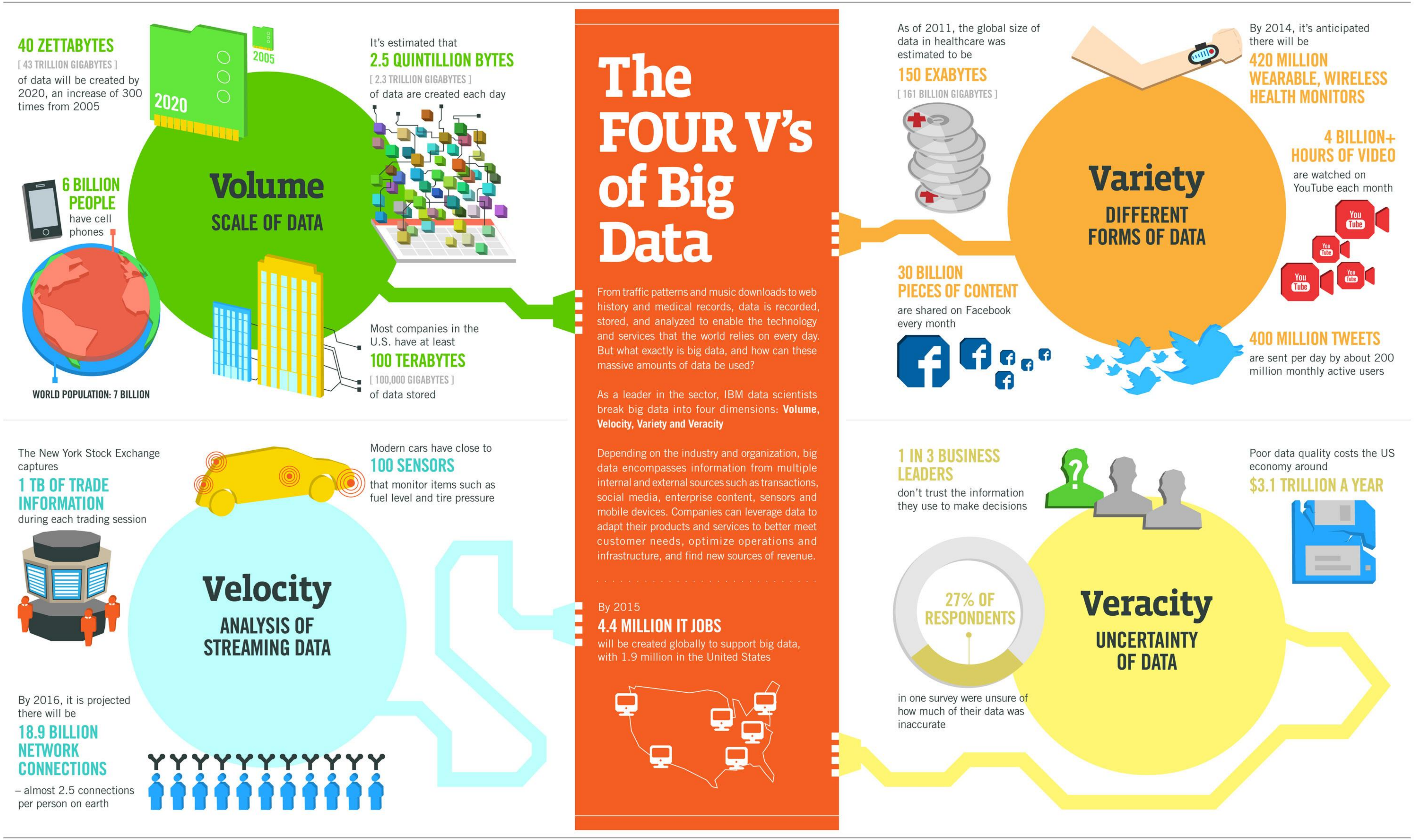
Source vector machines

Data exploration

Decision Trees

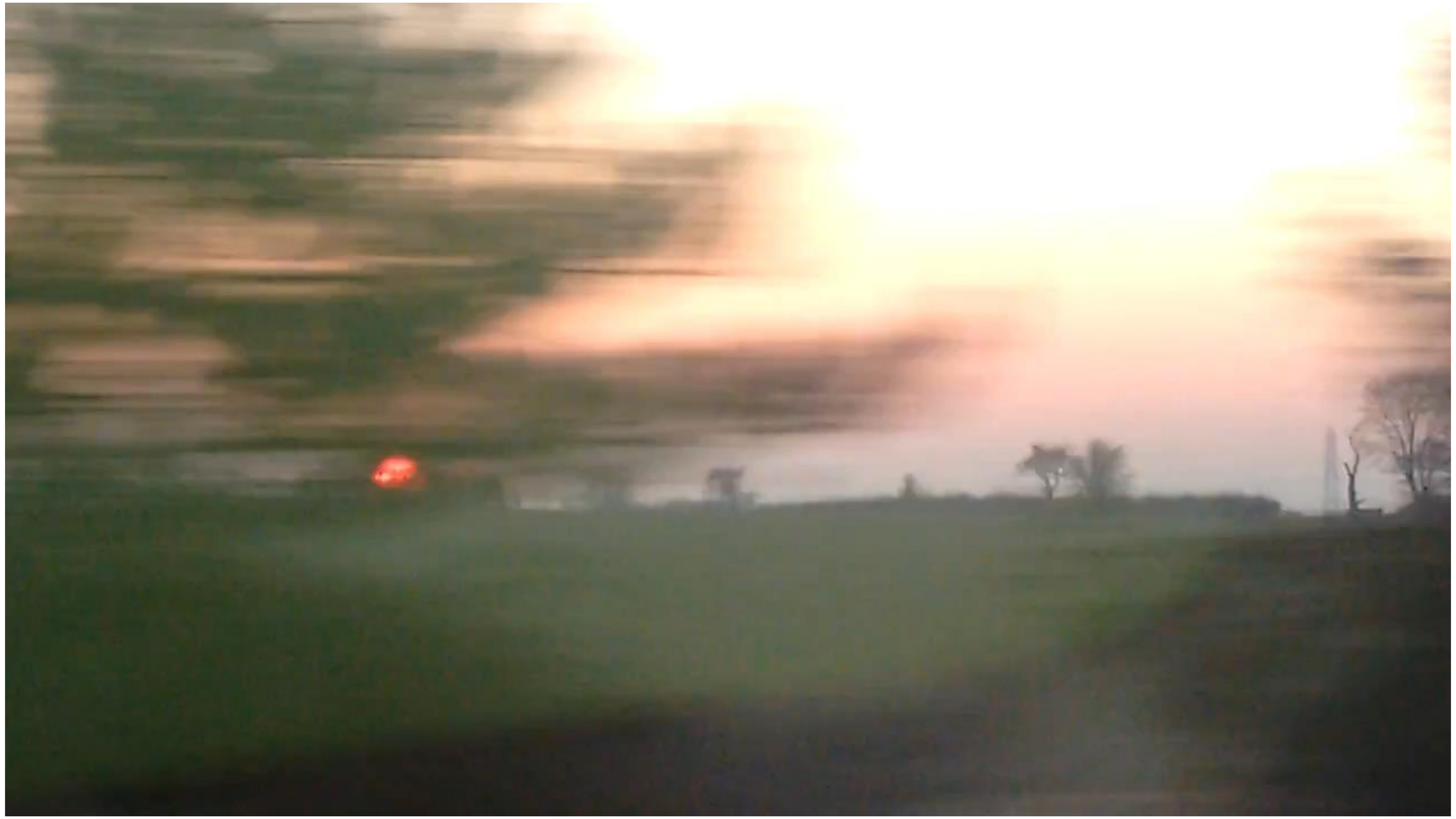
E/S/R

Science for Communities



https://www.ibmbigdatahub.com/sites/default/files/infographic_file/4-Vs-of-big-data.jpg

Fast moving landscape



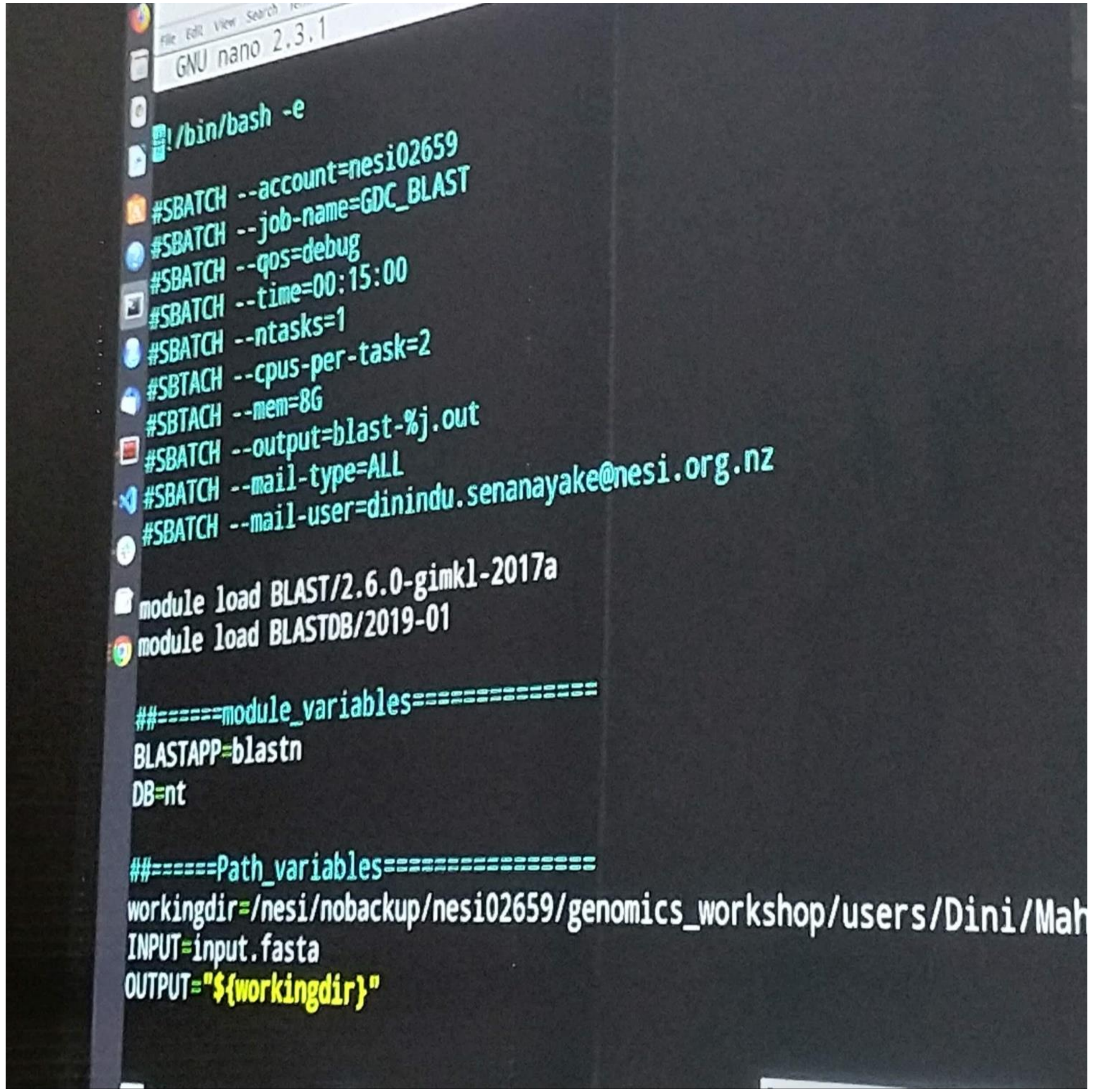
Working from the train, 2016

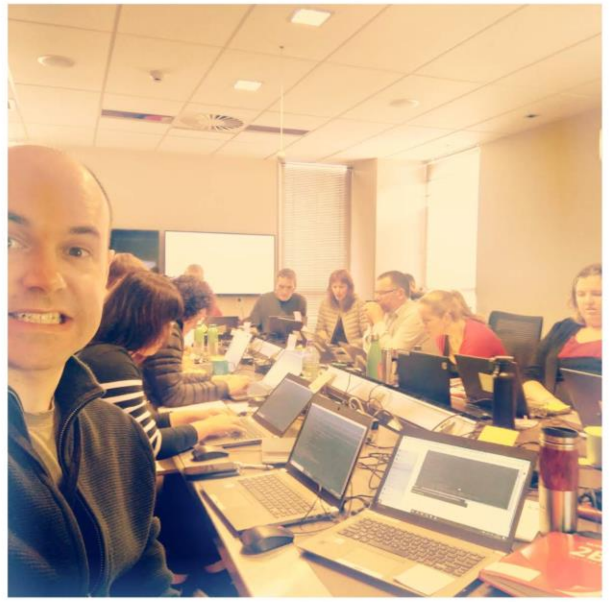
Science for Communities

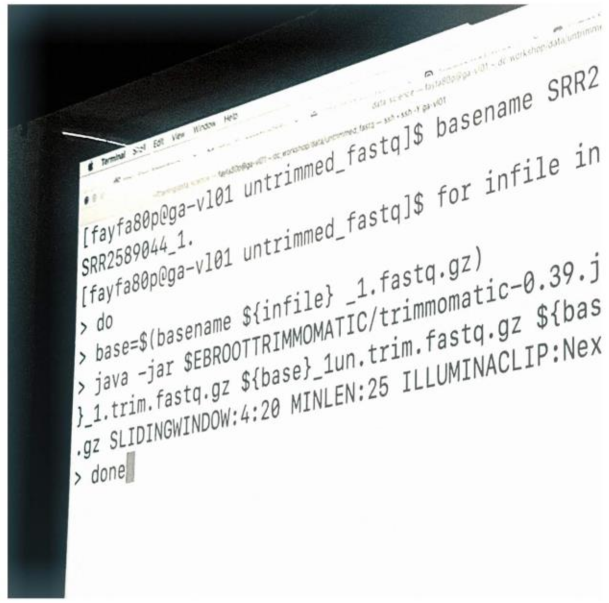
Data carpentries

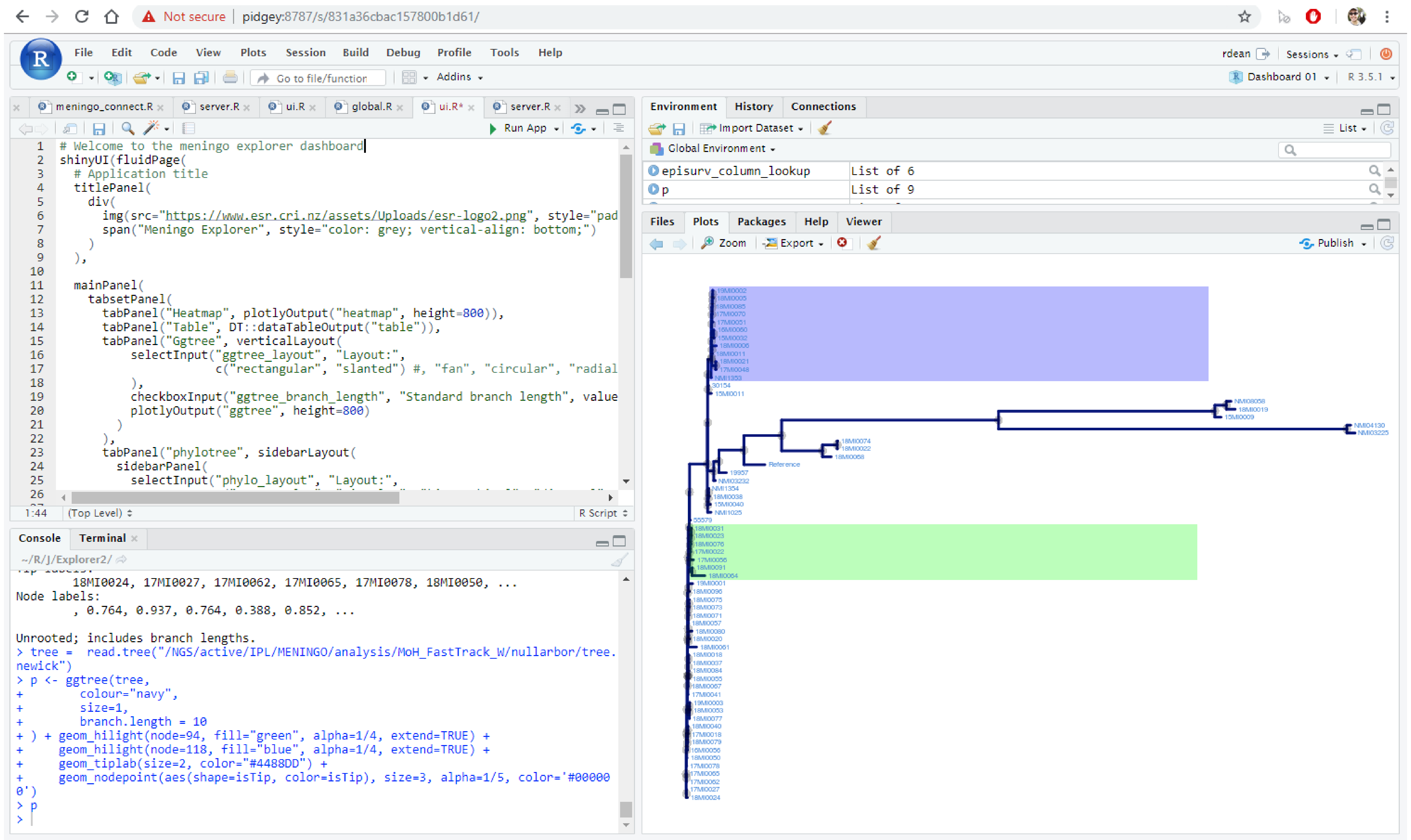
Science for Communities

Genomics carpentries

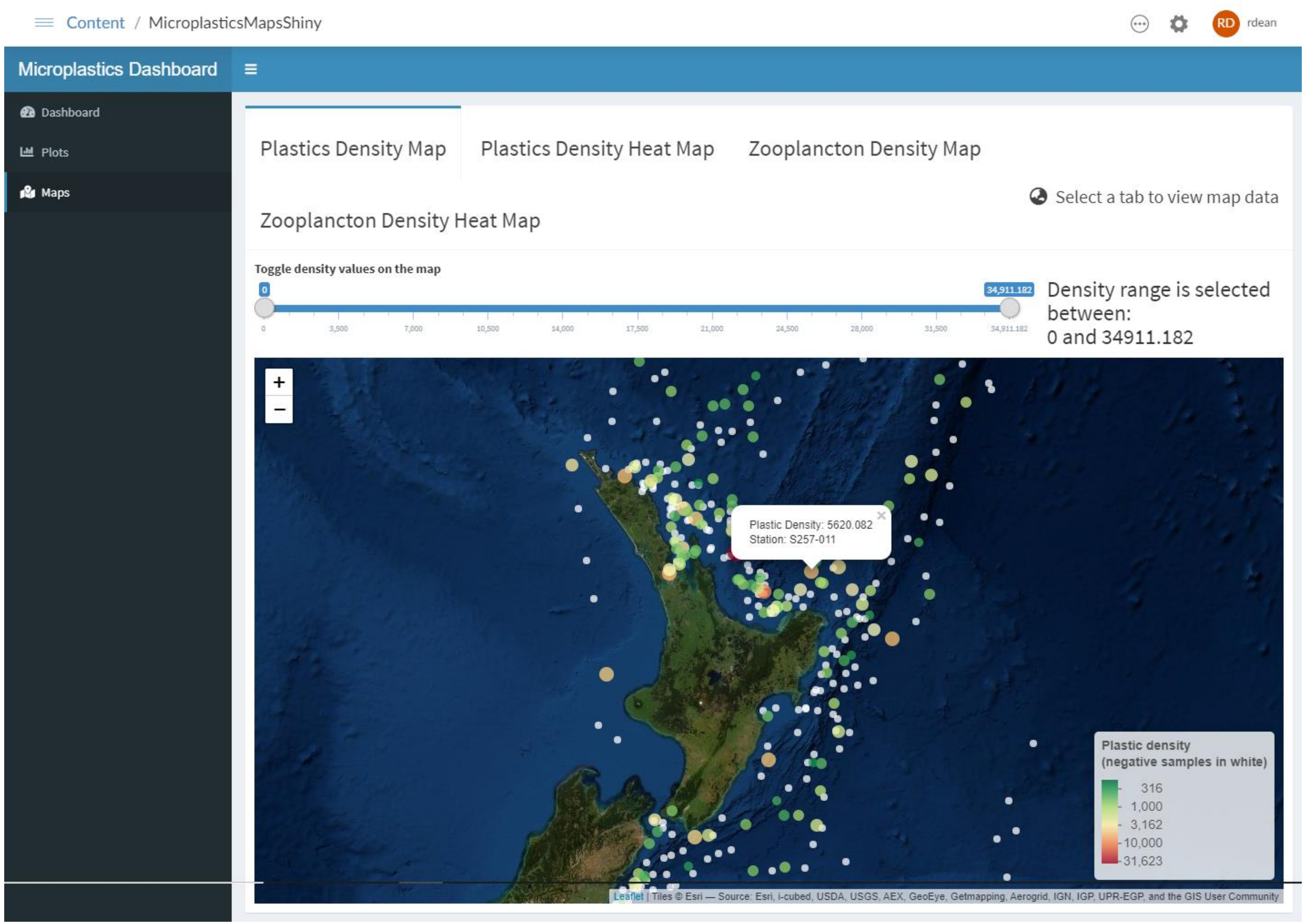






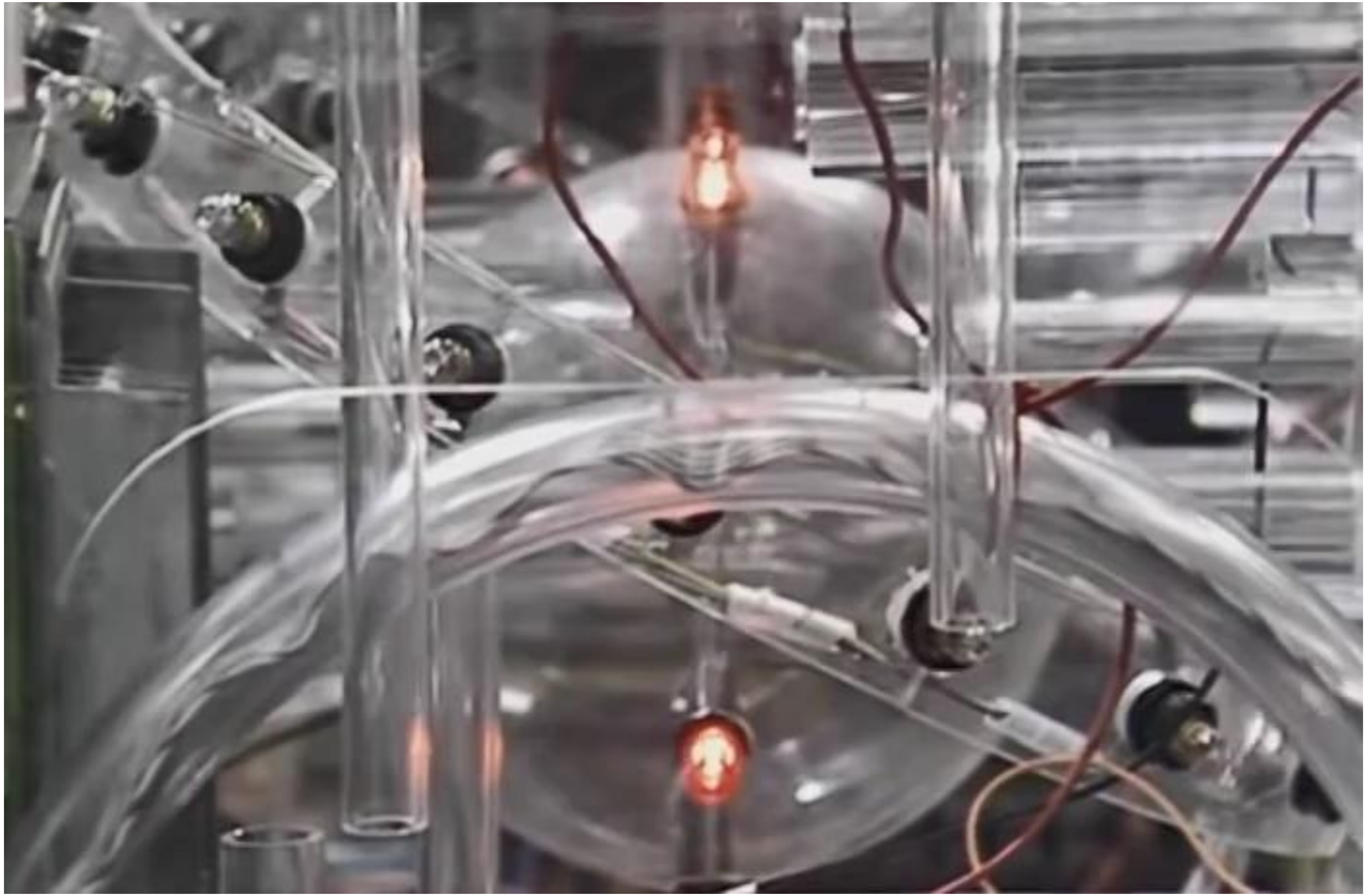


Science for Communities

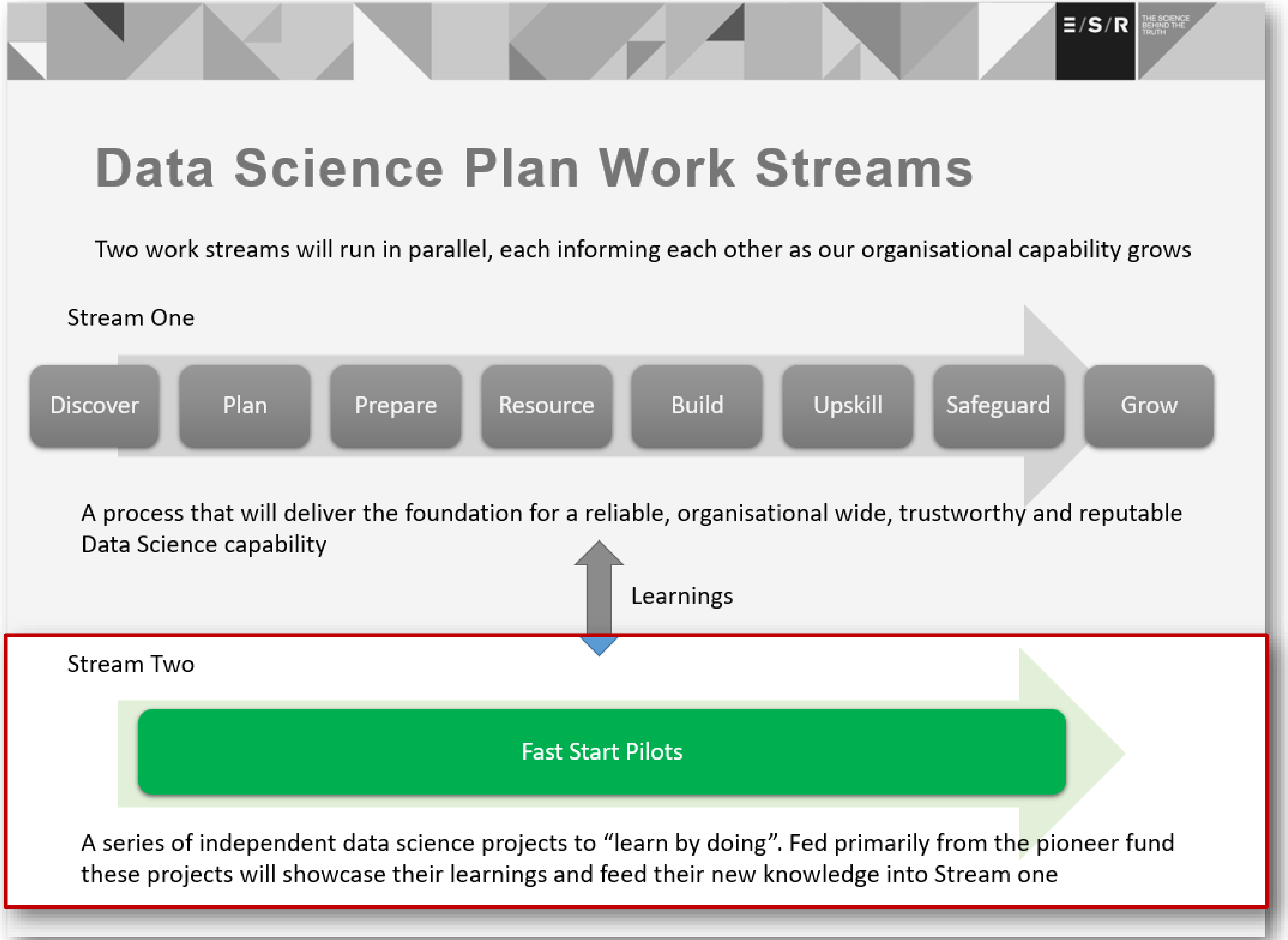


Credit - Amazon - [Pokemon Snorlax Onesie]

ESR
Science for Communities

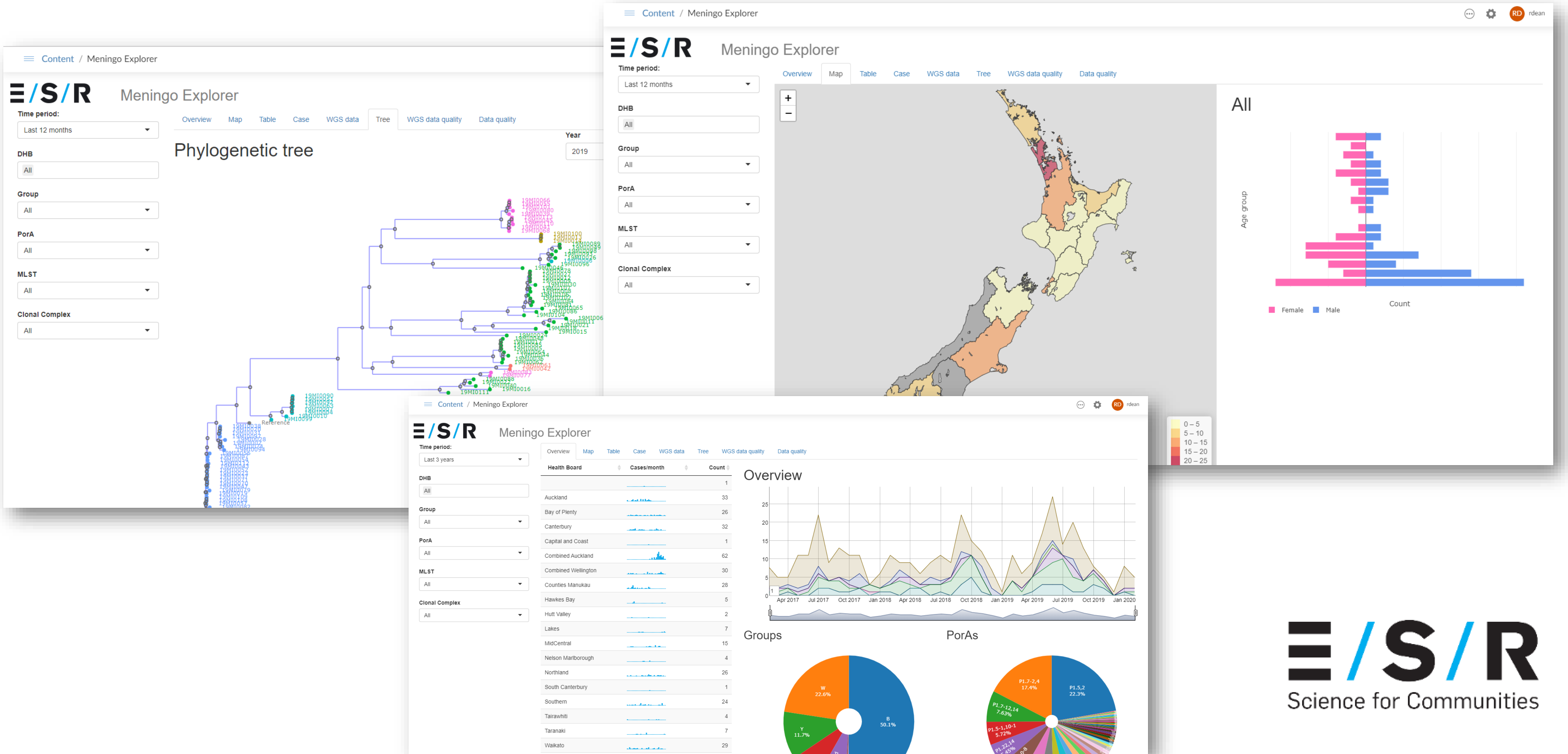


E/S/R
Science for Communities



ESR
Science for Communities

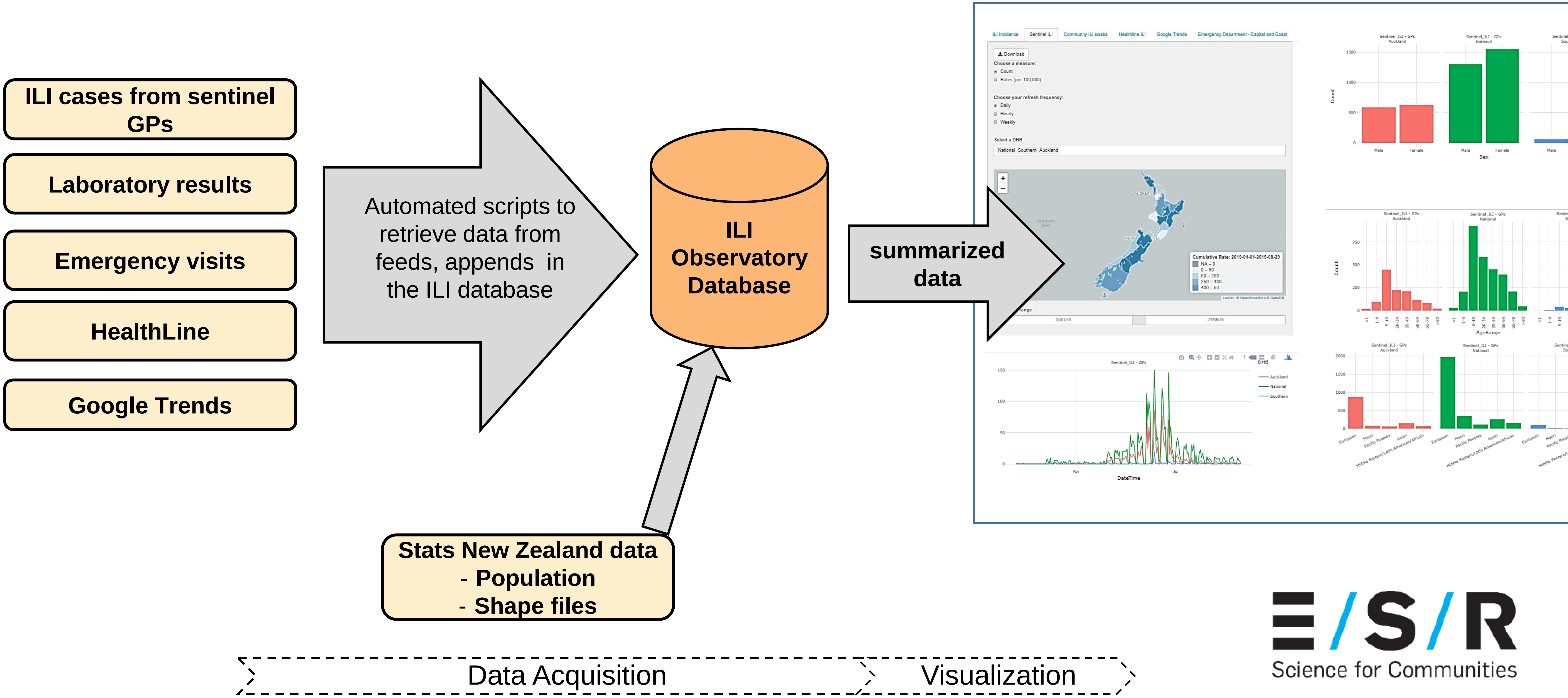
Genomics to clinical wisdom



E/S/R

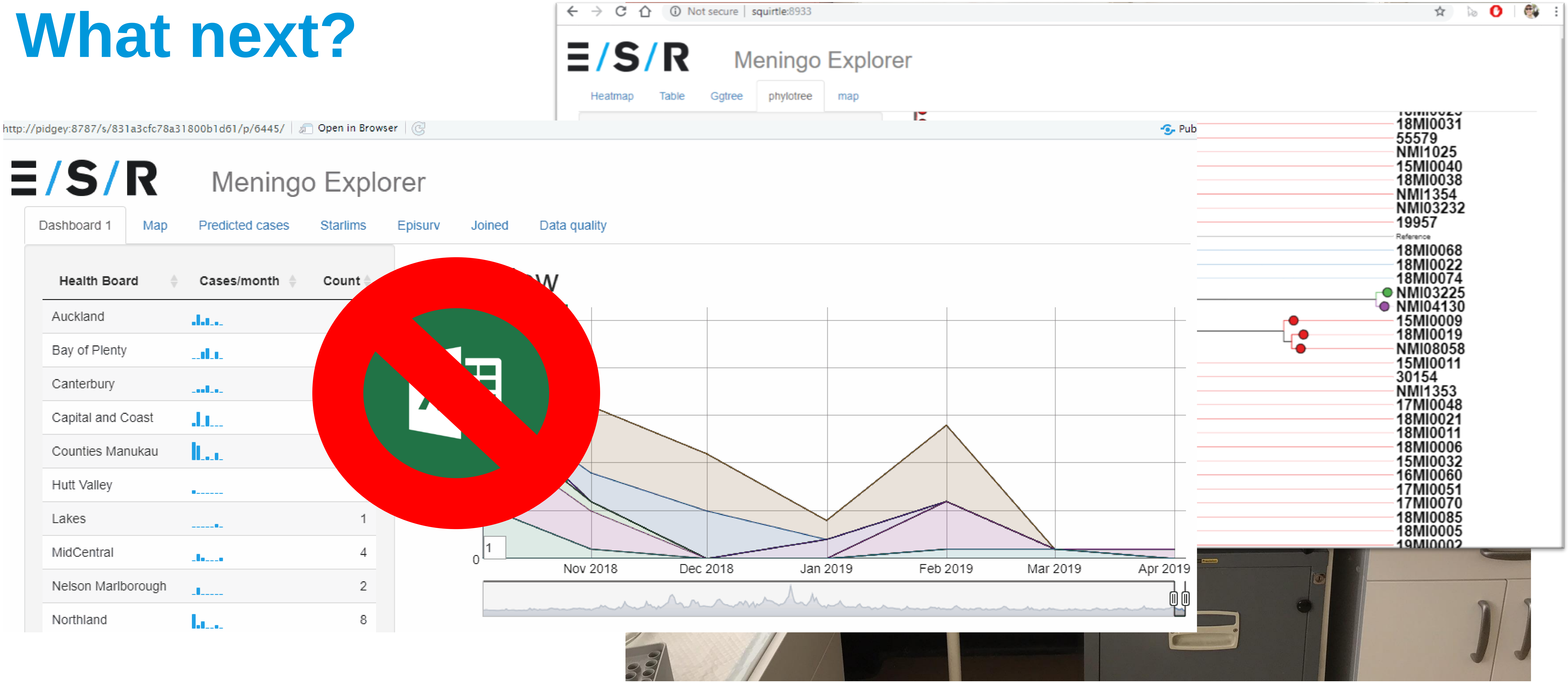
Science for Communities

Influenza Like Illness



ESR
Science for Communities

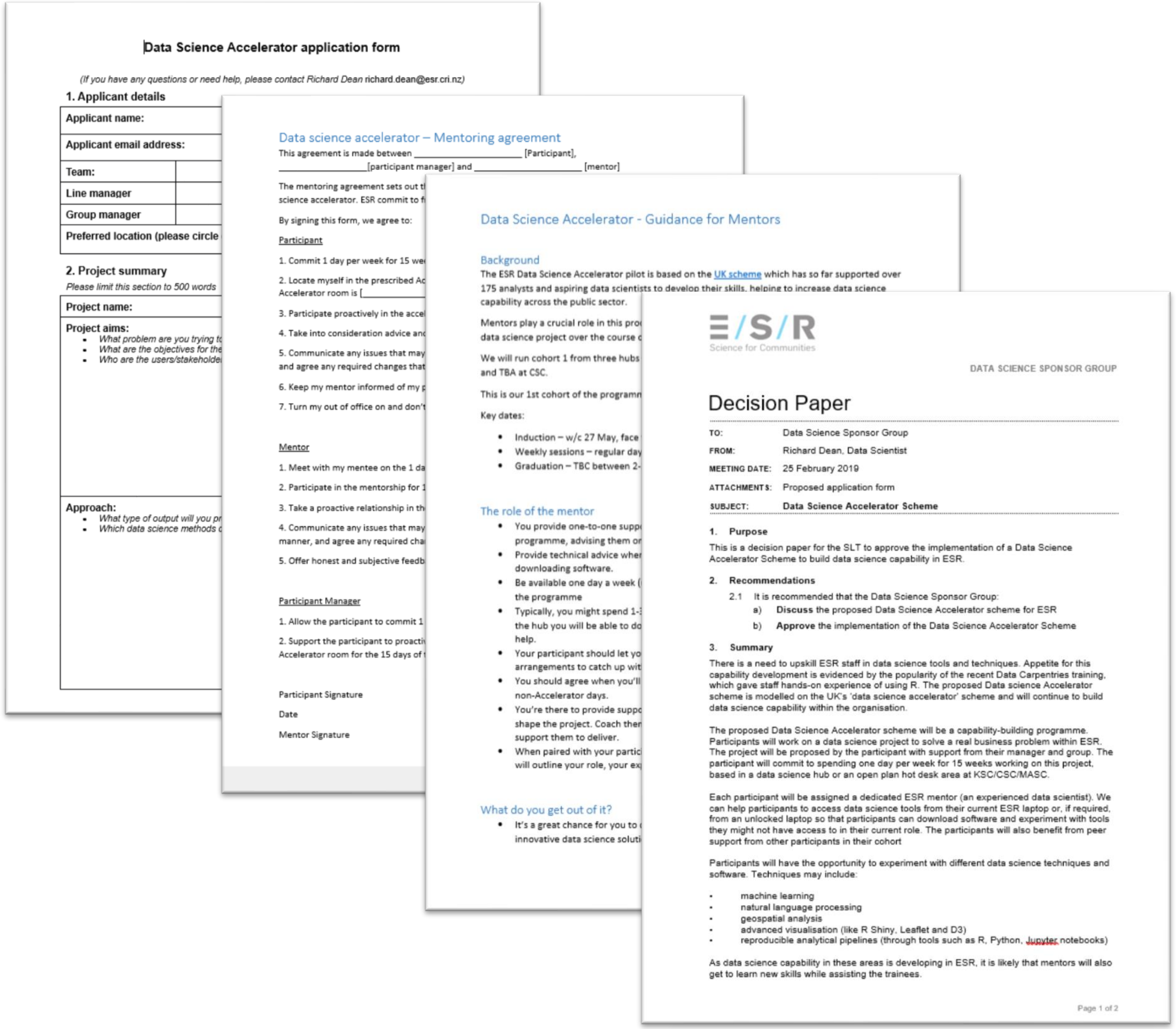
What next?



Science for Communities

Science for Communities

Preparing the groundwork



Data Science Accelerator application form

(If you have any questions or need help, please contact Richard Dean richard.dean@esr.cri.nz)

1. Applicant details

Applicant name:
Applicant email address:
Team:
Line manager
Group manager
Preferred location (please circle

2. Project summary
Please limit this section to 500 words

Project name:
Project aims:
- What problem are you trying t
- What are the objectives for the
- Who are the users/stakeholde

Approach:
- What type of output will you pr
- Which data science methods c

Data science accelerator – Mentoring agreement

This agreement is made between ________________ [Participant], ________________ [participant manager] and ________________ [mentor]

The mentoring agreement sets out t
science accelerator. ESR commit to f

By signing this form, we agree to:

Participant

1. Commit 1 day per week for 15 we
2. Locate myself in the prescribed Ac
Accelerator room is [
3. Participate proactively in the accel
4. Take into consideration advice an
5. Communicate any issues that may
and agree any required changes that
6. Keep my mentor informed of my p
7. Turn my out of office on and don'

Mentor

1. Meet with my mentee on the 1 da
2. Participate in the mentorship for 1
3. Take a proactive relationship in th
4. Communicate any issues that may
manner, and agree any required cha
5. Offer honest and subjective feedb

Participant Manager

1. Allow the participant to commit 1
2. Support the participant to proacti
Accelerator room for the 15 days of

Participant Signature
Date
Mentor Signature

Data Science Accelerator - Guidance for Mentors

Background

The ESR Data Science Accelerator pilot is based on the UK scheme which has so far supported over 175 analysts and aspiring data scientists to develop their skills, helping to increase data science capability across the public sector.

Mentors play a crucial role in this pro
data science project over the course c

We will run cohort 1 from three hubs
and TBA at CSC.

This is our 1st cohort of the programm

Key dates:

- Induction – w/c 27 May, face
- Weekly sessions – regular day
- Graduation – TBC between 2-

The role of the mentor

- You provide one-to-one supp
programme, advising them or
- Provide technical advice wher
downloading software.
- Be available one day a week (
the programme
- Typically, you might spend 1-3
the hub you will be able to do
help.
- Your participant should let yo
arrangements to catch up wit
- You should agree when you'll
non-Accelerator days.
- You're there to provide suppo
shape the project. Coach ther
support them to deliver.
- When paired with your partic
will outline your role, your ex

What do you get out of it?

- It's a great chance for you to
innovative data science soluti

E/S/R
Science for Communities

DATA SCIENCE SPONSOR GROUP

Decision Paper

TO:	Data Science Sponsor Group
FROM:	Richard Dean, Data Scientist
MEETING DATE:	25 February 2019
ATTACHMENTS:	Proposed application form
SUBJECT:	**Data Science Accelerator Scheme**

1. **Purpose**

This is a decision paper for the SLT to approve the implementation of a Data Science Accelerator Scheme to build data science capability in ESR.

2. **Recommendations**

2.1 It is recommended that the Data Science Sponsor Group:

a) **Discuss** the proposed Data Science Accelerator scheme for ESR

b) **Approve** the implementation of the Data Science Accelerator Scheme

3. **Summary**

There is a need to upskill ESR staff in data science tools and techniques. Appetite for this capability development is evidenced by the popularity of the recent Data Carpentries training, which gave staff hands-on experience of using R. The proposed Data science Accelerator scheme is modelled on the UK's 'data science accelerator' scheme and will continue to build data science capability within the organisation.

The proposed Data Science Accelerator scheme will be a capability-building programme. Participants will work on a data science project to solve a real business problem within ESR. The project will be proposed by the participant with support from their manager and group. The participant will commit to spending one day per week for 15 weeks working on this project, based in a data science hub or an open plan hot desk area at KSC/CSC/MASC.

Each participant will be assigned a dedicated ESR mentor (an experienced data scientist). We can help participants to access data science tools from their current ESR laptop or, if required, from an unlocked laptop so that participants can download software and experiment with tools they might not have access to in their current role. The participants will also benefit from peer support from other participants in their cohort

Participants will have the opportunity to experiment with different data science techniques and software. Techniques may include:

- machine learning
- natural language processing
- geospatial analysis
- advanced visualisation (like R Shiny, Leaflet and D3)
- reproducible analytical pipelines (through tools such as R, Python, Jupyter notebooks)

As data science capability in these areas is developing in ESR, it is likely that mentors will also get to learn new skills while assisting the trainees.

Page 1 of 2

Native vegetation wastewater remote sensor trial 2019

E/S/R
Science for Communities

The pitch…



#007
Squirtle
Water

(Level 16)

#008
Wartortle
Water

(Level 36)

#009
Blastoise
Water

E/S/R
Science for Communities

Some data science accelerator ideas

Last Saturday's walk along the Paekakariki escarpment track

Science for Communities

4 Projects

Christchurch
Creating a data visualization platform from **genomics** and **foodborne pathogens**

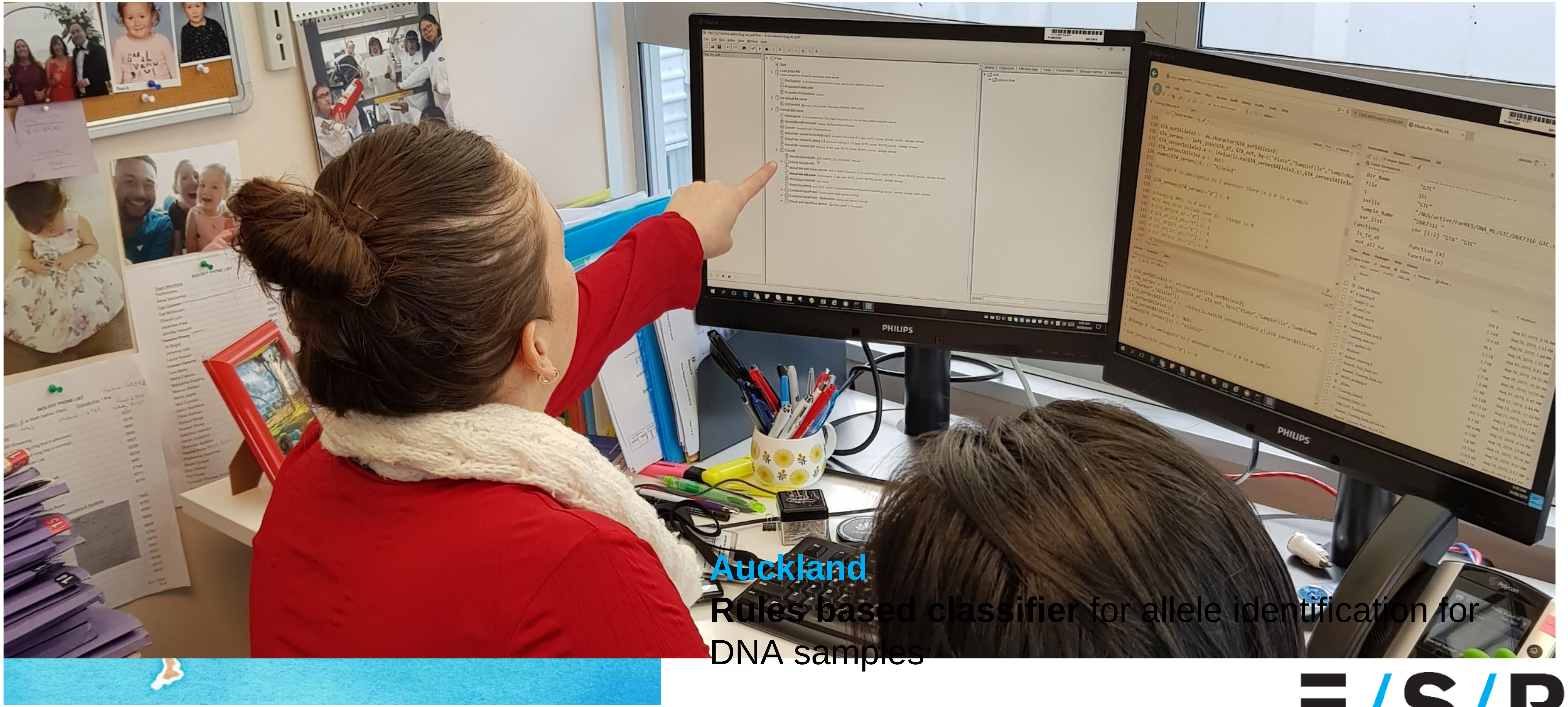
Wellington
Improving health intelligence reporting and visualisation: A prototype for **invasive pneumococcal disease**

Effective use of data from bacteria with acquired **carbapenemase genes**

Auckland
Rules based classifier for allele identification for DNA samples

Science for Communities

4 Projects

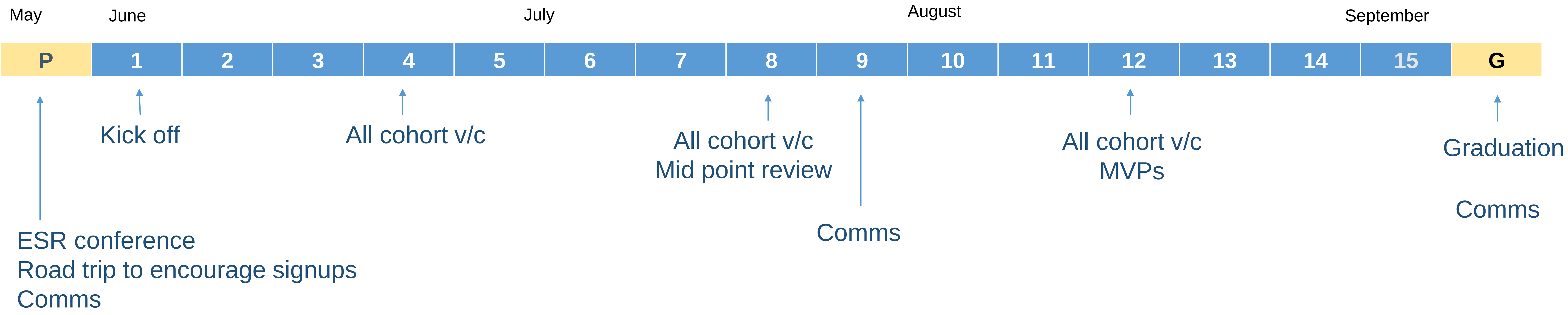


Auckland

Rules based classifier for allele identification for DNA samples

Science for Communities

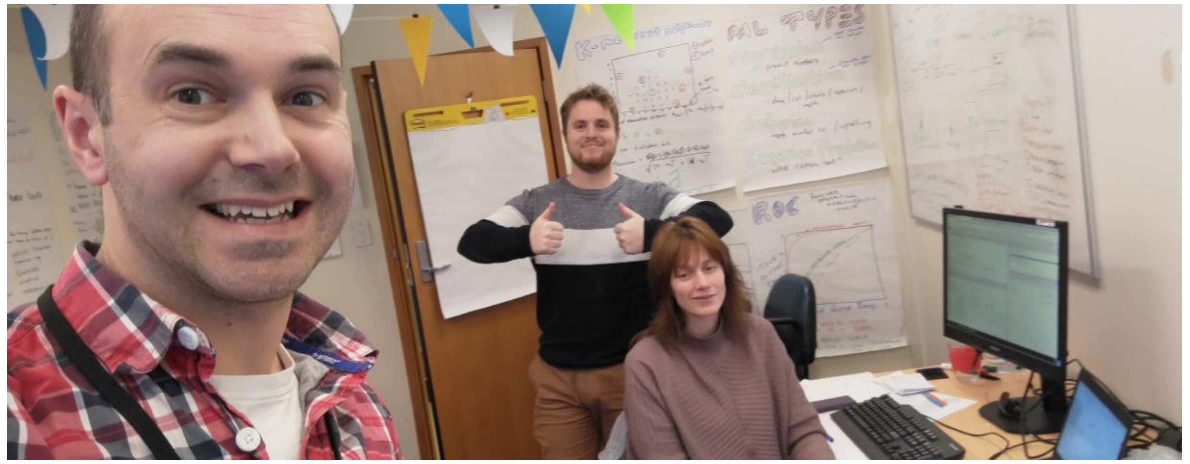
How we managed things



- ESR systems
- GIT
- Diary

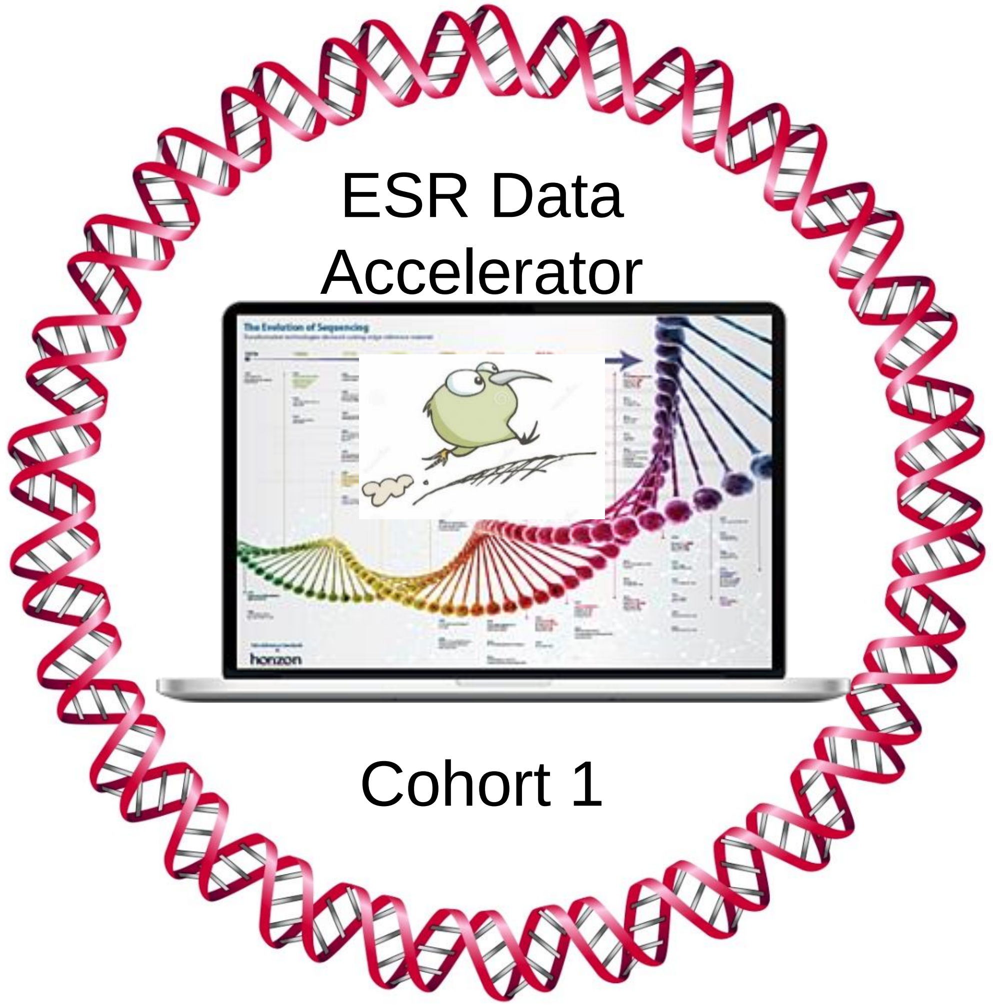
- Video conferences
- Graduation

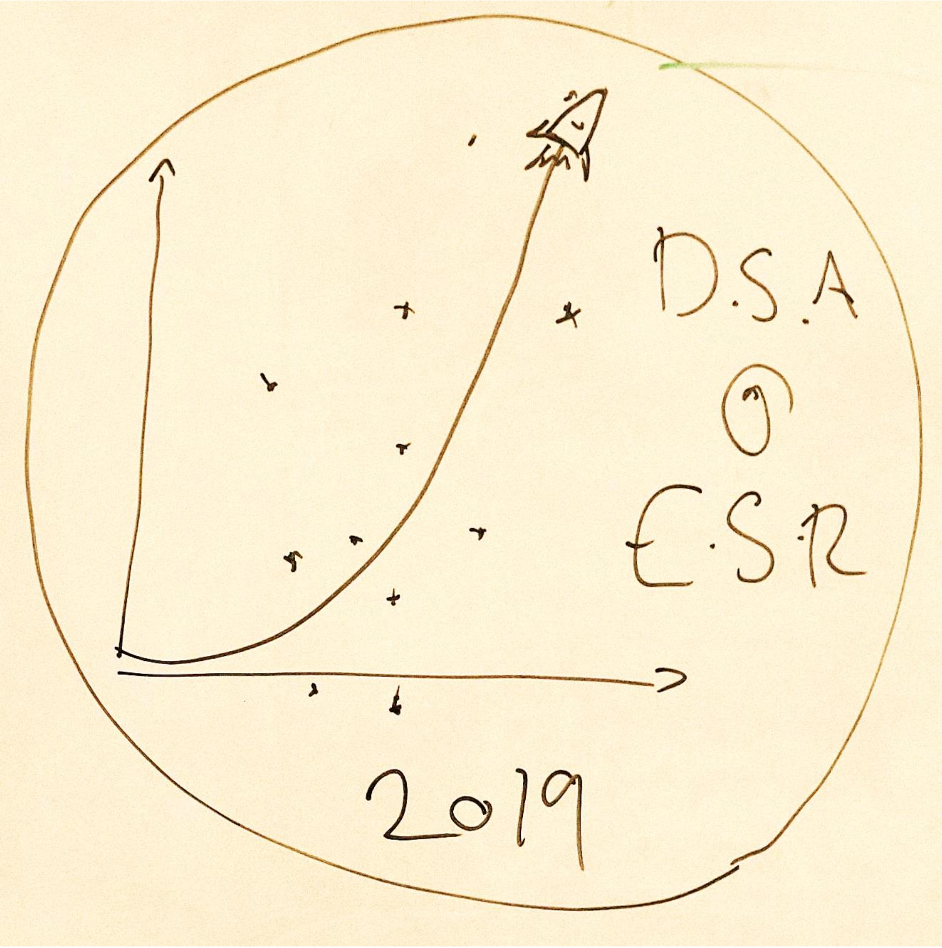
Fun!

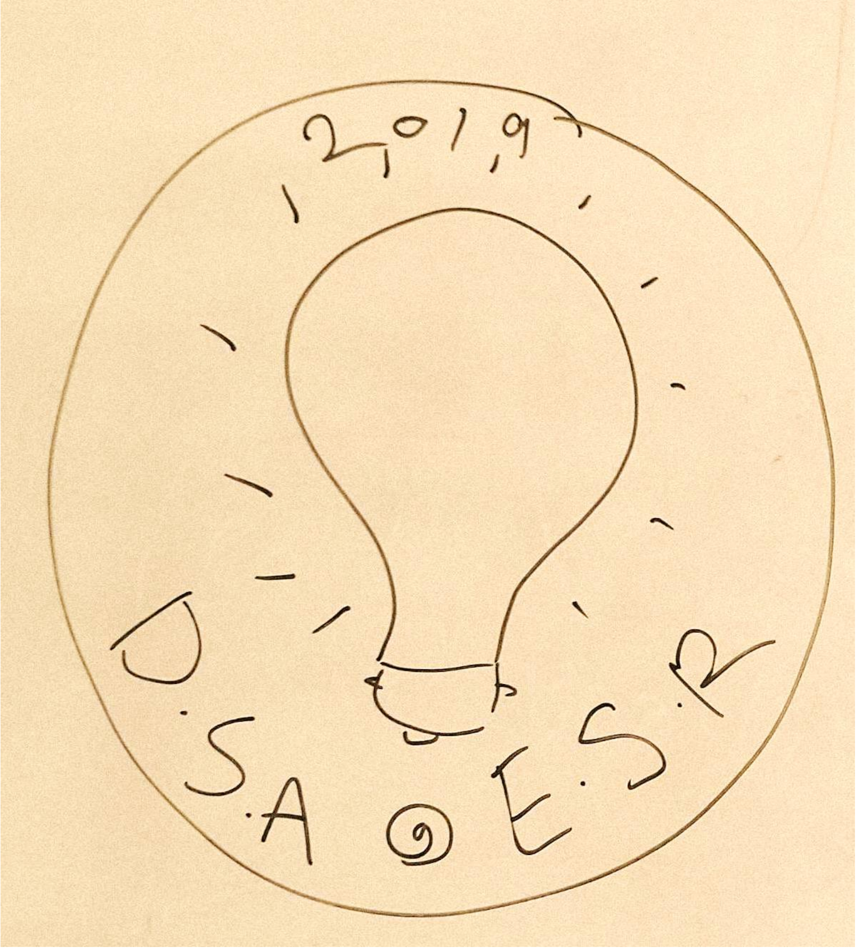


E/S/R
Science for Communities

E/S/R
Science for Communities





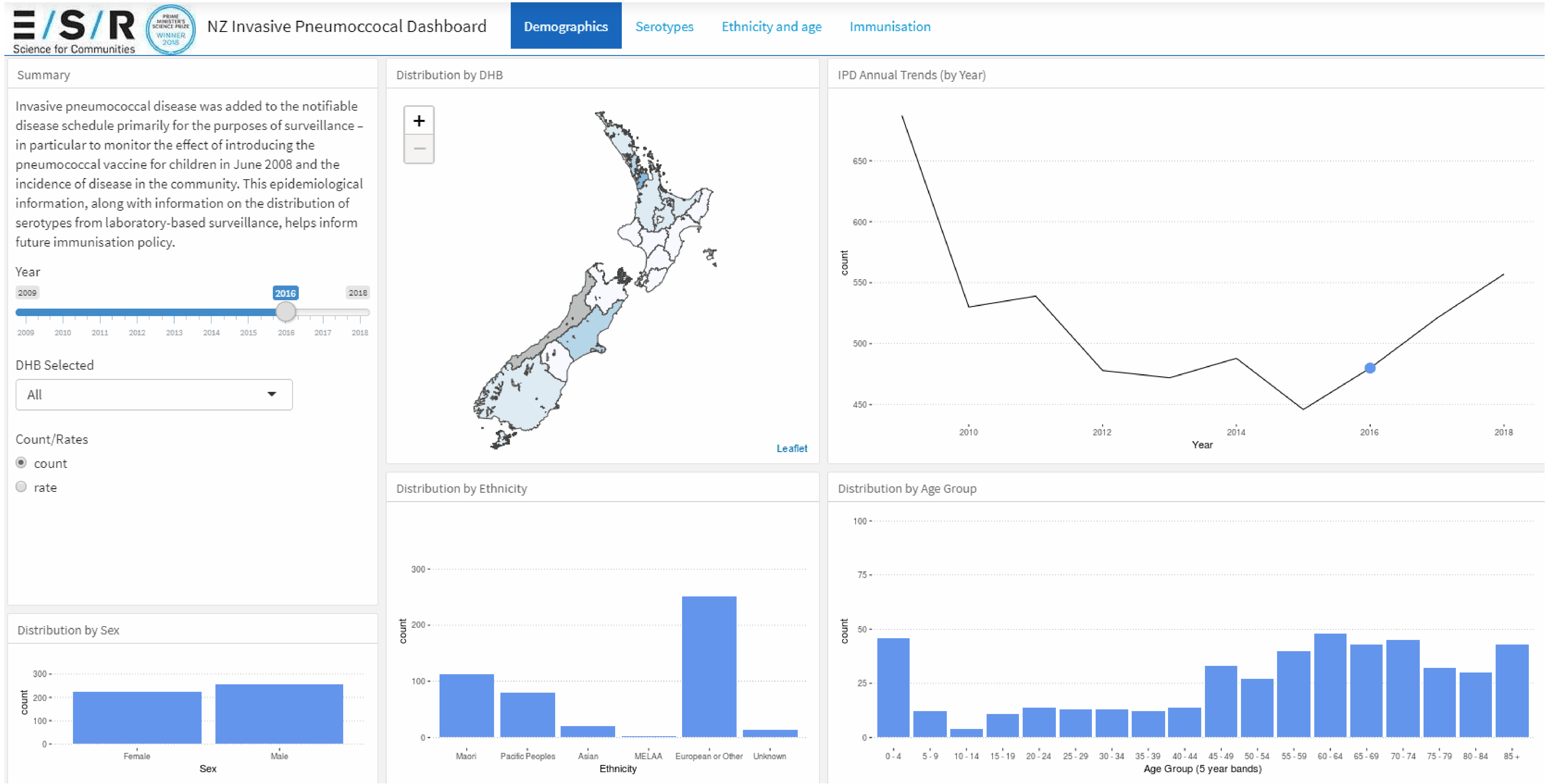


E/S/R
Science for Communities

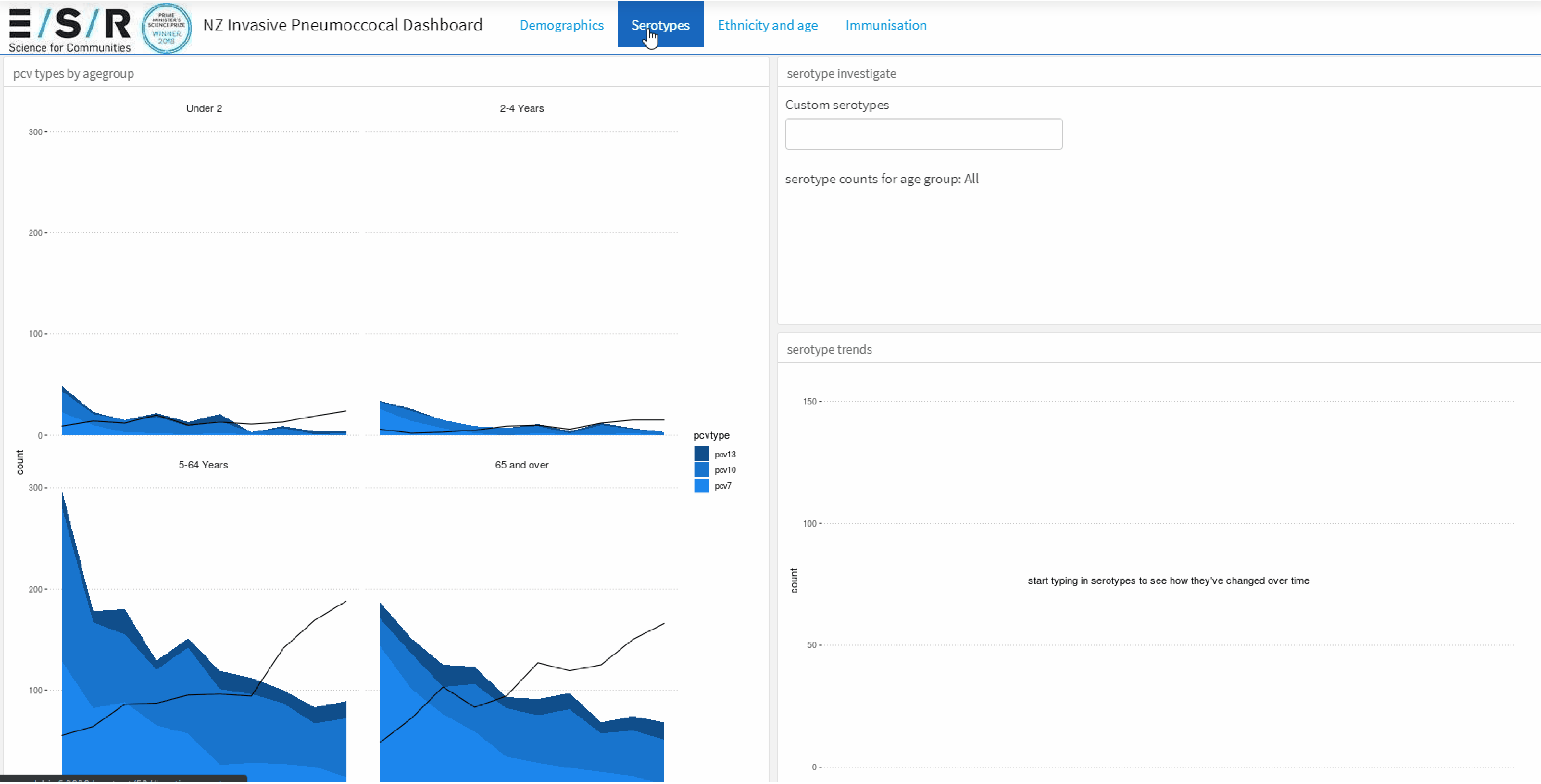
Results

Giles Graham

Investigating methods to automate some production of the invasive pneumococcal disease report



Science for Communities



http://kscprod-bio6:3939/connect/#/apps/58/access

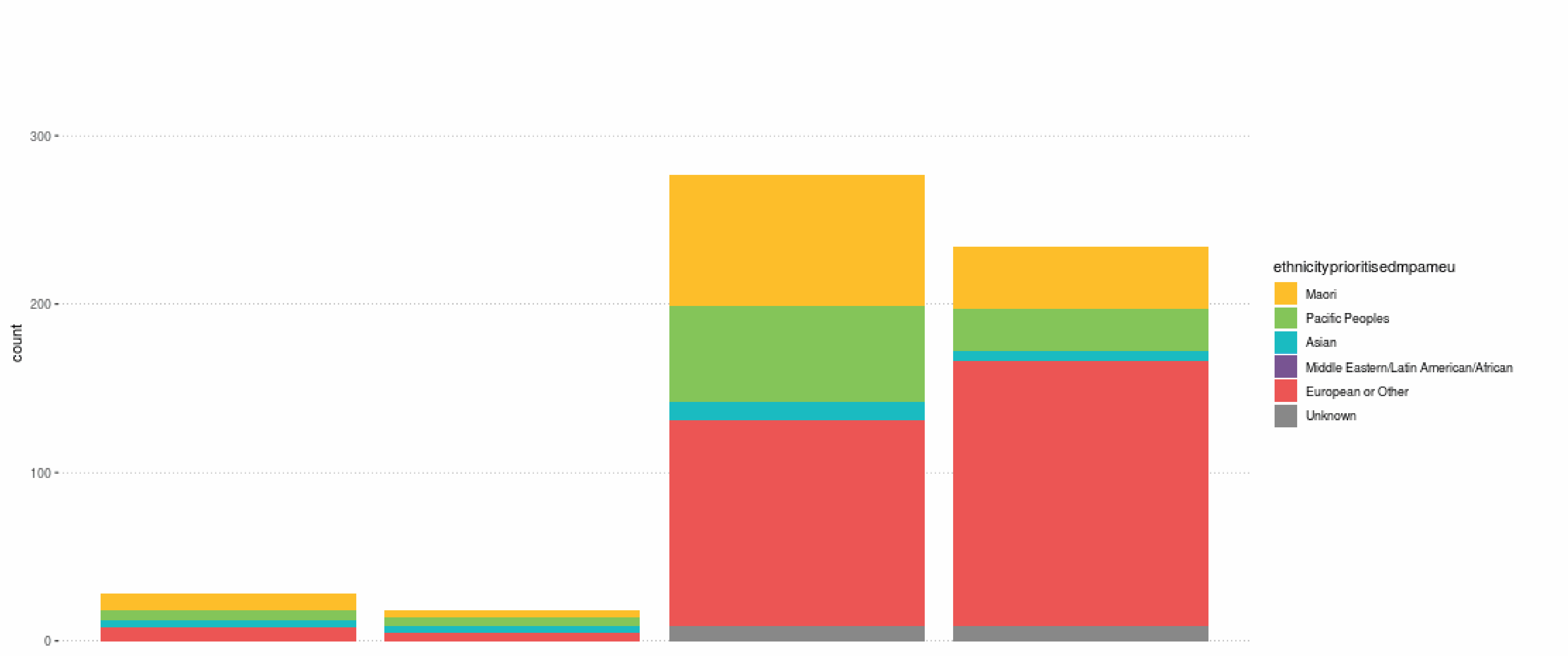
Science for Communities

Science for Communities

NZ Invasive Pneumococcal Dashboard

Demographics | Serotypes | Ethnicity and age | Immunisation

Ethnicity by Age Graph



Age Distribution by Ethnicity

Under 2	10.00	6.00	4.00	0.00	8.00	0.00
2-4 Years	4.00	5.00	4.00	0.00	5.00	0.00
5-64 Years	78.00	57.00	11.00	0.00	122.00	9.00
65 and over	37.00	25.00	6.00	0.00	157.00	9.00

Controls

Year

2009 2018

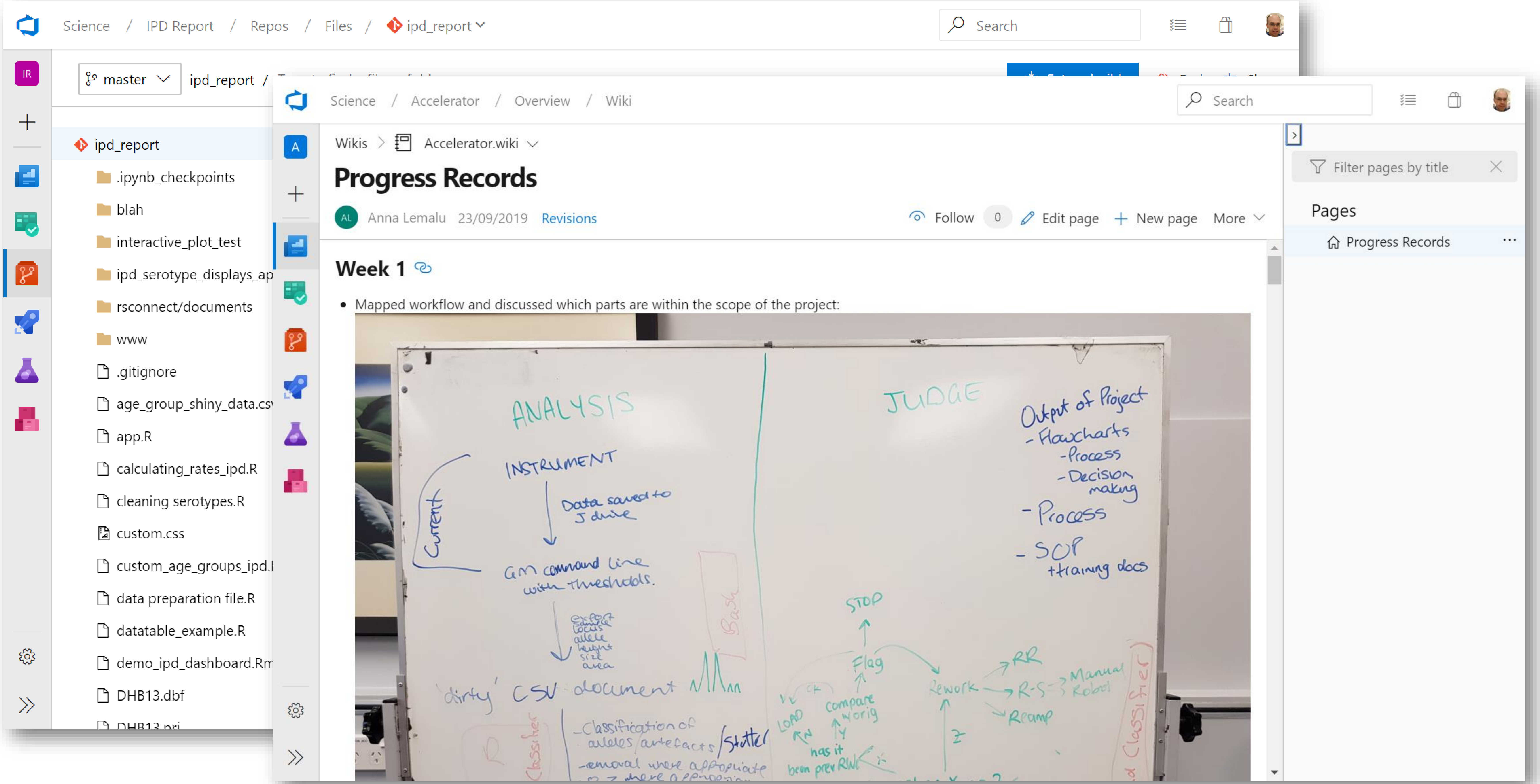
2013 2017

Discussion

a bunch of stuff about age groups and ethnicities will go here

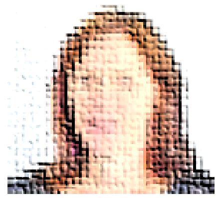
http://kscprod-bio6:3939/connect/#/apps/58/access

Science for Communities

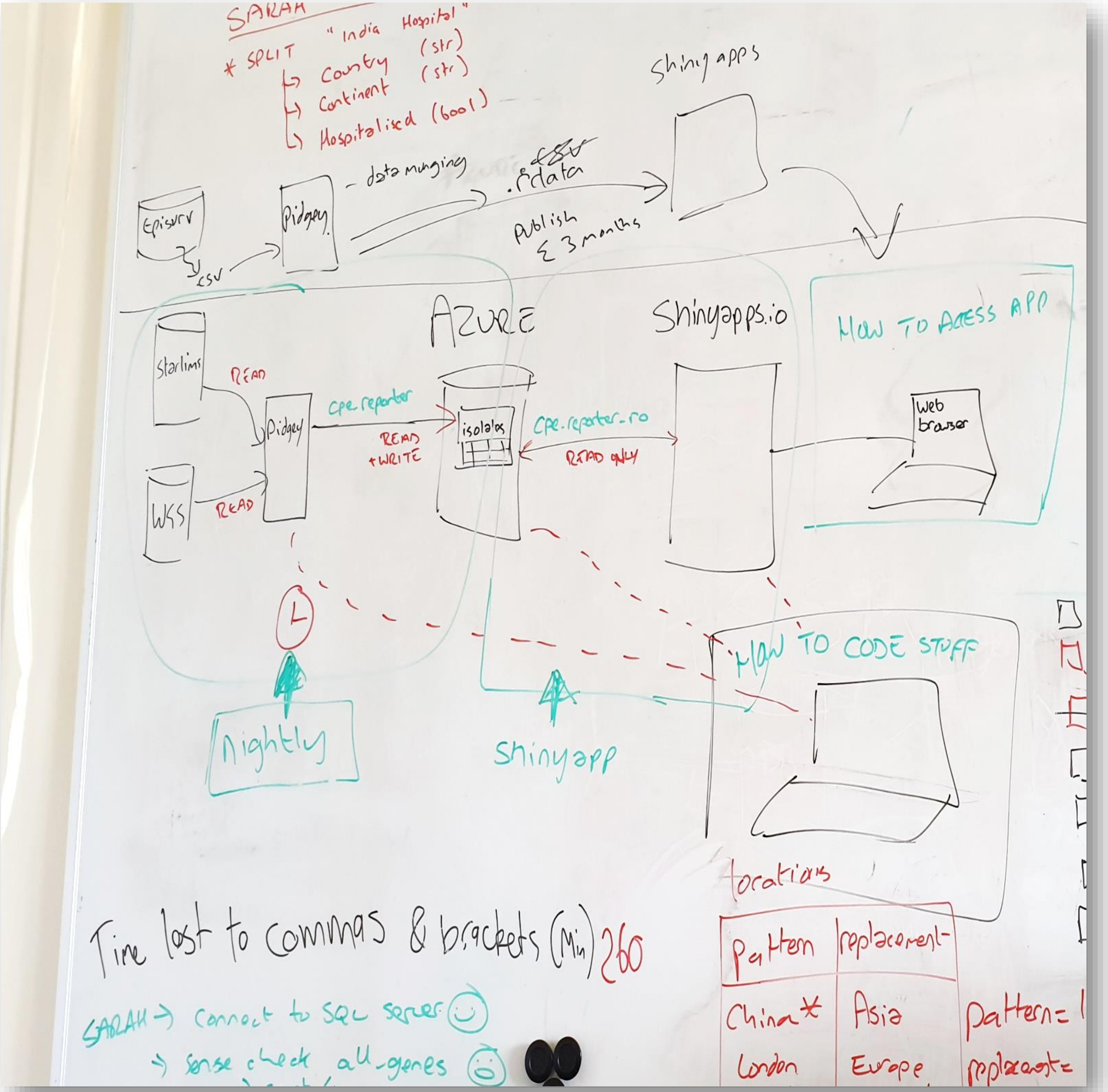


Science for Communities

Effective use of data from bacteria with acquired **carbapenemase** genes



Sarah Bakker

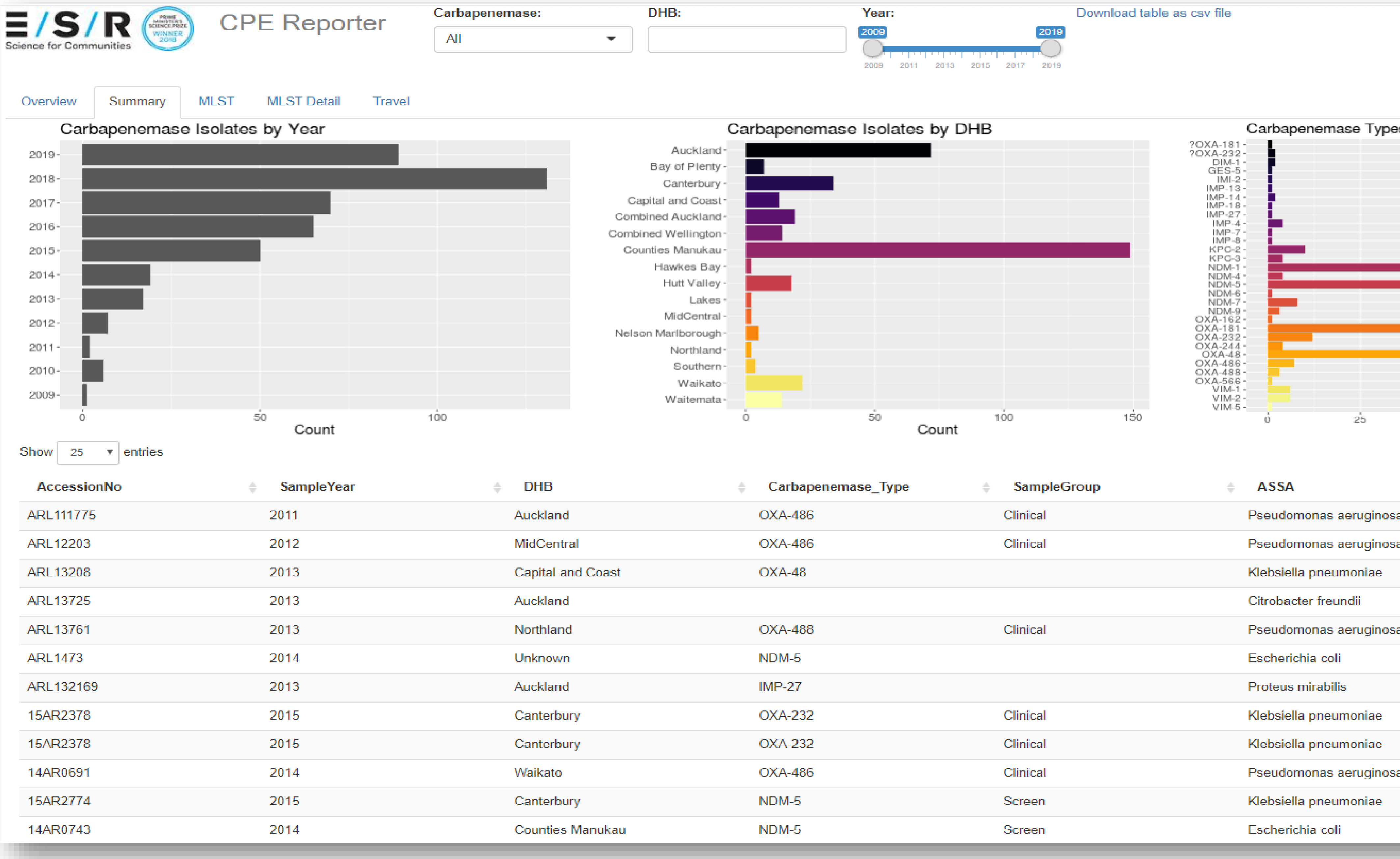


Week 1-5

- Steep learning curve!
- Getting access to equipment and permission to folders and software.
- Learning R.
- Scoping different platforms for displaying data.
- Learning R shinyApp.
- Request for data from BI team.
- Lots of homework!!

Weeks 6-10

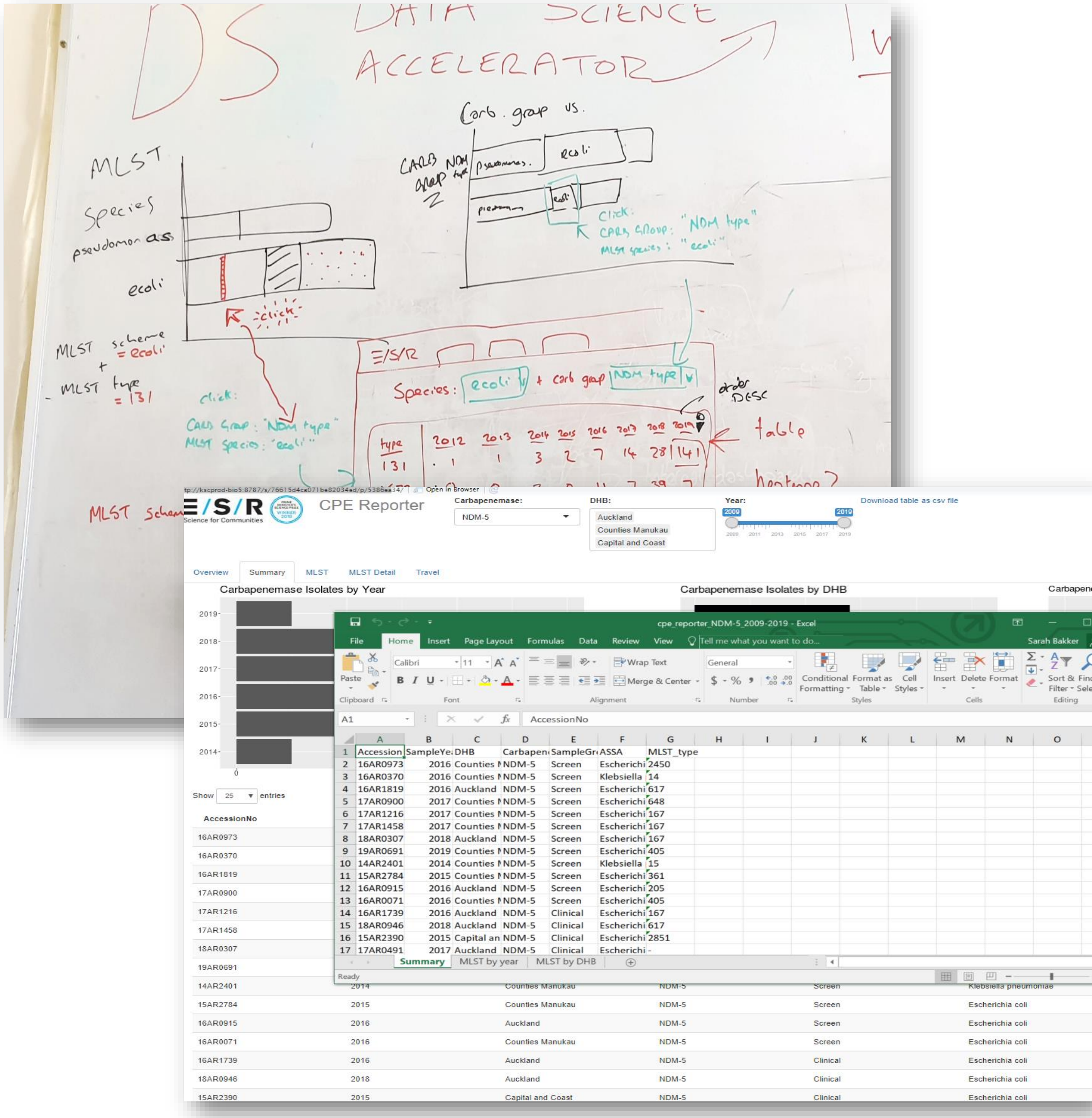
- More scoping of platforms for displaying the data.
- Develop more R coding knowledge while waiting for SQL data from BI team
  - data transformation
  - data wrangling
  - data visualization
  - Building an R shinyApp



Science for Communities

Weeks 11-15

- After lots of discussions, finally get useable data from BI team.
- However, data not user friendly and will need a lot of cleaning…
- Remaining time on programme has been spent cleaning data, re-assessing, more cleaning – iterative process that is not finished yet!



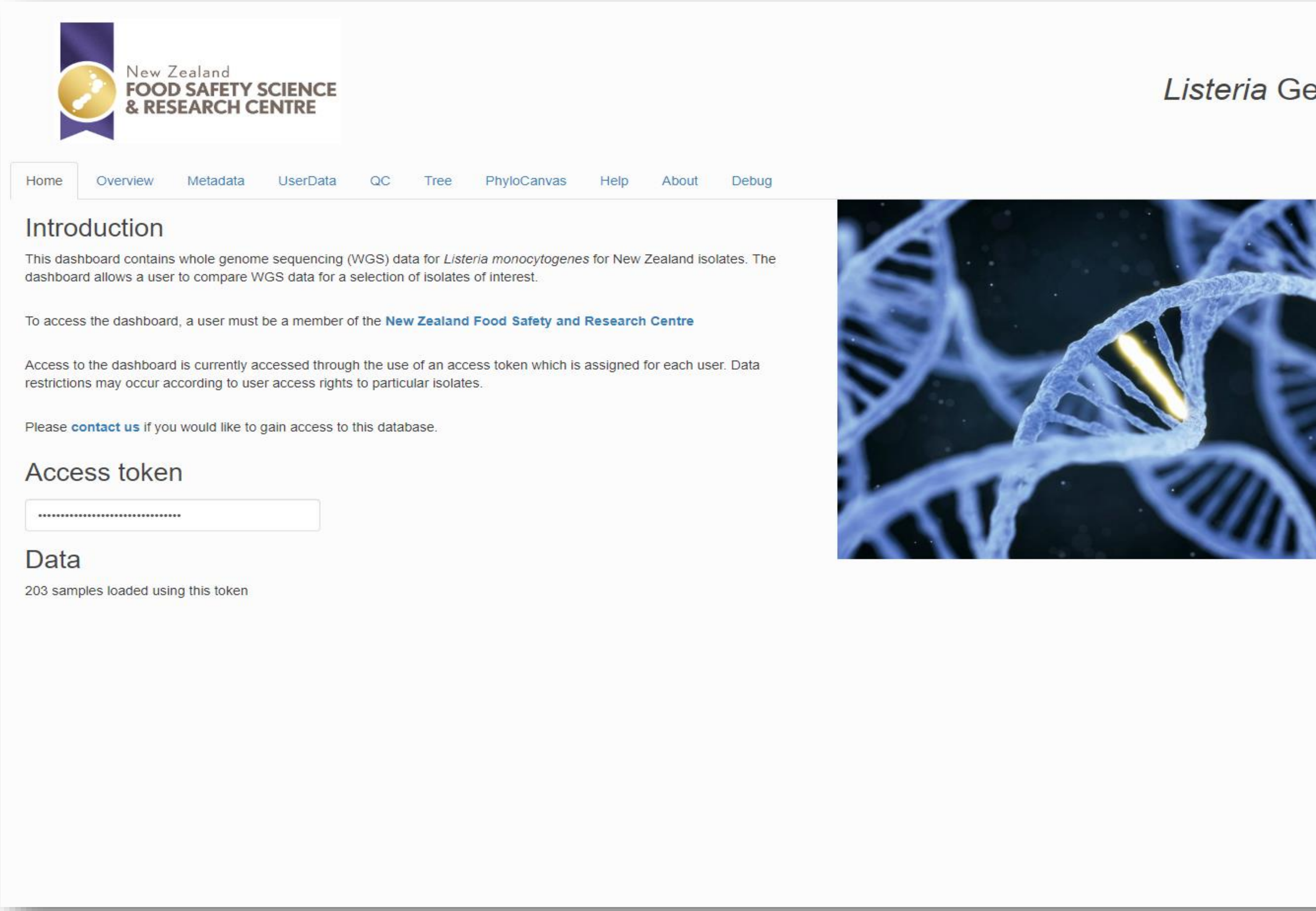
Creating a **data visualization** platform for **genomics** and **foodborne pathogens**


Lucia Rivas


Pierre Dupont

Collating whole genome sequencing data for Listeria across industry and research sectors

- Listeria genomic database via API
- Creating a dashboard for industry for food safety purposes
- External authentication
- Blog: NZFSSRC Listeria



E/S/R
Science for Communities

New Zealand FOOD SAFETY SCIENCE & RESEARCH CENTRE

Agreed level of meta data through consultation with owners of data (MPI, MoH, industry)

Listeria Genomics Dashboard

Logged in under MPI user group – restricted information specific for client is in view but not for others

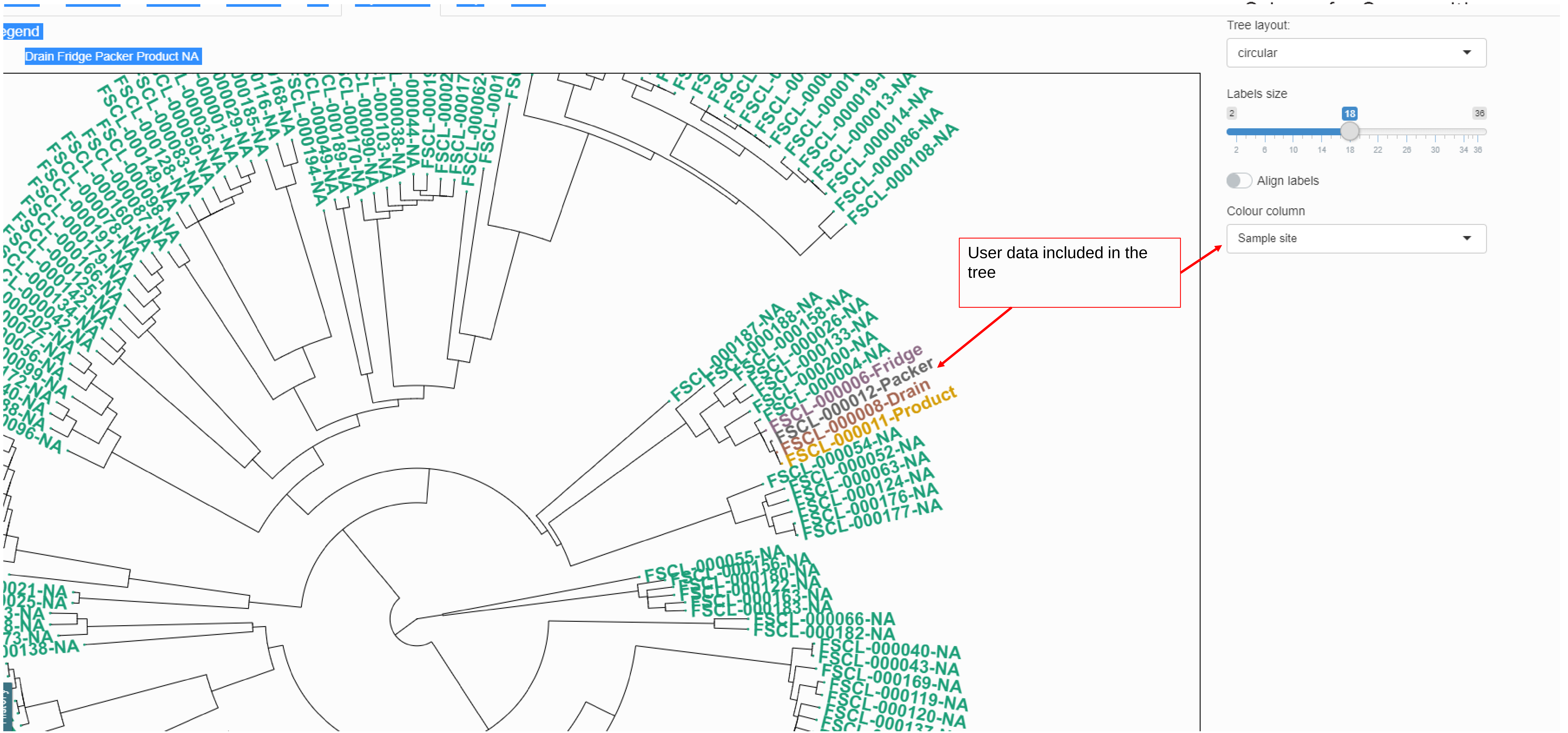
Home | Overview | Metadata | UserData | QC | PhyloCanvas | Help | About

Select/Deselect all

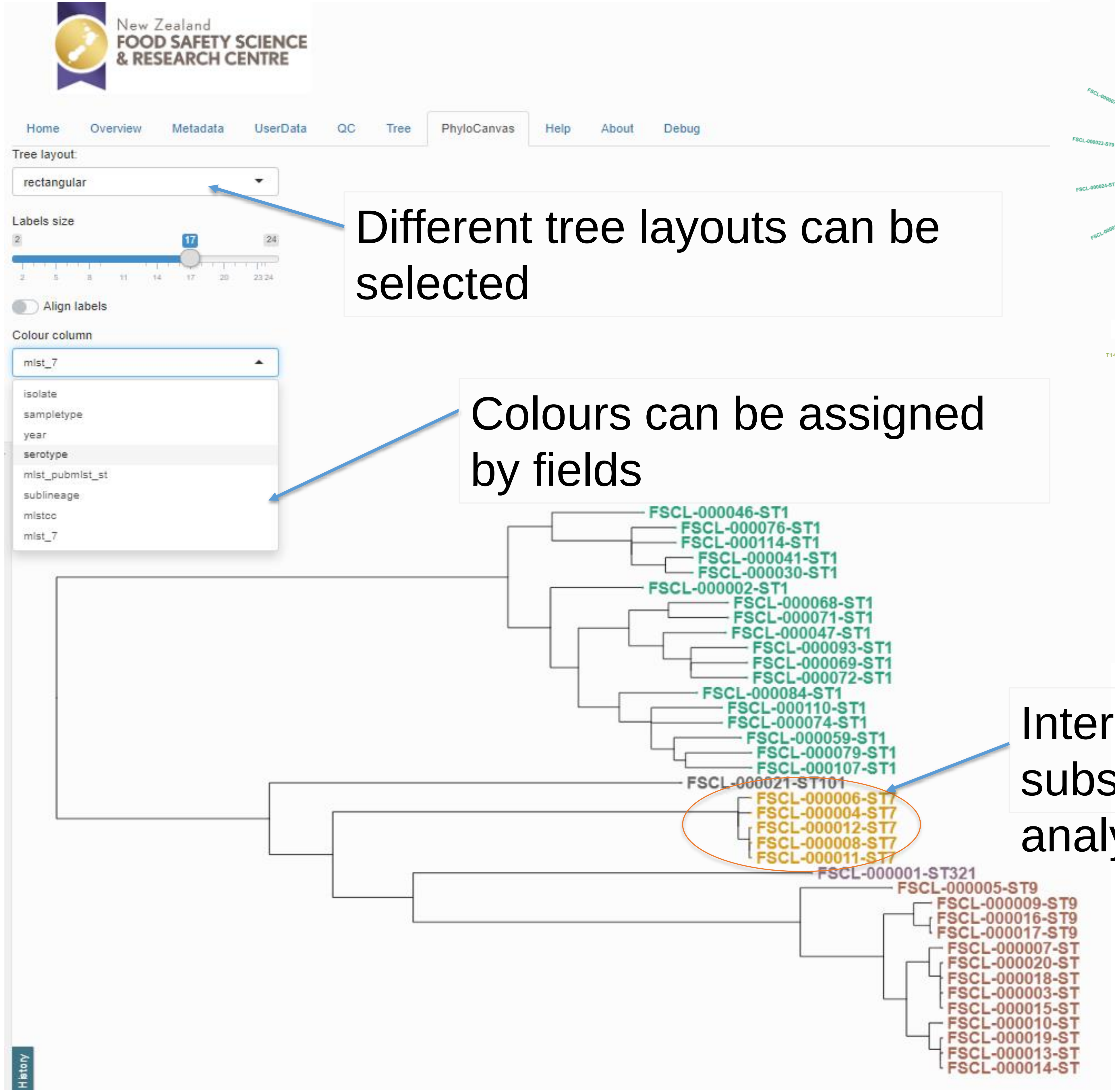
Show 25 entries

Search:

	Isolate	Species	Serotype	SampleType	Country	Year	Identifier	ProjectClient	SequenceType	Sublineage	AlleleCode	CC	Lineage
	All		All	All		All	All	All	All	All	All	All	All
1	FSCL-000001	Listeria monocytogenes	O1/2	Non clinical	New Zealand	2012	CPH1212437-2441	MPI	ST321	SL321	14.16.17	CC321	II
2	FSCL-000005	Listeria monocytogenes	O1/2	Non clinical	New Zealand	2012	CPH1213390-3392-13	MPI	ST9	SL9	19.42.89	CC9	II
3	FSCL-000006	Listeria monocytogenes	O1/2	Non clinical	New Zealand	2012	CPH1212412	MPI	ST7	SL7	23.52.65.52.40	CC7	II
4	FSCL-000007	Listeria monocytogenes	O1/2	Non clinical	New Zealand	2012	CPH1212413	MPI	ST9	SL9	19.53.66.53.41	CC9	II
5	FSCL-000008	Listeria monocytogenes	O1/2	Non clinical	New Zealand	2012	CPH1213402-3404-4	MPI	ST7	SL7	23.52.65.52.40.49	CC7	II
6	FSCL-000009	Listeria monocytogenes	O1/2	Non clinical	New Zealand	2012	CPH1213393-3397-13	MPI	ST9	SL9	19.53.88.86.73	CC9	II
7	FSCL-000010	Listeria monocytogenes	O1/2	Non clinical	New Zealand	2012	CPH1212387	MPI	ST9	SL9	19.53.66	CC9	II
8	FSCL-000011	Listeria monocytogenes	O1/2	Non clinical	New Zealand	2012	CPH1212392	MPI	ST7	SL7	23.52.65	CC7	II
9	FSCL-000012	Listeria	O1/2	Non clinical	New Zealand	2012	CPH1212393	MPI	ST7	SL7	23.52.65.52.40.49	CC7	II

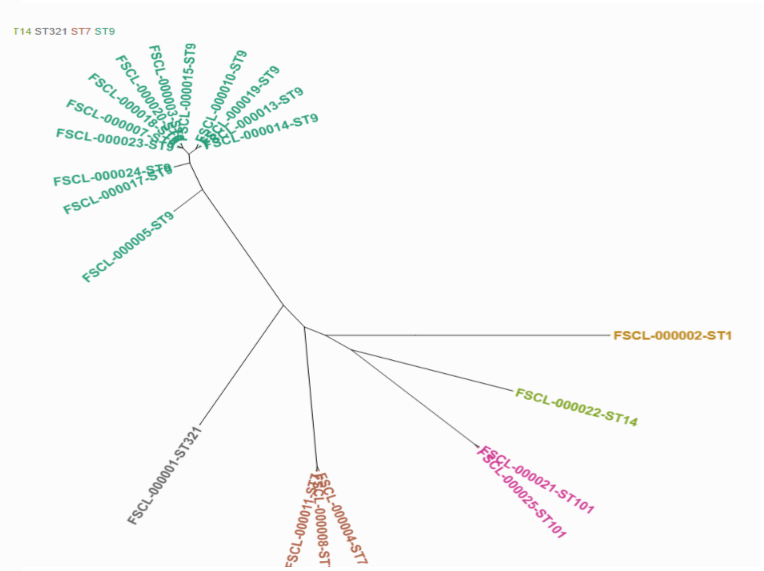


egend
Drain Fridge Packer Product NA
Tree layout:
circular
Labels size
2
18
36
Align labels
Colour column
Sample site
User data included in the tree
FSCL-000006-Fridge
FSCL-000012-Packer
FSCL-000008-Drain
FSCL-000011-Product



Different tree layouts can be selected

Colours can be assigned by fields



Interactive – can select a subset for further analysis

QC metrics for WGS

Home | Overview | Metadata | UserData | QC | PhyloCanvas | Help | About

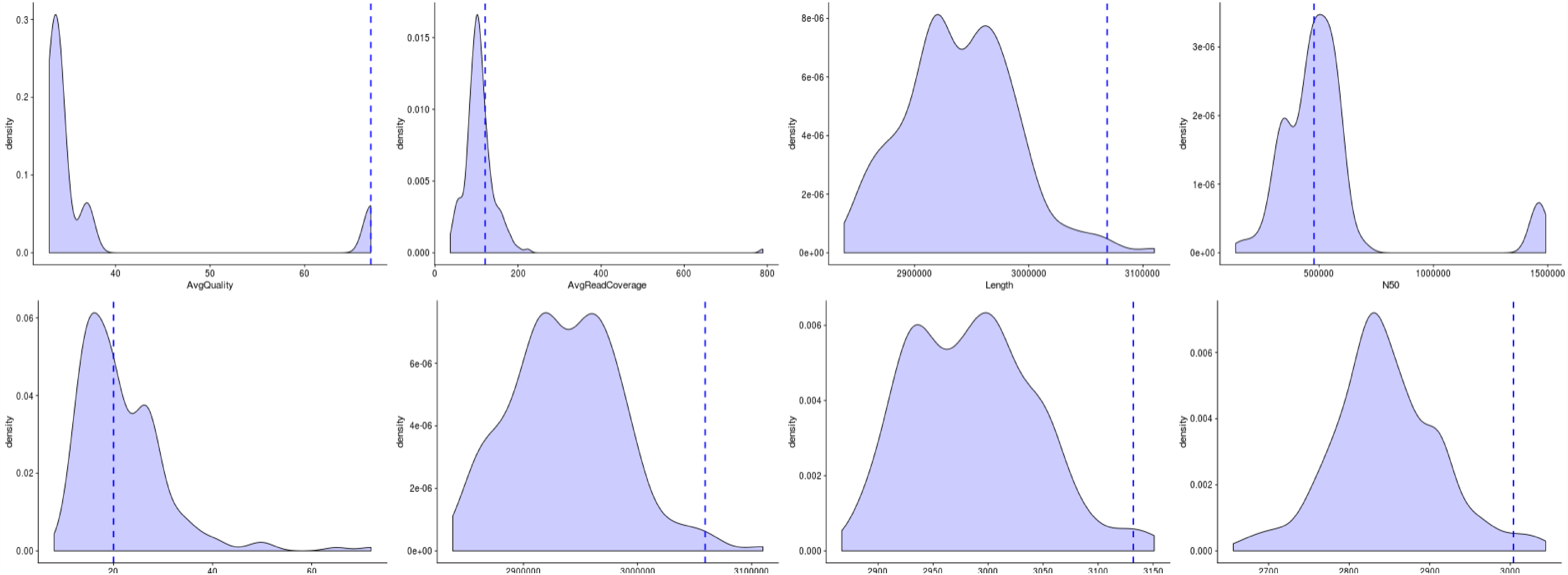
Show 5 entries

Search:

	Isolate	AvgQuality	AvgReadCoverage	Length	N50	Contigs	ACTG_bases	BAFPresent	BAFPerfect
1	FSCL-000001	67	121	3068674	477952	20	3059237	3132	3004
2	FSCL-000005	67	84	2948865	449139	19	2943356	3030	2917
3	FSCL-000006	67	202	2876044	1458346	16	2872490	2937	2815
4	FSCL-000007	67	191	2946098	604853	13	2940982	3021	2908
5	FSCL-000008	67	168	2876408	1458317	16	2872507	2937	2816

Showing 1 to 5 of 203 entries

Previous 1 2 3 4 5 ... 41 Next

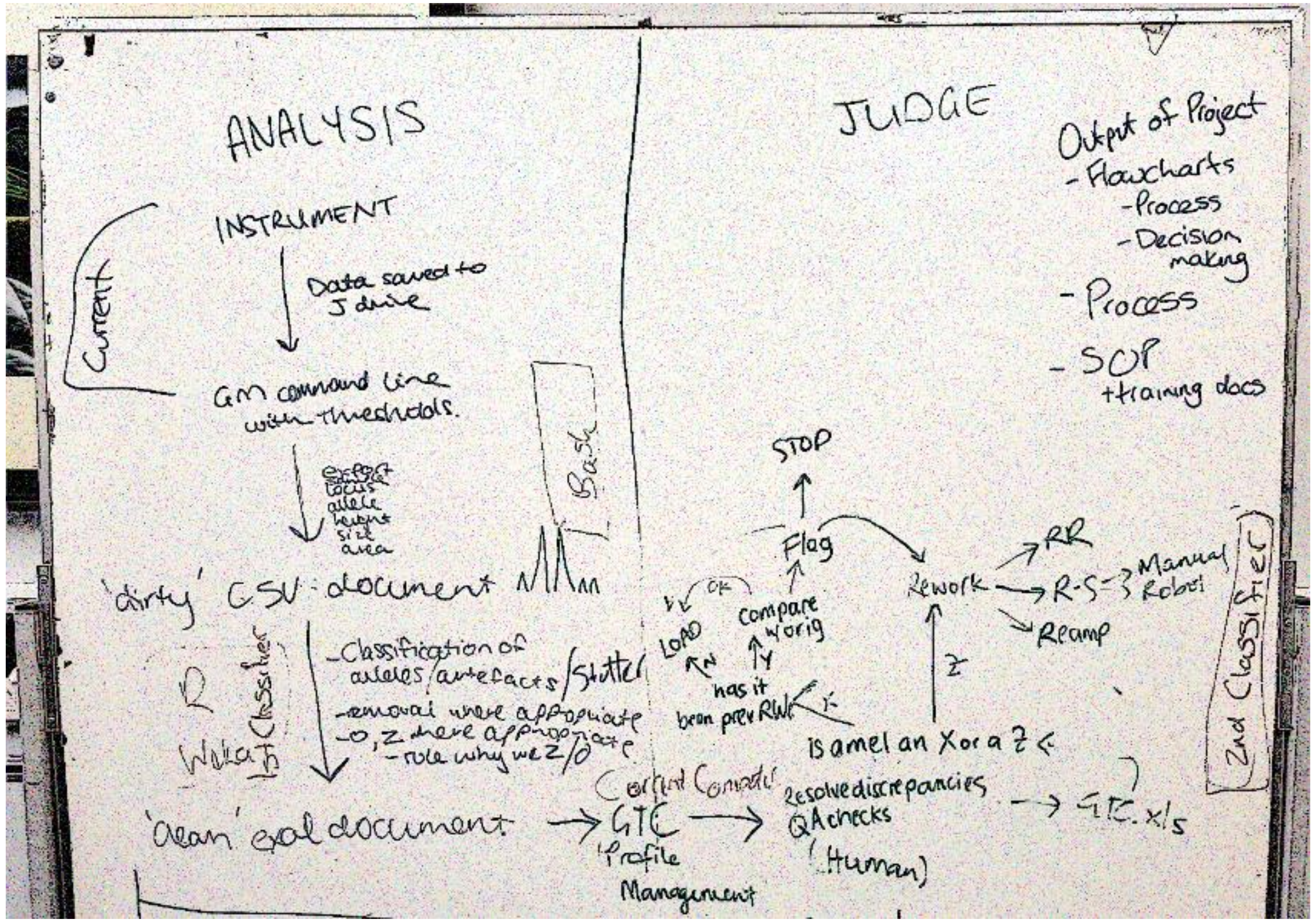


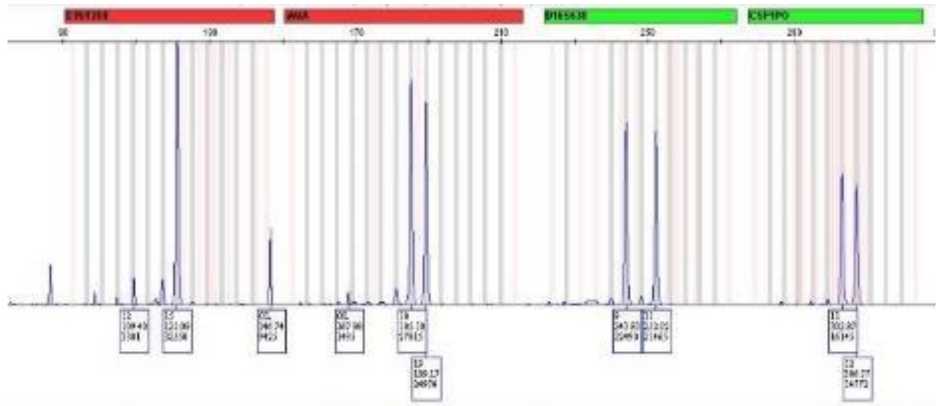
Rules based classifier for allele identification for DNA samples


Maria van der Salm


Anna Lemalu

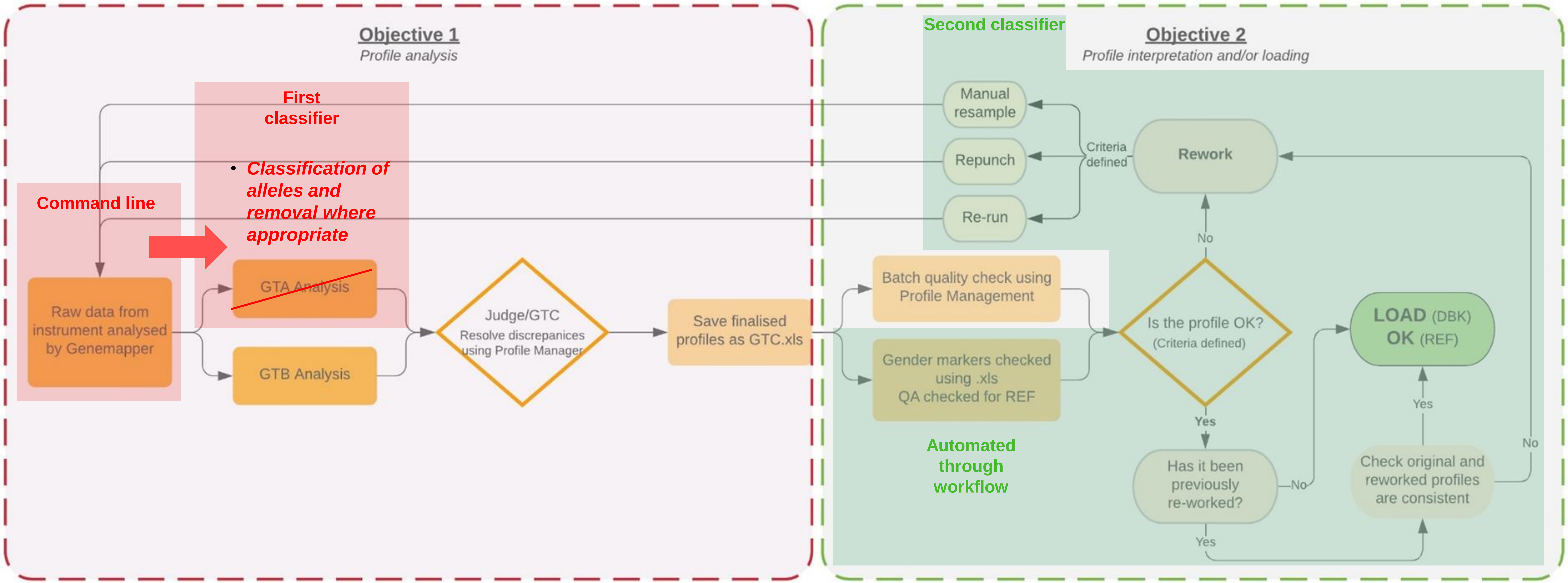
- Mentored by 2 data scientists at our Mt Albert Science Centre
- Weka for Machine Learning
- R for manipulation
- Command line for analysis methods
- Problems with unbalanced data
- Surprise visitors!



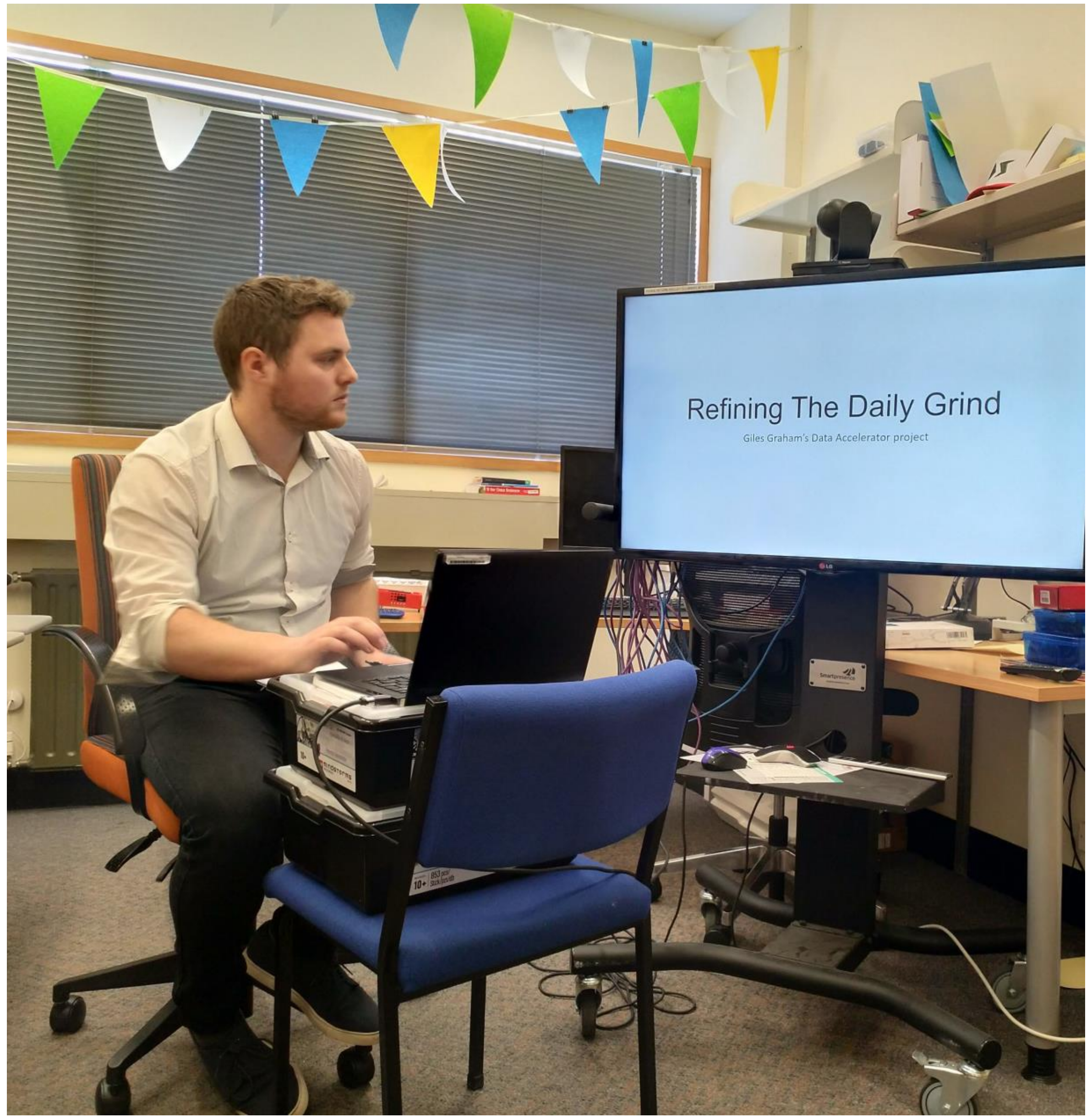


Science for Communities

Rules based classifier for allele identification for DNA samples

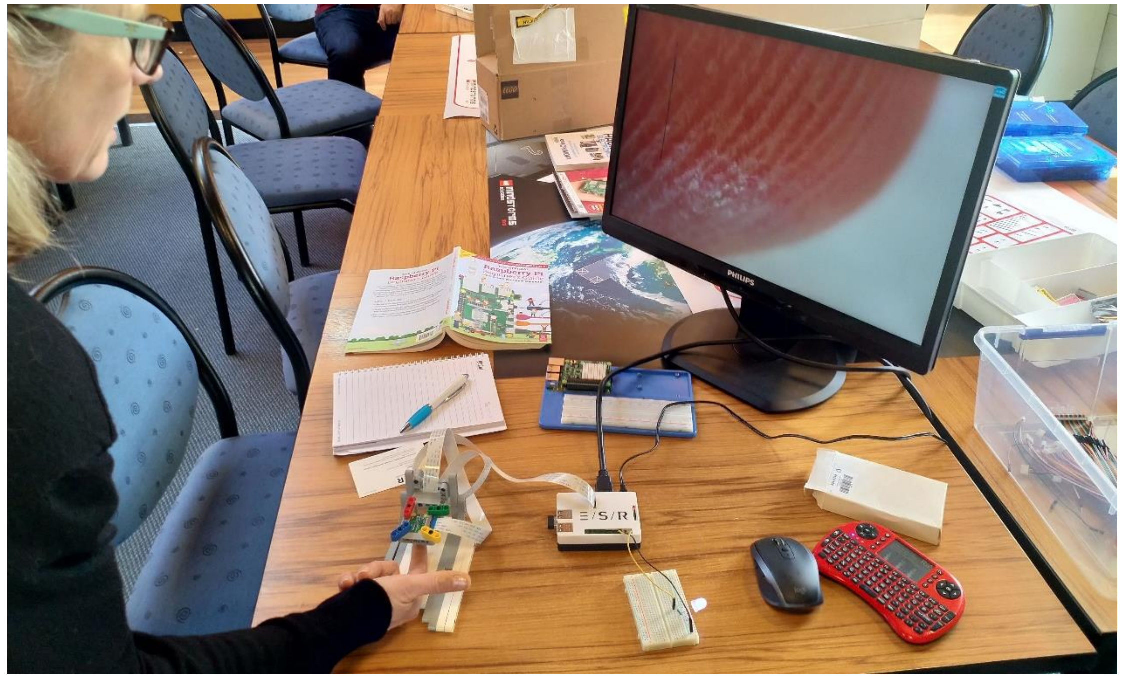


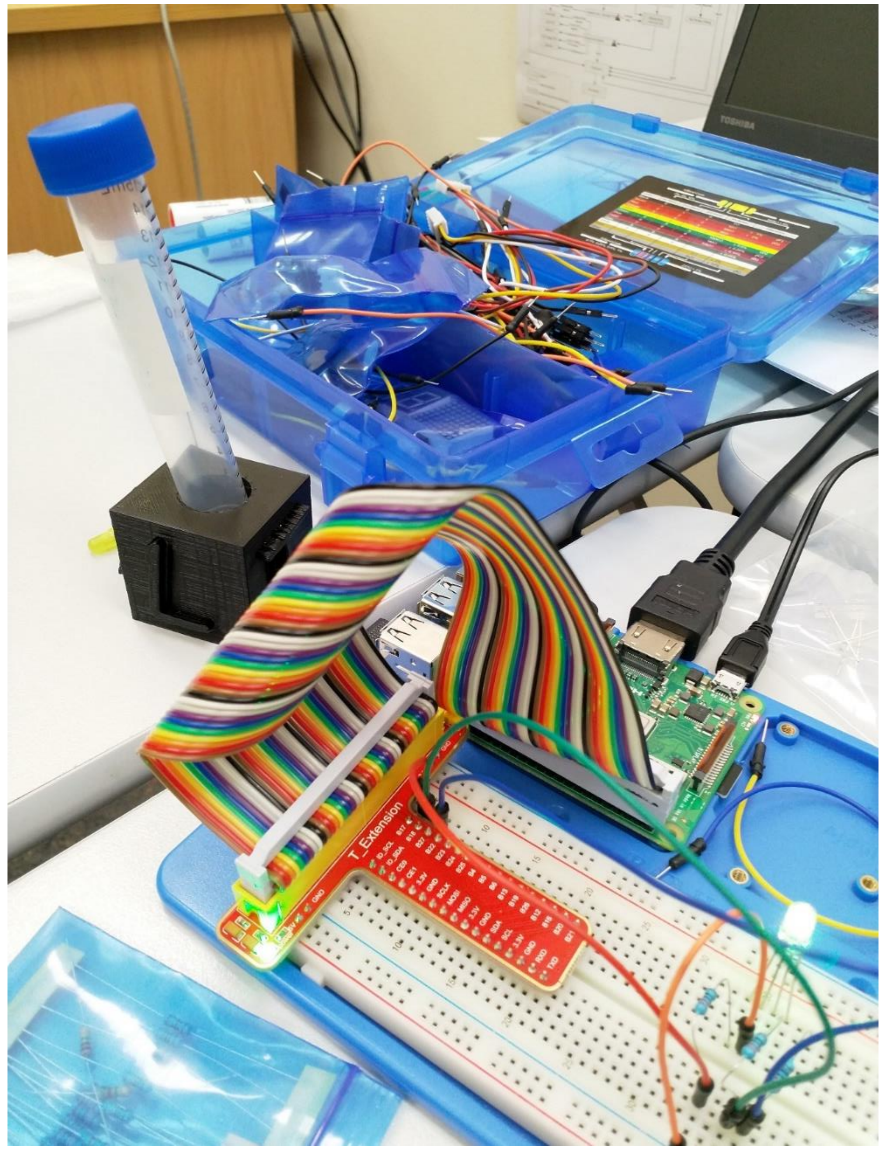
Graduation



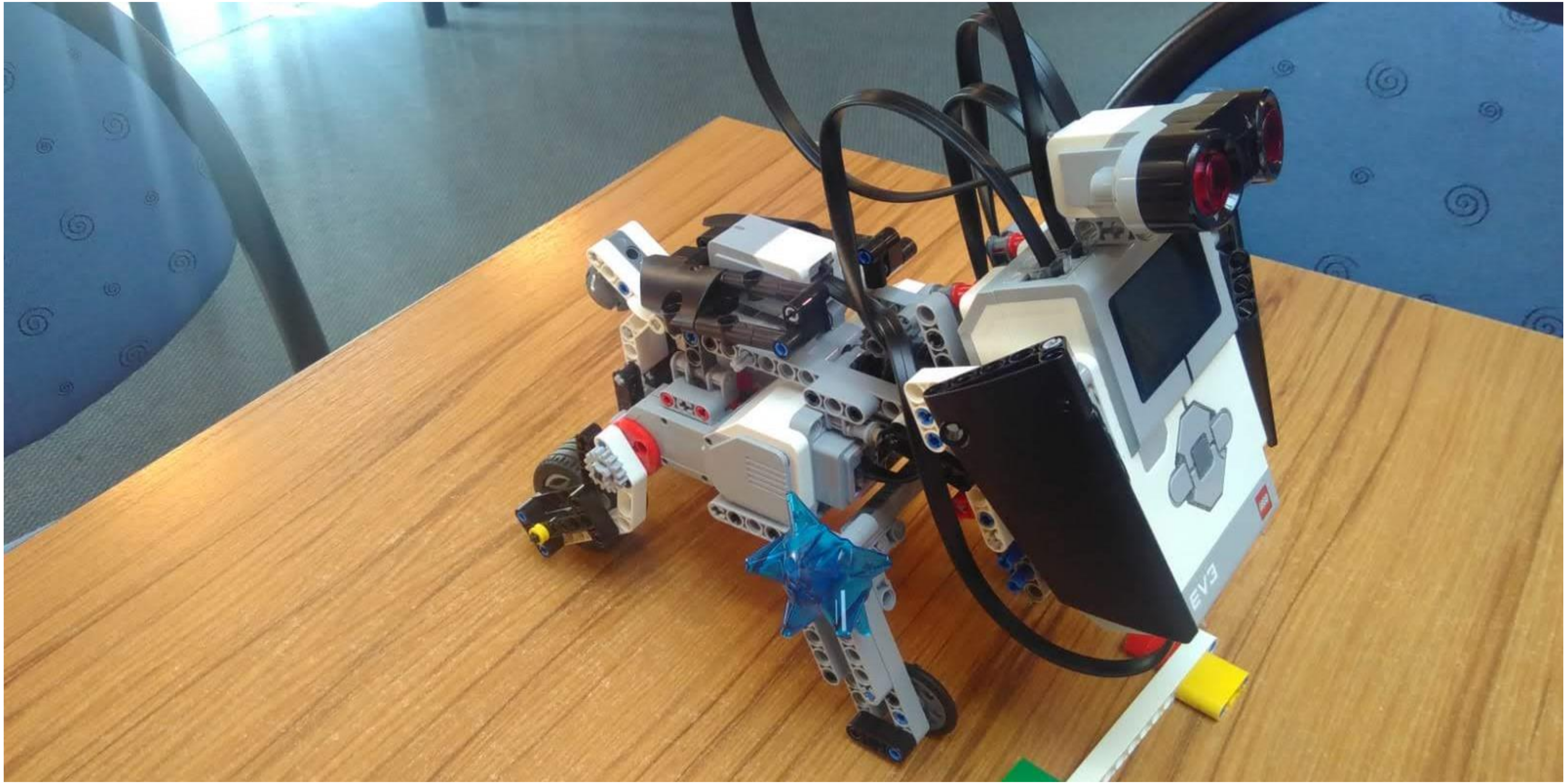
E/S/R
Science for Communities

E.r.i.c.a.

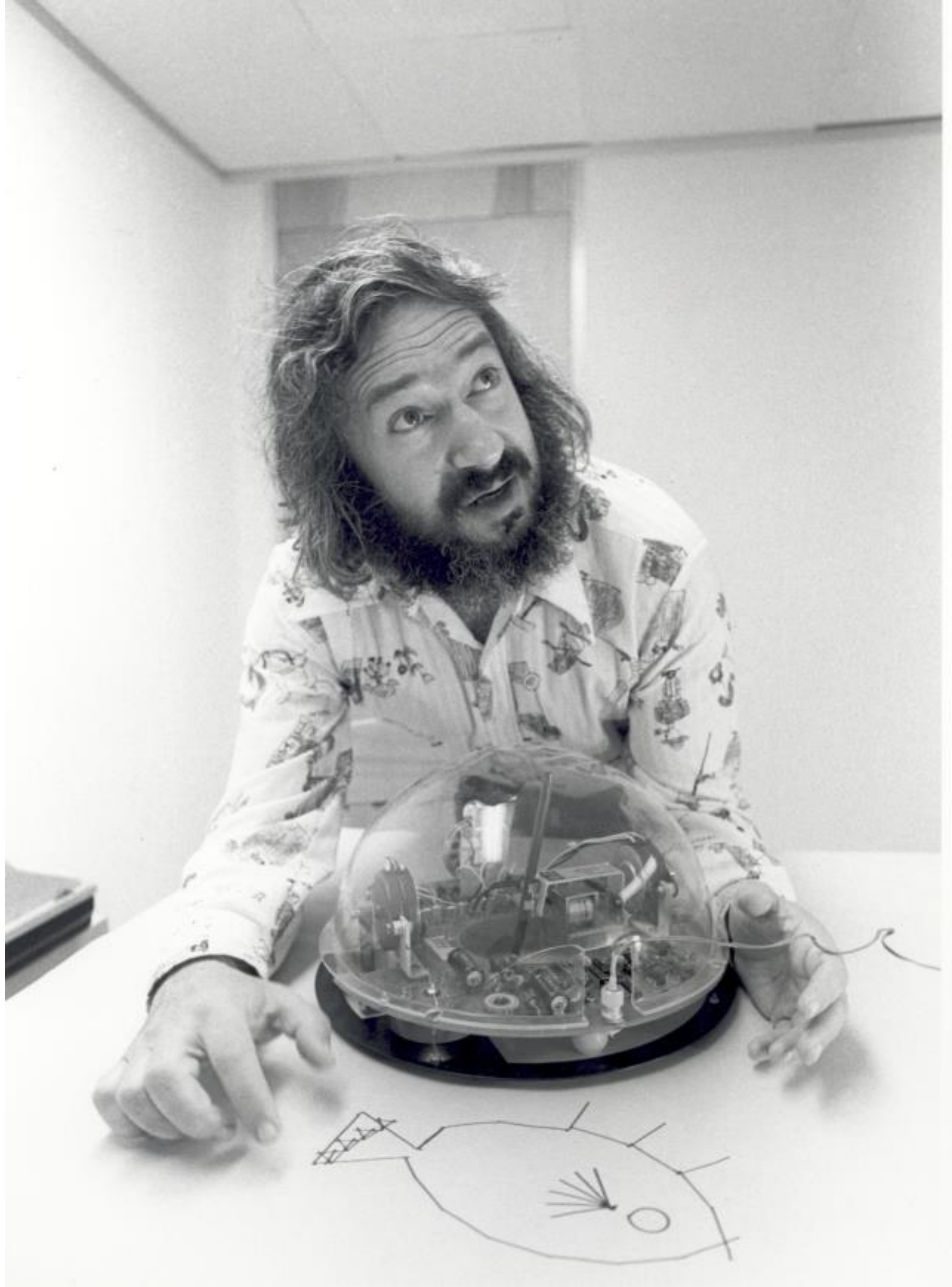


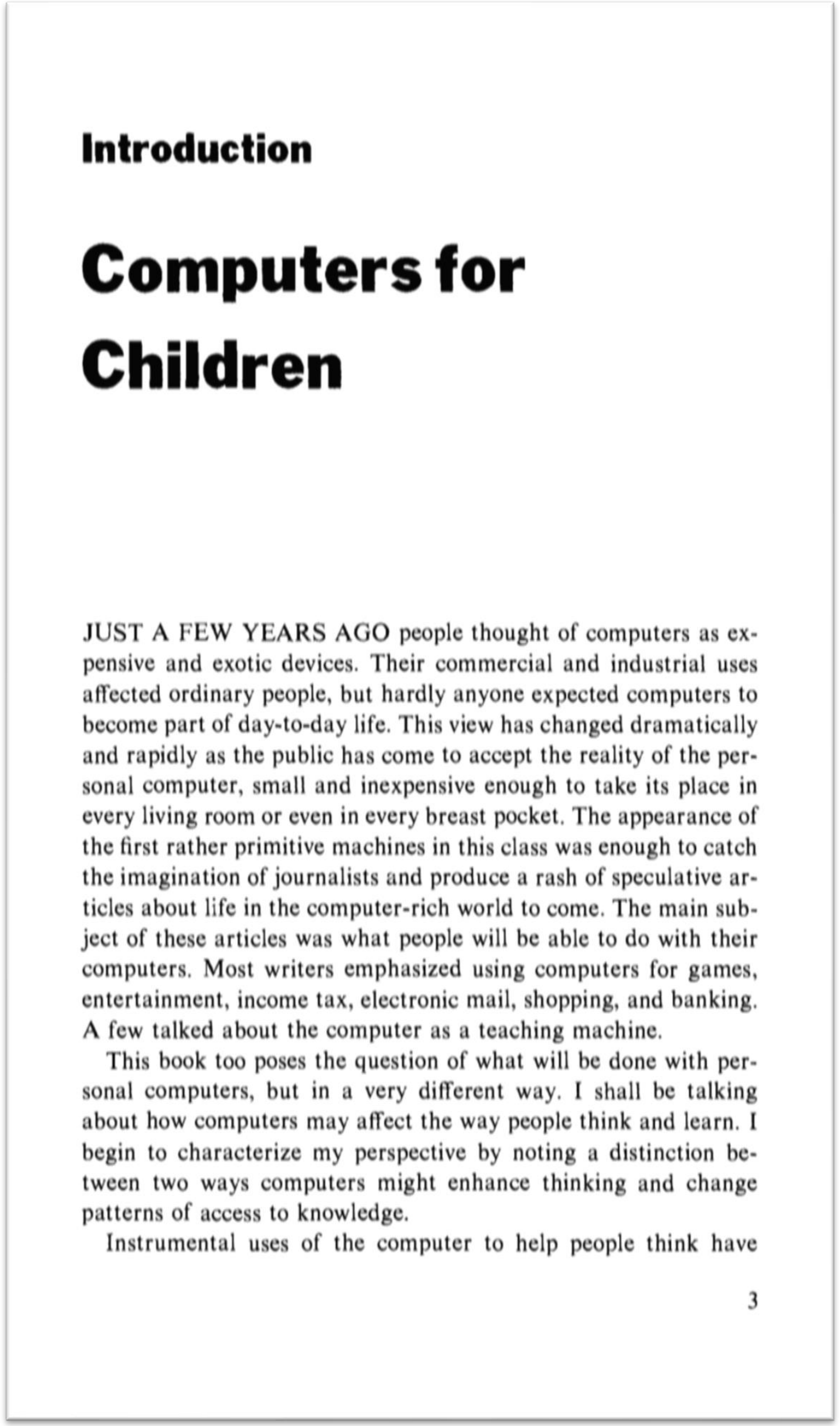


Science for Communities



E/S/R
Science for Communities





Introduction

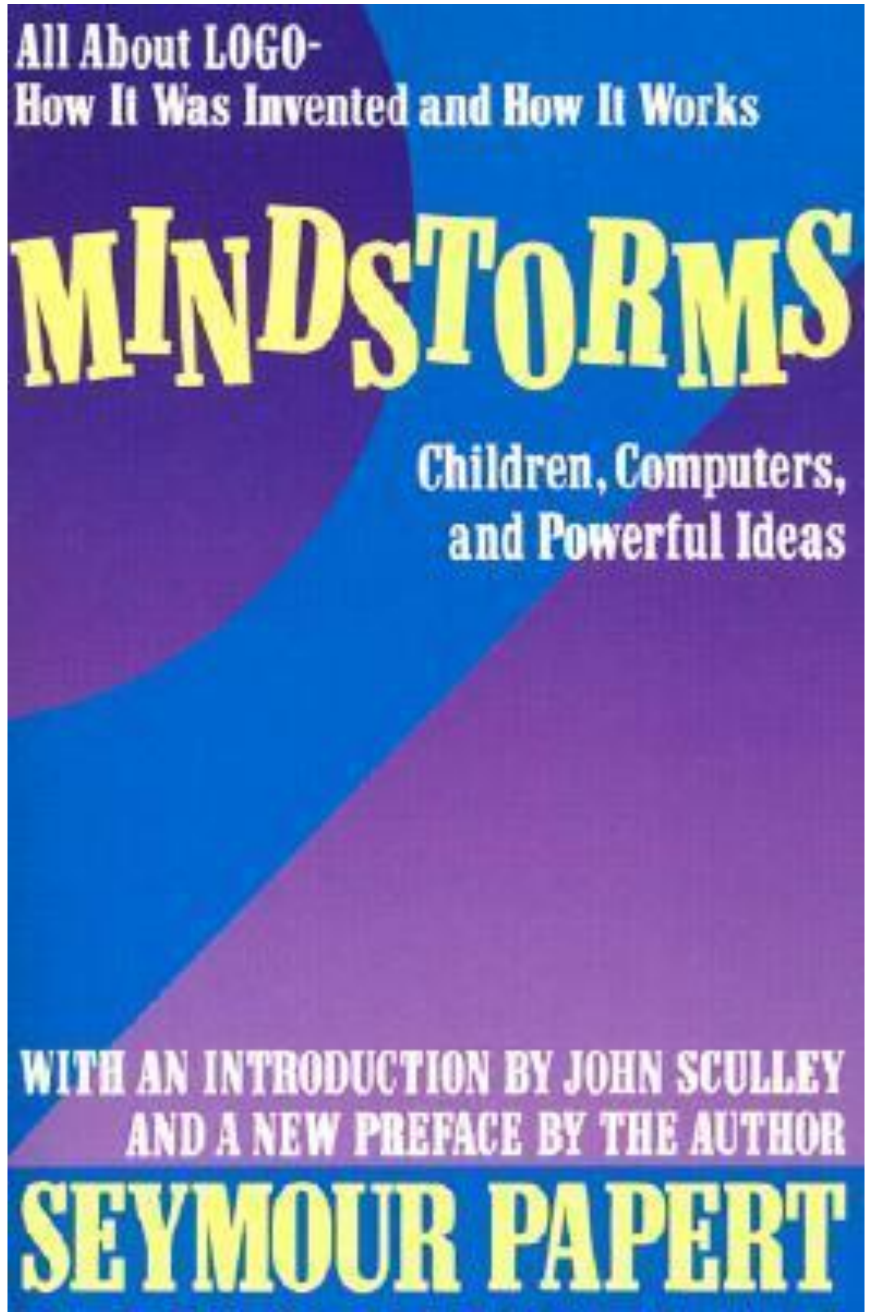
Computers for Children

JUST A FEW YEARS AGO people thought of computers as expensive and exotic devices. Their commercial and industrial uses affected ordinary people, but hardly anyone expected computers to become part of day-to-day life. This view has changed dramatically and rapidly as the public has come to accept the reality of the personal computer, small and inexpensive enough to take its place in every living room or even in every breast pocket. The appearance of the first rather primitive machines in this class was enough to catch the imagination of journalists and produce a rash of speculative articles about life in the computer-rich world to come. The main subject of these articles was what people will be able to do with their computers. Most writers emphasized using computers for games, entertainment, income tax, electronic mail, shopping, and banking. A few talked about the computer as a teaching machine.

This book too poses the question of what will be done with personal computers, but in a very different way. I shall be talking about how computers may affect the way people think and learn. I begin to characterize my perspective by noting a distinction between two ways computers might enhance thinking and change patterns of access to knowledge.

Instrumental uses of the computer to help people think have

3



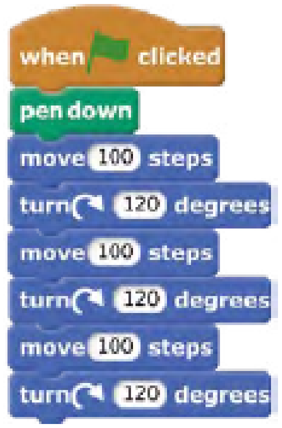
Science for Communities

https://mindstorms.media.mit.edu/

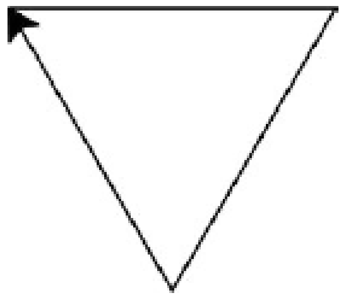
Simon Inns | https://youtu.be/XIdOR9n398c

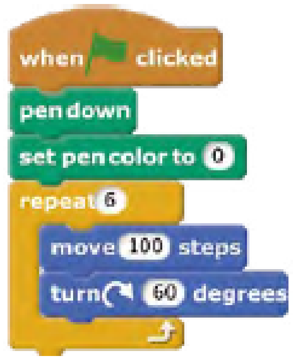
E/S/R
Science for Communities

Scratch	Python	Output

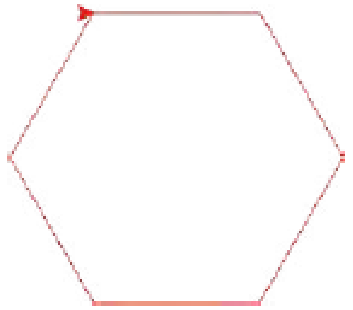


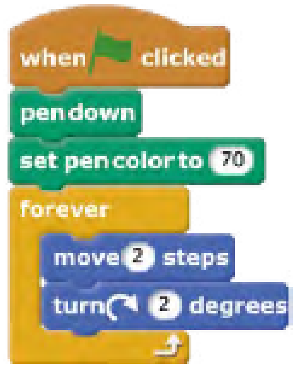

```
from turtle import *
forward(100)
right(120)
forward(100)
right(120)
forward(100)
```



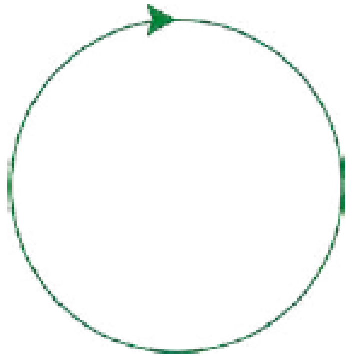


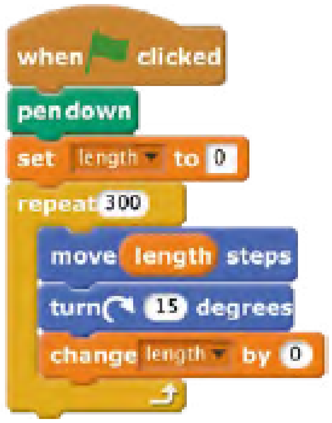

```
from turtle import *
color('red')
for i in range(6):
    forward(100)
    right(60)
```



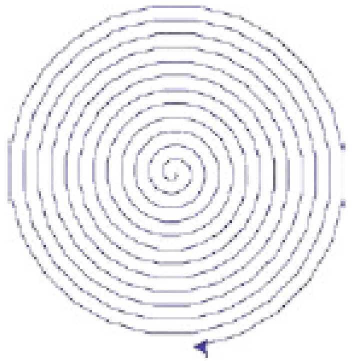



```
from turtle import *
color('green')
while True:
    forward(2)
    right(2)
```



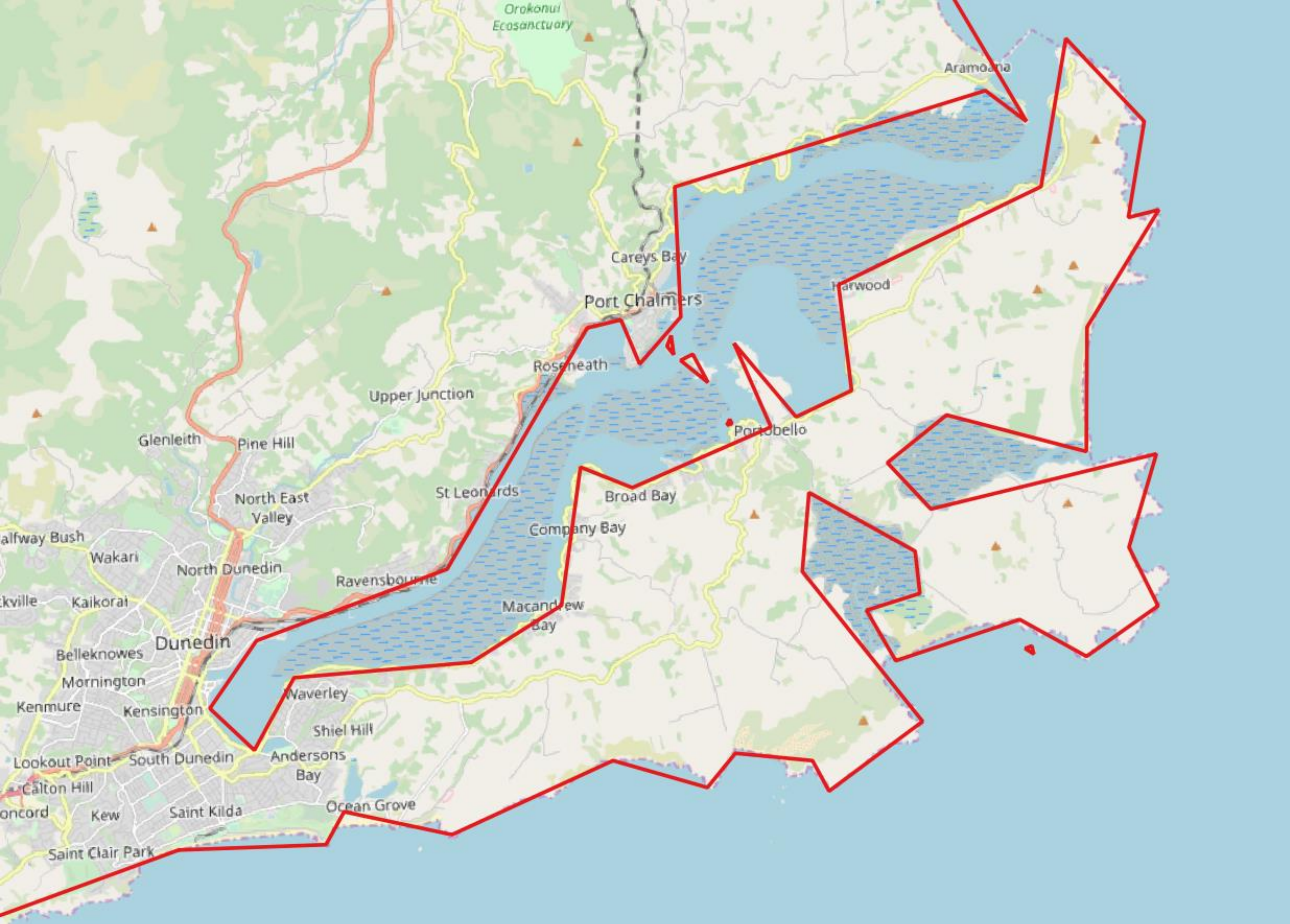



```
from turtle import *
color('blue')
length = 0
for i in range(300):
    forward(length)
    right(15)
    length = length + 0.1
```

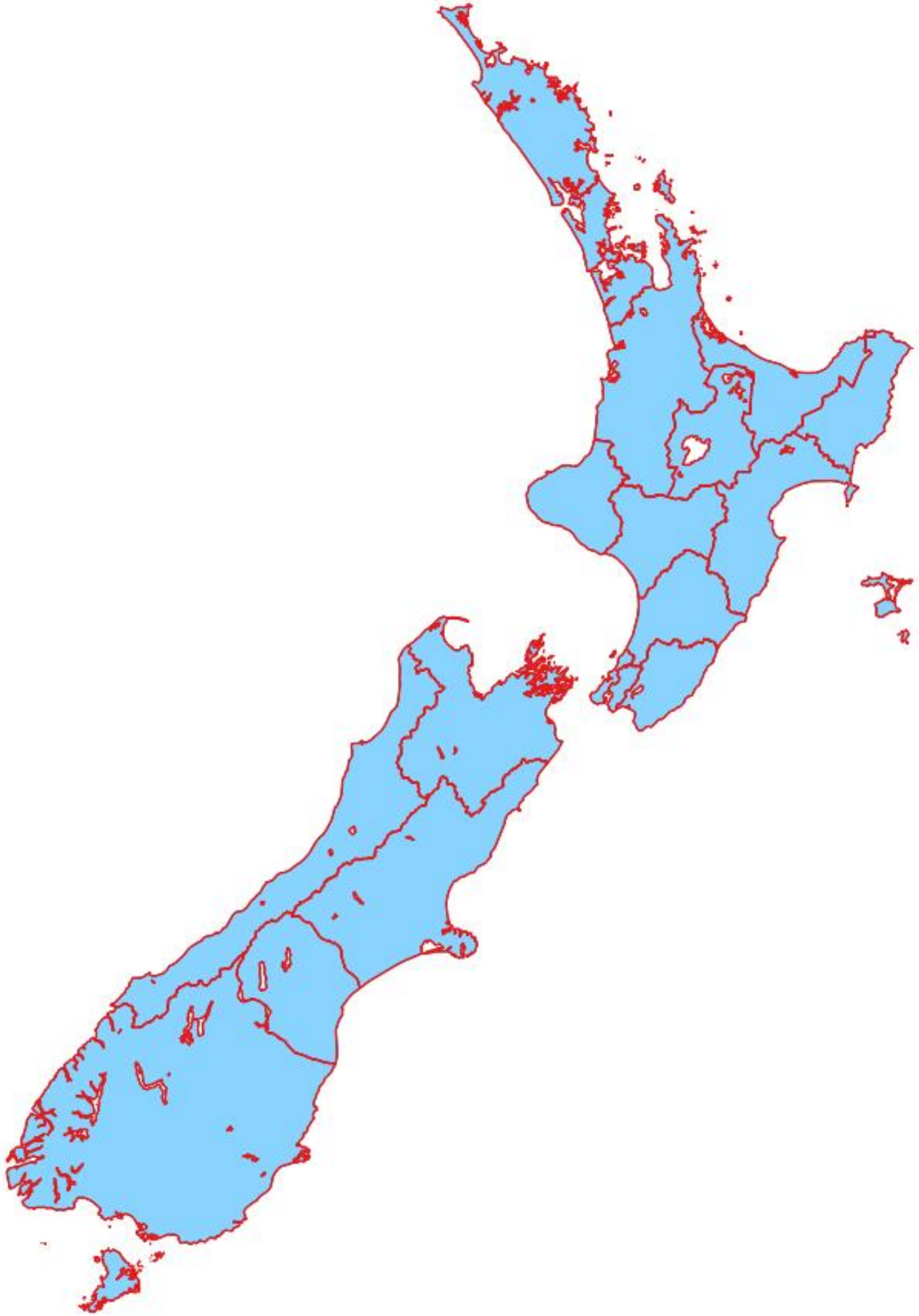


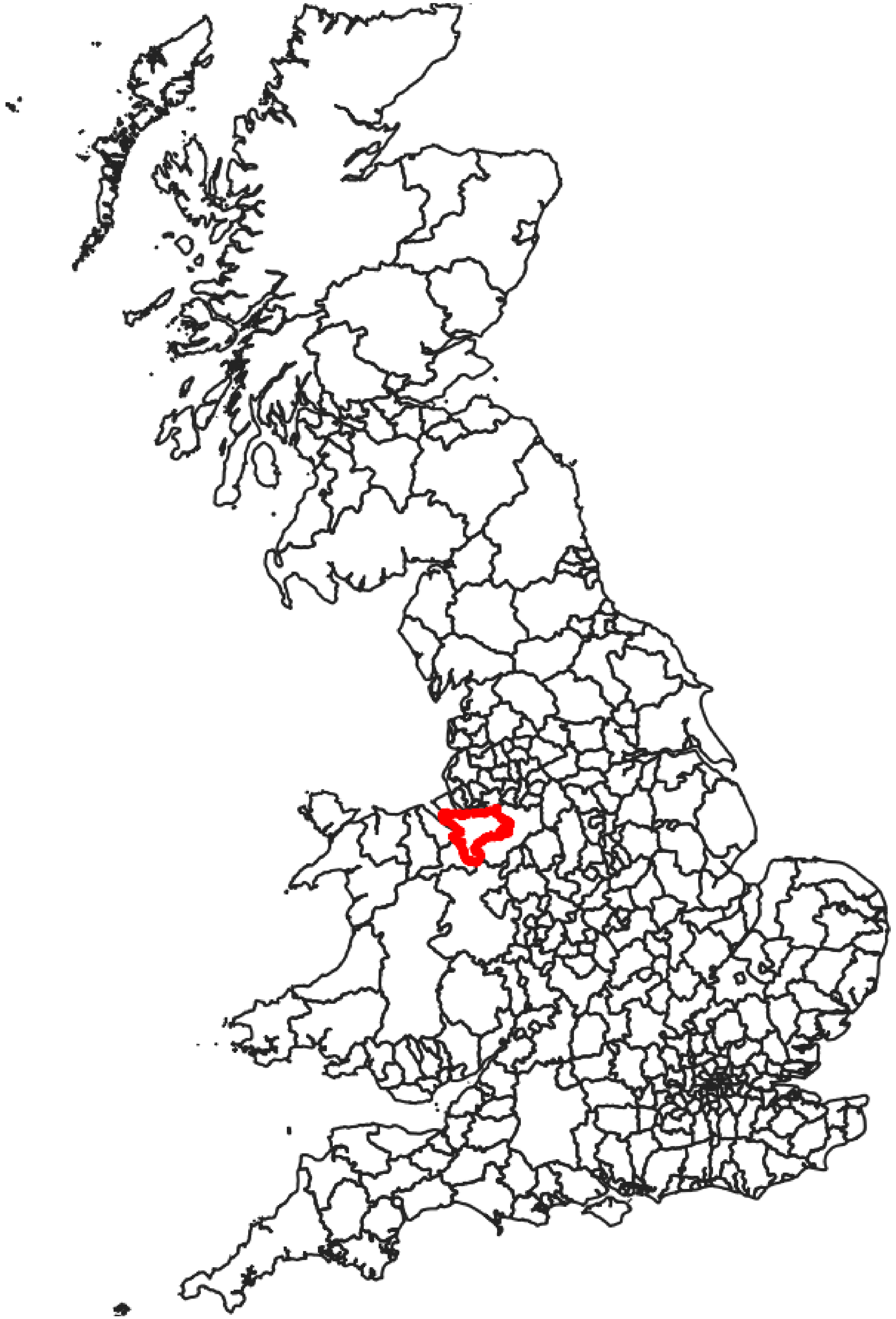
E/S/R
Science for Communities

https://helloworld.raspberrypi.org/issues/1/pdf



Science for Communities





Science for Communities

Public Health England

Protecting and improving the nation's health

Cheshire West and Chester

Unitary authority

This profile was published on 4th July 2017
Deprivation map (page 2) revised on 4th April 2018

Health Profile 2017

Health in summary
The health of people in Cheshire West and Chester is varied compared with the England average. About 16% (9,200) of children live in low income families. Life expectancy for both men and women is similar to the England average.

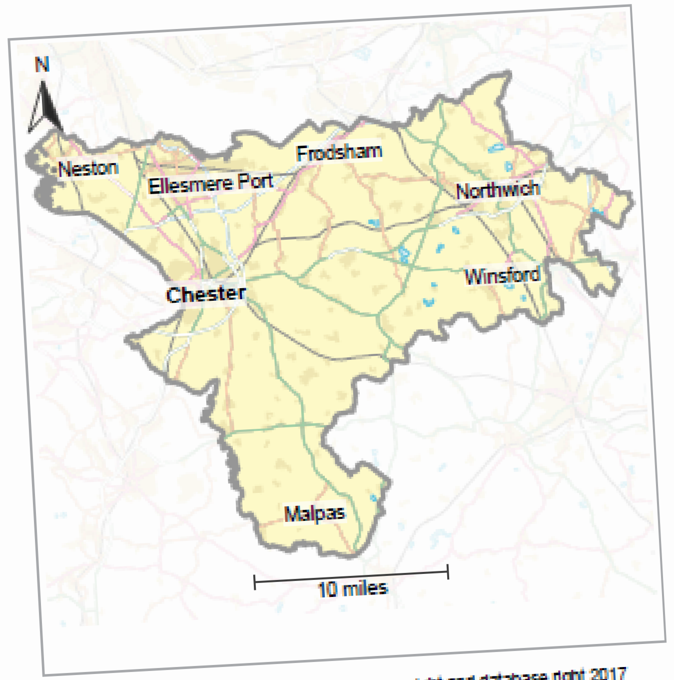
Health inequalities
Life expectancy is 10.0 years lower for men and 8.7 years lower for women in the most deprived areas of Cheshire West and Chester than in the least deprived areas.

Child health
In Year 6, 18.8% (623) of children are classified as obese. The rate of alcohol-specific hospital stays among those under 18 is 37*. This represents 24 stays per year. Levels of breastfeeding initiation are worse than the England average. Levels of GCSE attainment are better than the England average.

Adult health
The rate of alcohol-related harm hospital stays is 606*, better than the average for England. This represents 2,038 stays per year. The rate of self-harm hospital stays is 200*. This represents 653 stays per year. The rate of smoking related deaths is 270*, better than the average for England. This represents 552 deaths per year. Estimated levels of adult smoking are better than the England average. The rate of people killed and seriously injured on roads is worse than average. Rates of sexually transmitted infections and TB are better than average. Rates of statutory homelessness, violent crime, long term unemployment and early deaths from cardiovascular diseases are better than average.

Local priorities
Priorities in Cheshire West and Chester include reducing inequalities, improving mental health and wellbeing and addressing key lifestyle issues (reducing smoking and substance misuse, and improving healthy eating and physical activity). For more information see www.valeroyalccg.nhs.uk, www.westcheshireccg.nhs.uk and www.cheshirewestandchester.gov.uk/jsna

* rate per 100,000 population

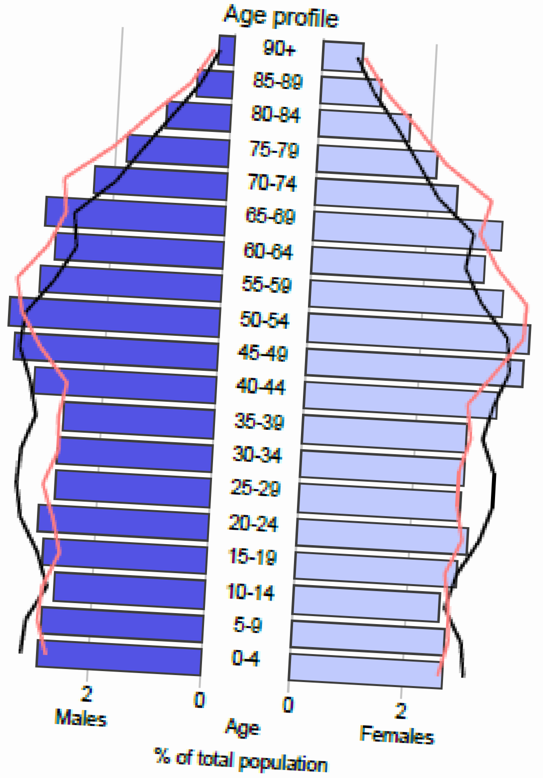


This profile gives a picture of people's health in Cheshire West and Chester. It is designed to help local government and health services understand their community's needs, so that they can work together to improve people's health and reduce health inequalities.

Visit www.healthprofiles.info for more profiles, more information and interactive maps and tools.

Follow @PHE_uk on Twitter

Population: summary characteristics



Cheshire West and Chester (population in thousands)	Males	Females	Persons
Population (2015):	163	171	334
Projected population (2020):	164	173	337
% people from an ethnic minority group:	1.8%	1.8%	1.8%
Dependency ratio (dependants / working population) x 100			65.0%
England (population in thousands)			
Population (2015):	27,029	27,757	54,786
Projected population (2020):	28,157	28,706	56,862
% people from an ethnic minority group:	13.1%	13.4%	13.2%
Dependency ratio (dependants / working population) x 100			60.7%

The age profile and table present demographic information for the residents of the area and England. They include a 2014-based population projection (to 2020), the percentage of people from an ethnic minority group (Annual Population Survey, October 2014 to September 2015) and the dependency ratio.

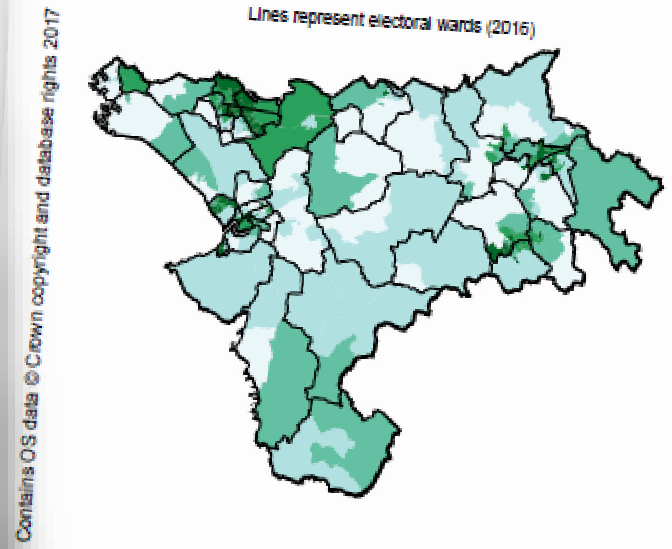
The dependency ratio estimates the number of dependants in an area by comparing the number of people considered less likely to be working (children aged under 16 and those of state pension age or above) with the working age population. A high ratio suggests the area might want to commission a greater level of services for older or younger people than those areas with a low ratio.

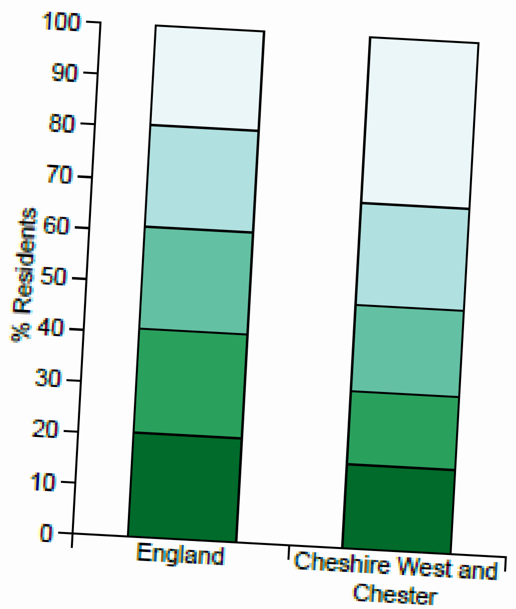
- Cheshire West and Chester 2015 (Male)
- Cheshire West and Chester 2015 (Female)
- England 2015
- Cheshire West and Chester 2020 estimate

Deprivation: a national view

The map shows differences in deprivation in this area based on national comparisons, using national quintiles (fifths) of the Index of Multiple Deprivation 2015 (IMD 2015), shown by lower super output area. The darkest coloured areas are some of the most deprived neighbourhoods in England.

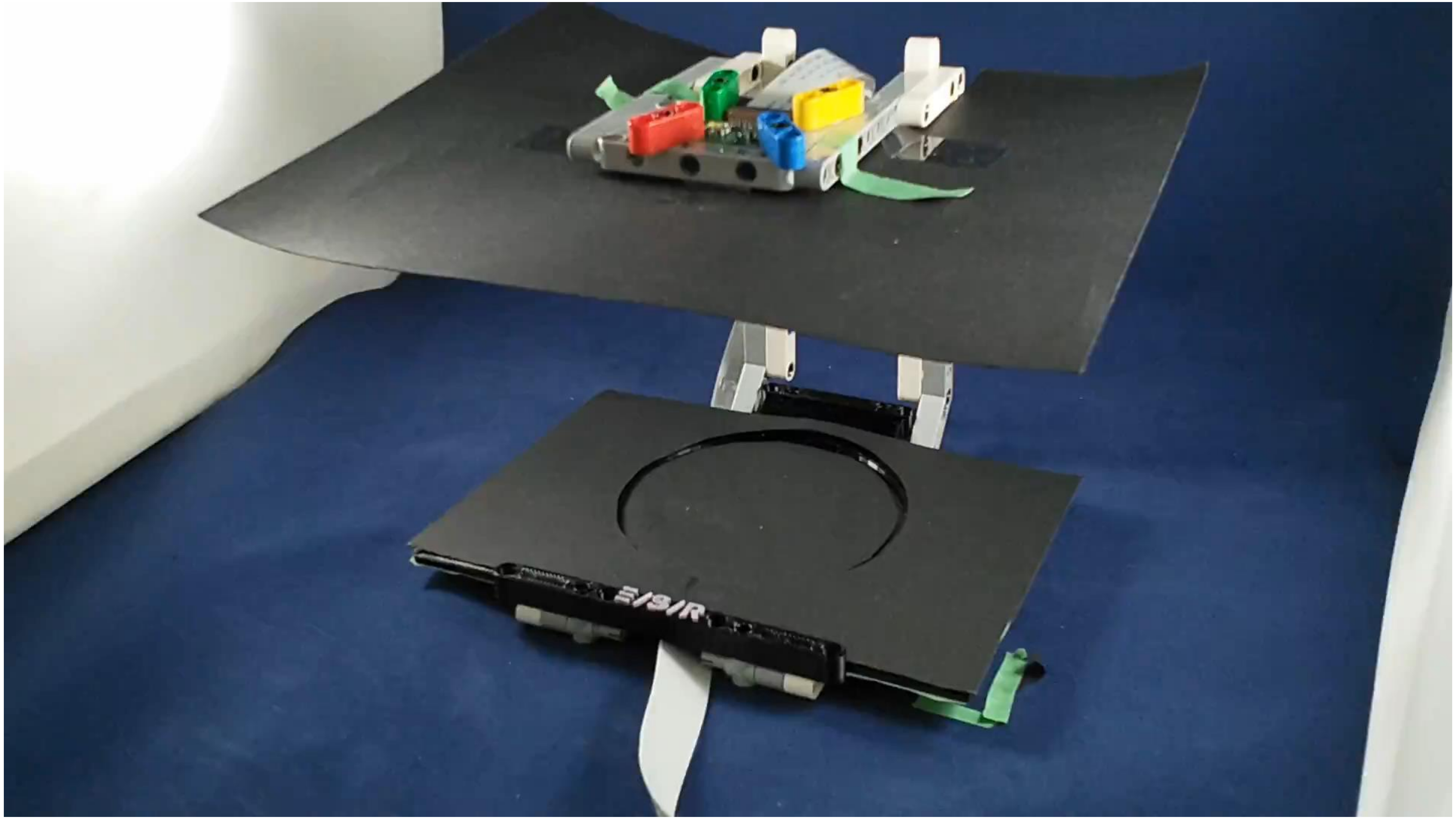
This chart shows the percentage of the population who live in areas at each level of deprivation.





Most deprived quintile — Least deprived quintile

Science for Communities


E/S/R

Recap

2019

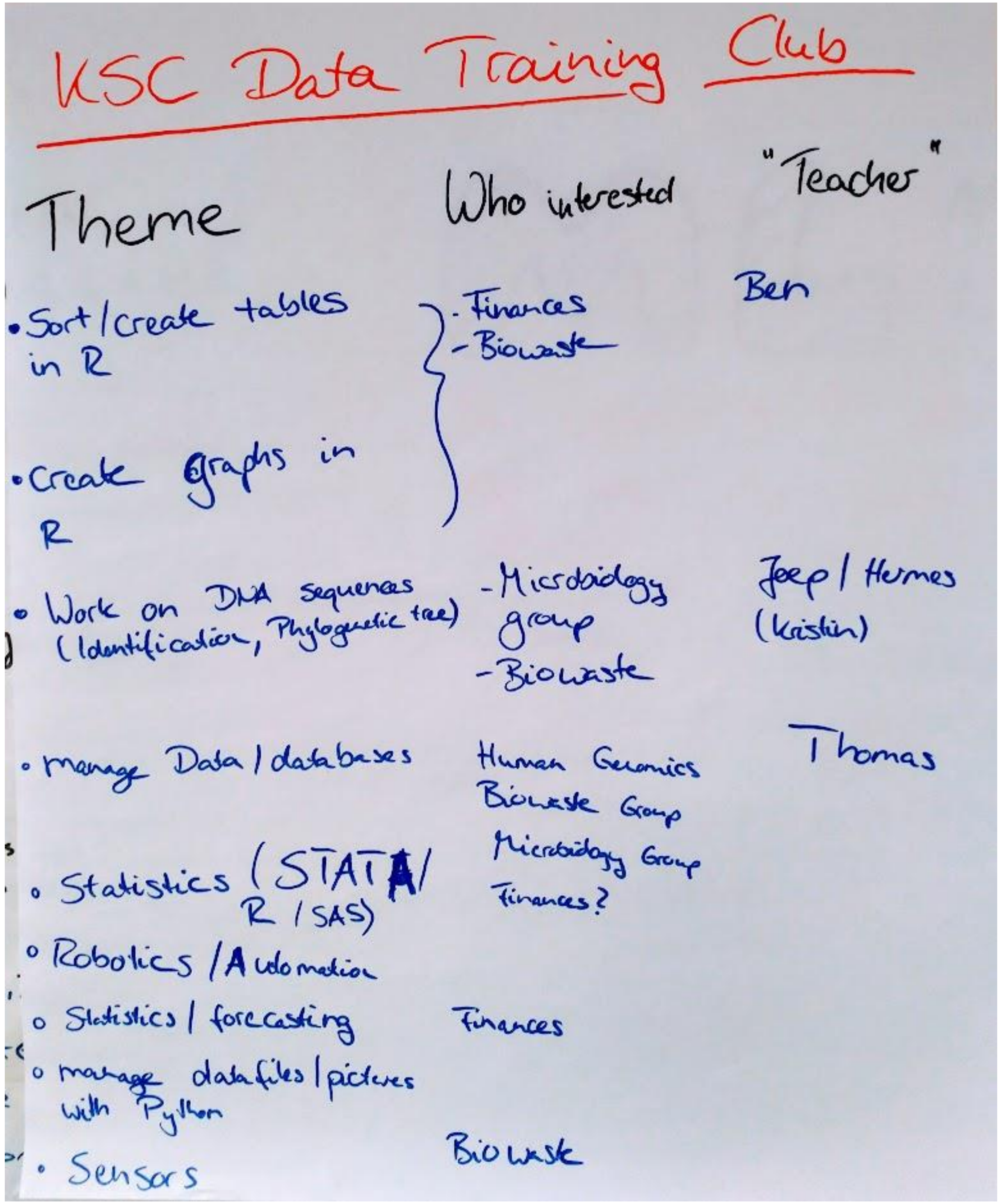
Science for Communities

Coming up in

2020

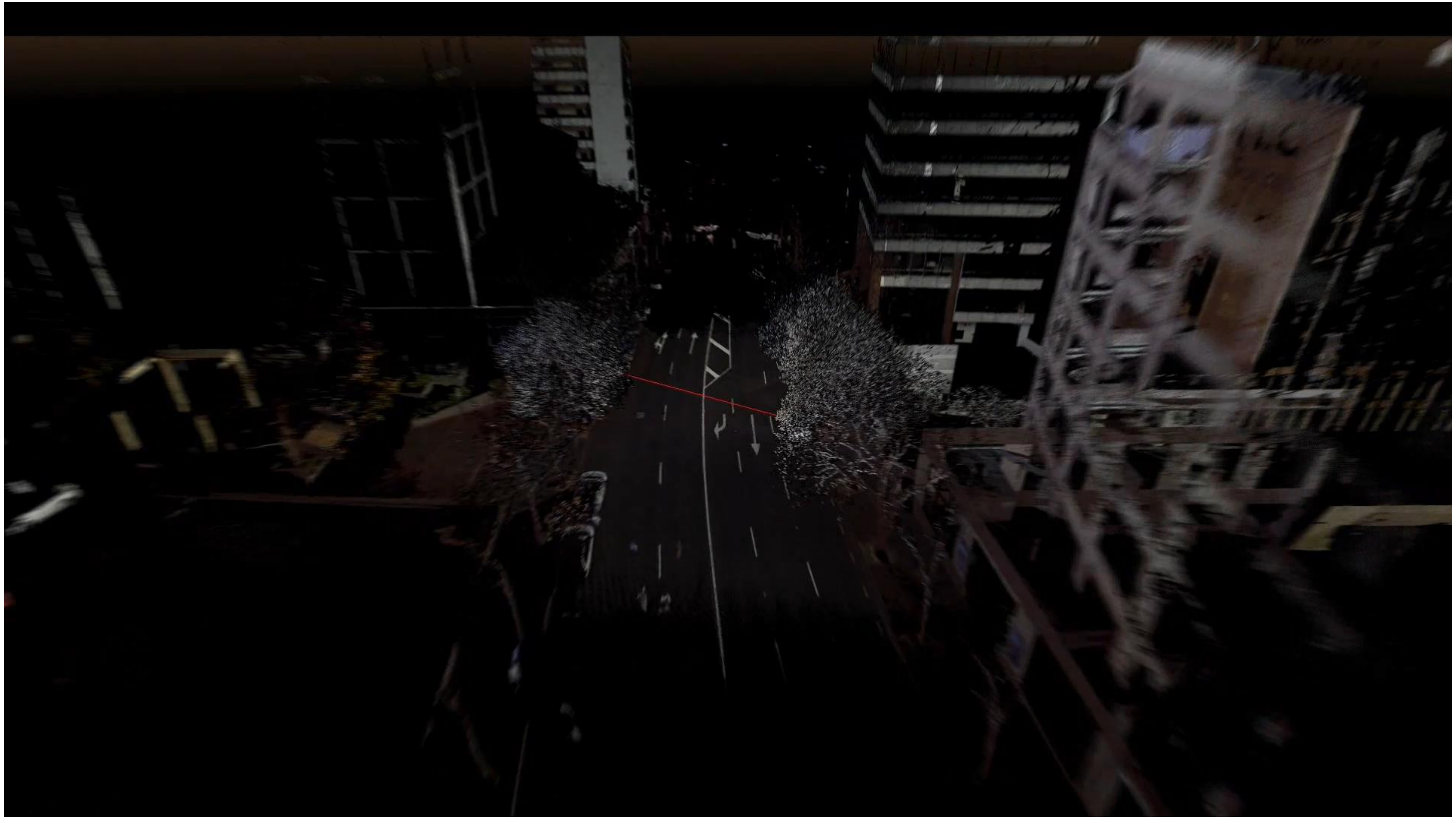
ESR
Science for Communities

Science for Communities

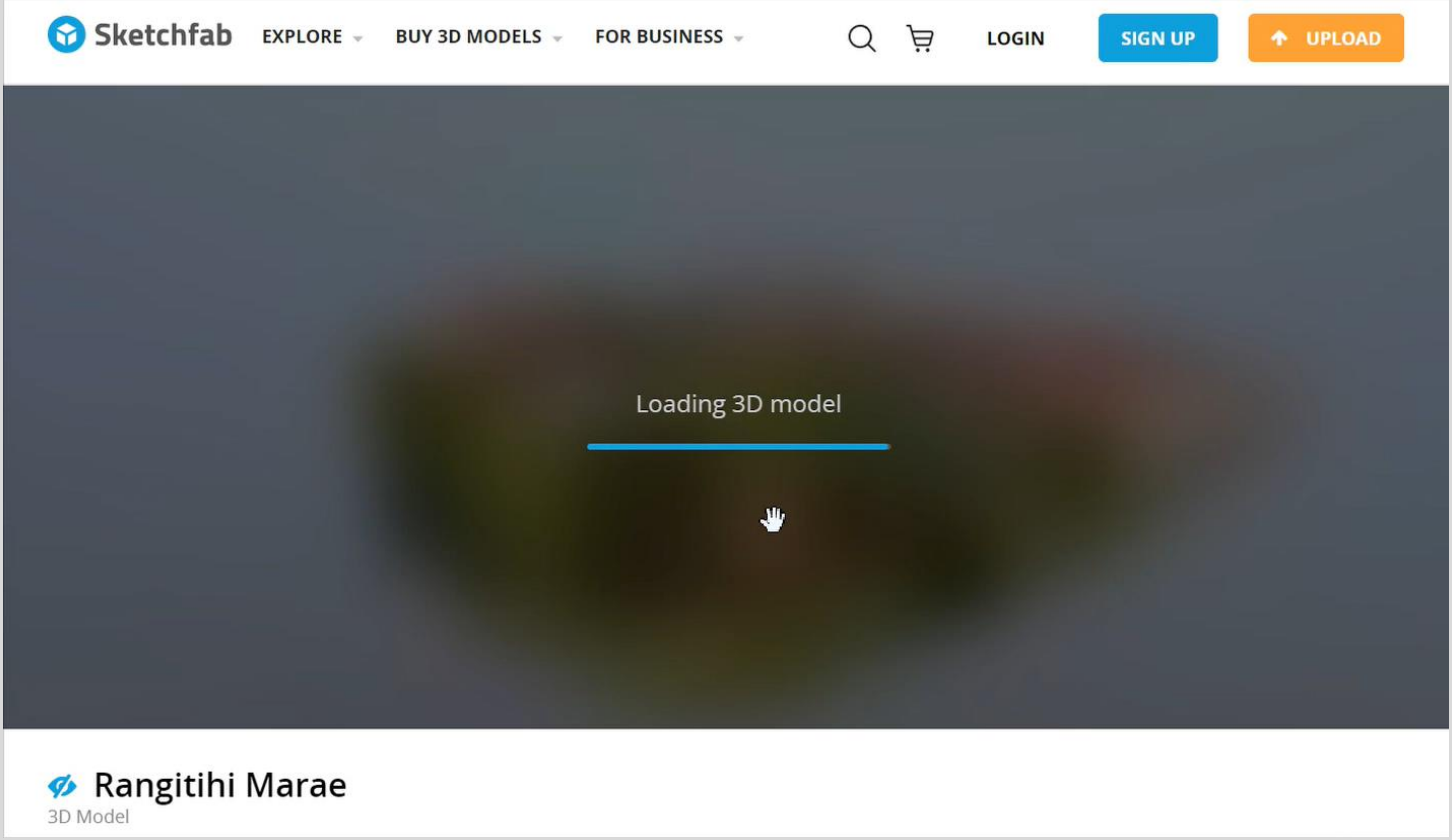


E/S/R
Science for Communities

3D



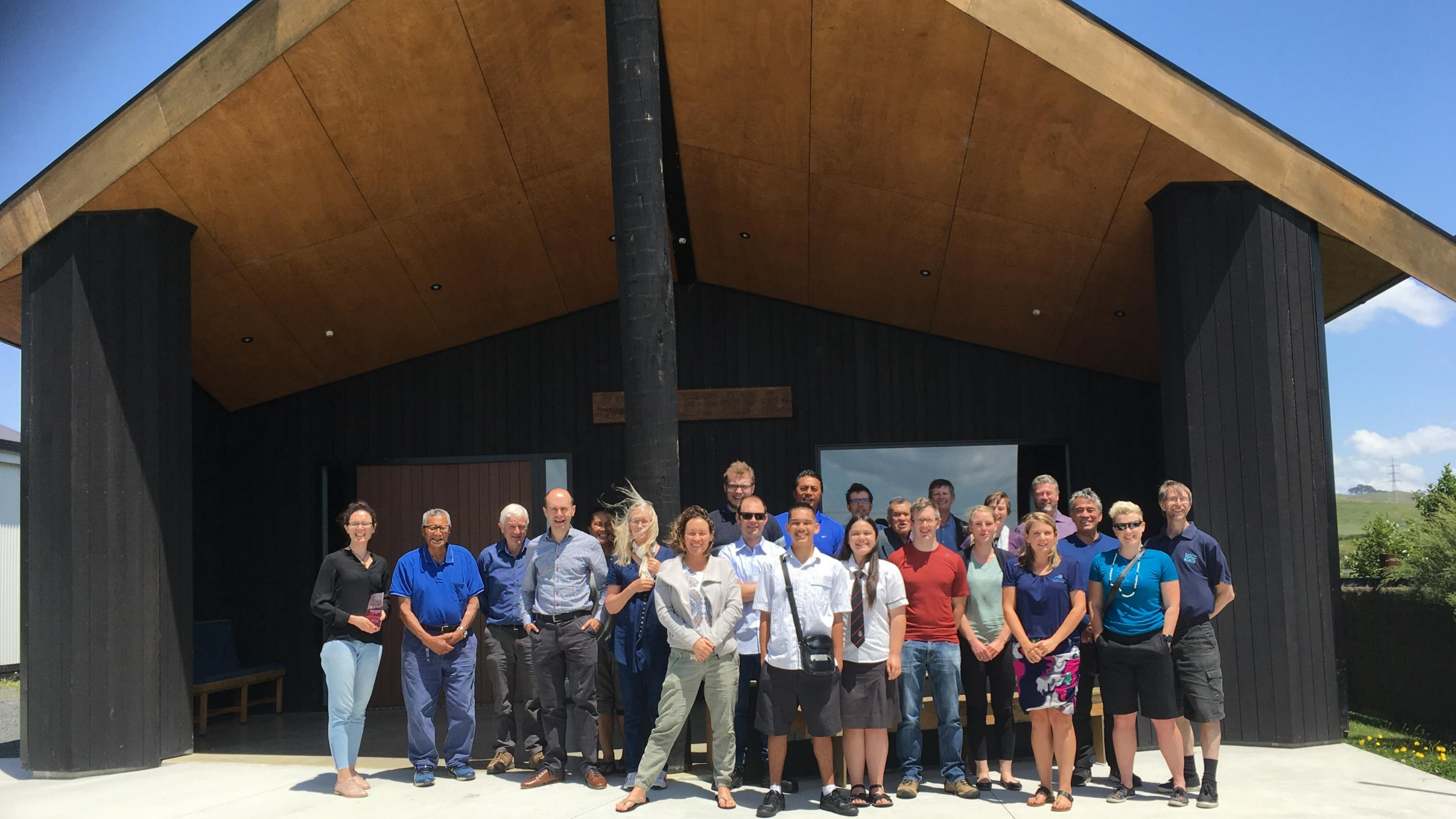
ESR
Science for Communities


Sketchfab
EXPLORE
BUY 3D MODELS
FOR BUSINESS
LOGIN
SIGN UP
UPLOAD
Loading 3D model
Rangitihi Marae
3D Model


finally:
Shell

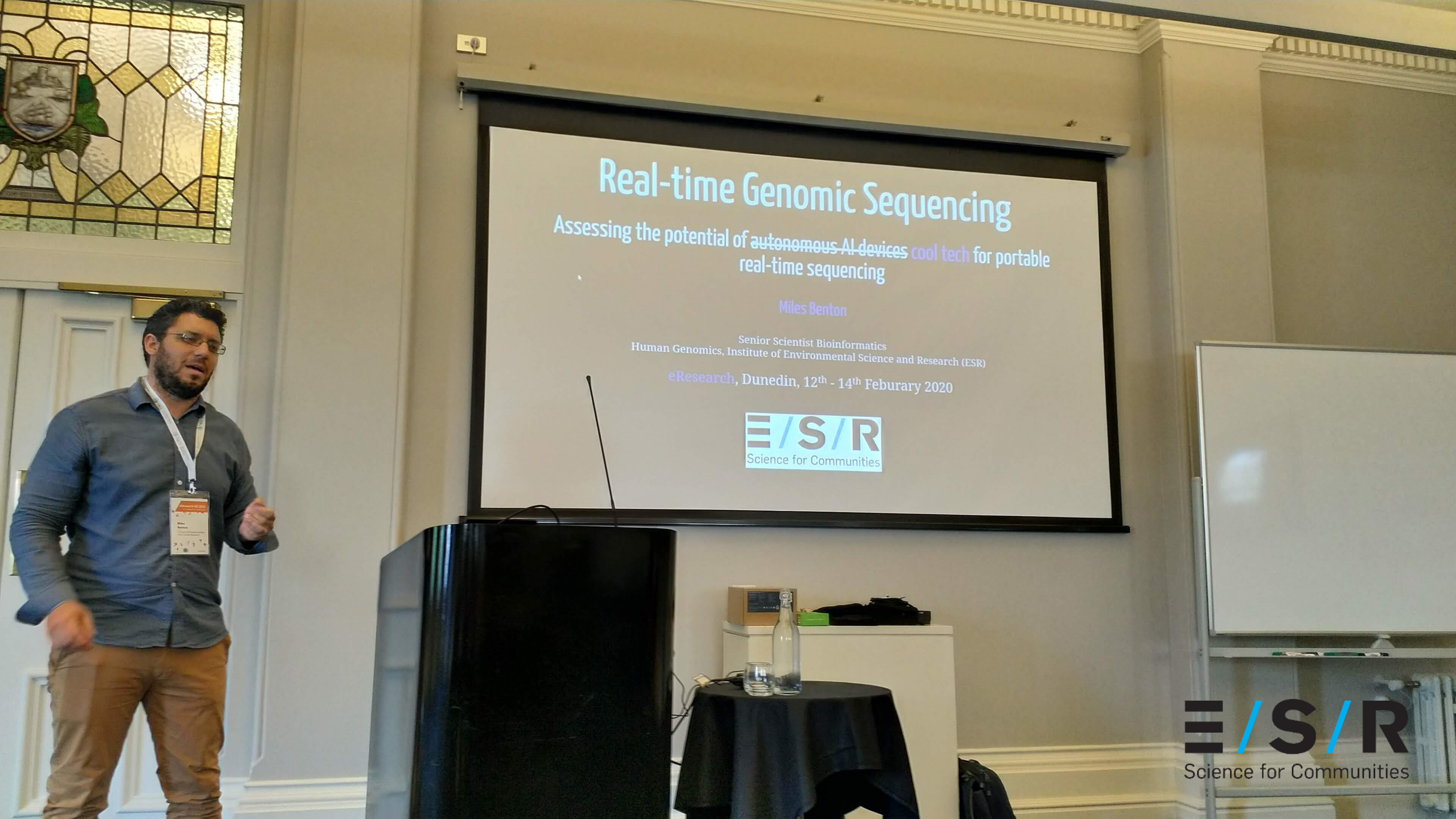
E/S/R
Science for Communities

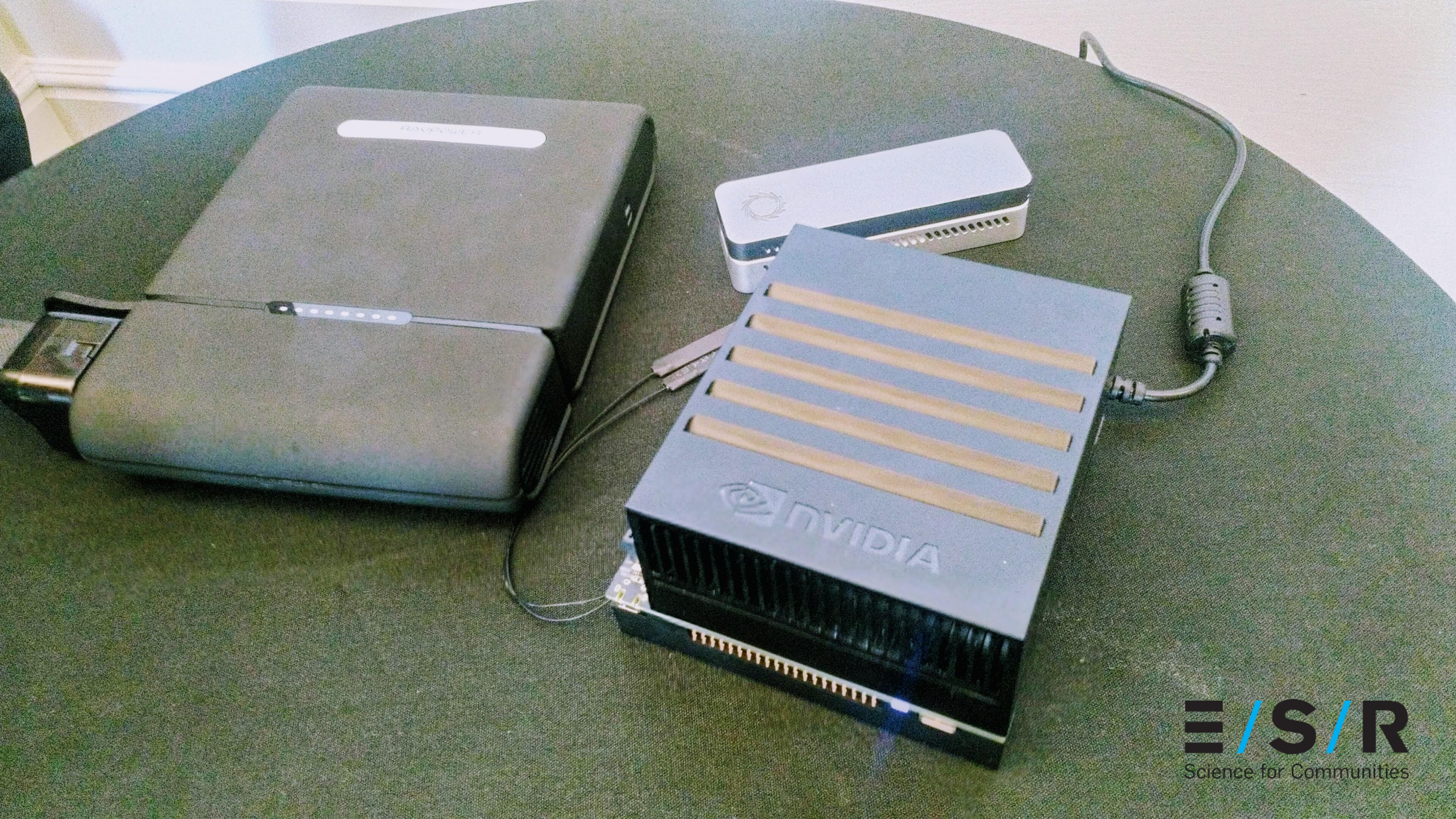

Our Land, Our Future
Tō tātou whenua, mō āpōpō
Tūī in your backyard!
E/S/R
Science for Communities


PUMA

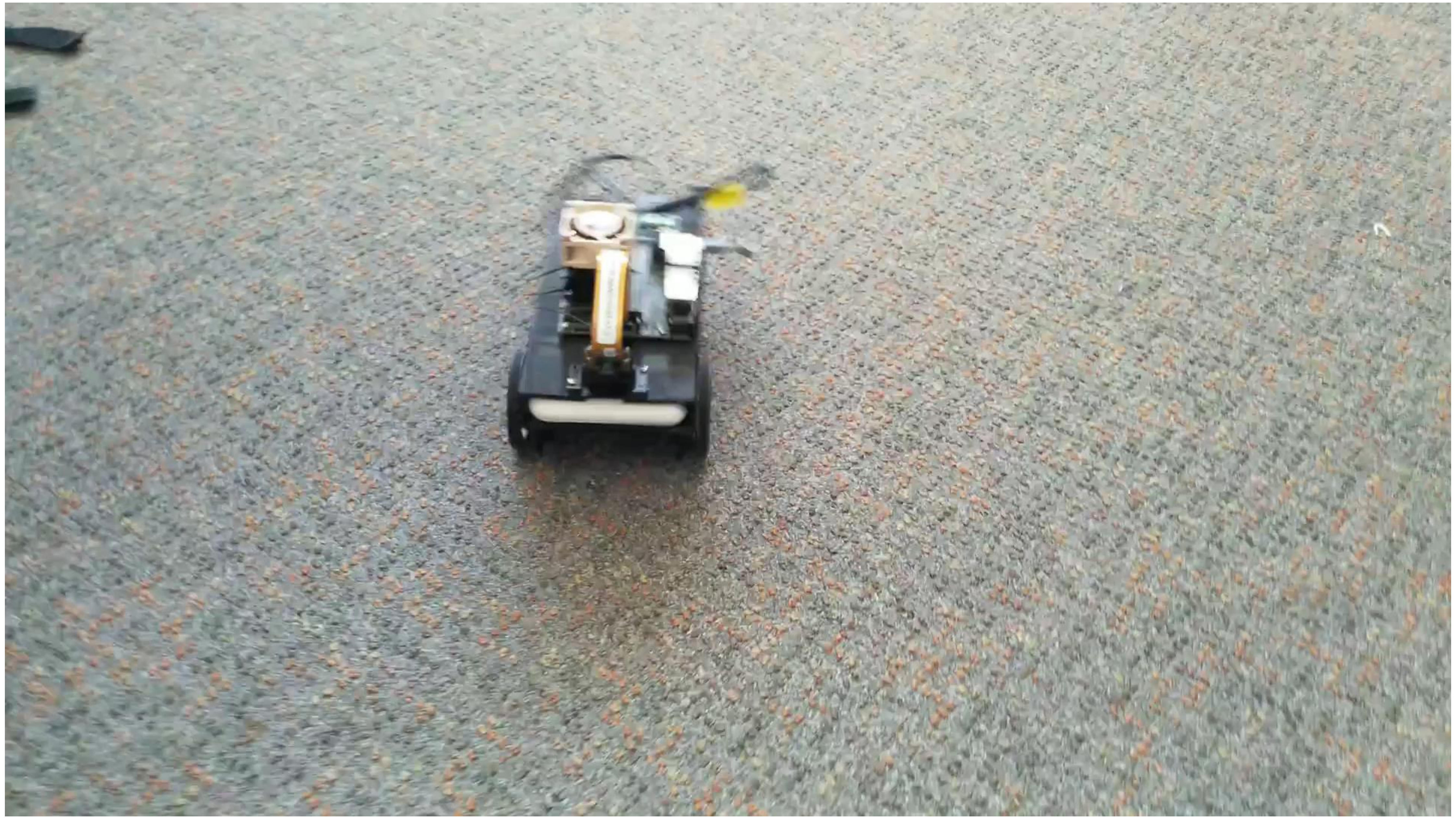
AI computing at the edge

E/S/R
Science for Communities


Real-time Genomic Sequencing
Assessing the potential of ~~autonomous AI devices~~ cool tech for portable real-time sequencing
Miles Benton
Senior Scientist Bioinformatics
Human Genomics, Institute of Environmental Science and Research (ESR)
eResearch, Dunedin, 12th - 14th Feburary 2020
E/S/R
Science for Communities
E/S/R
Science for Communities

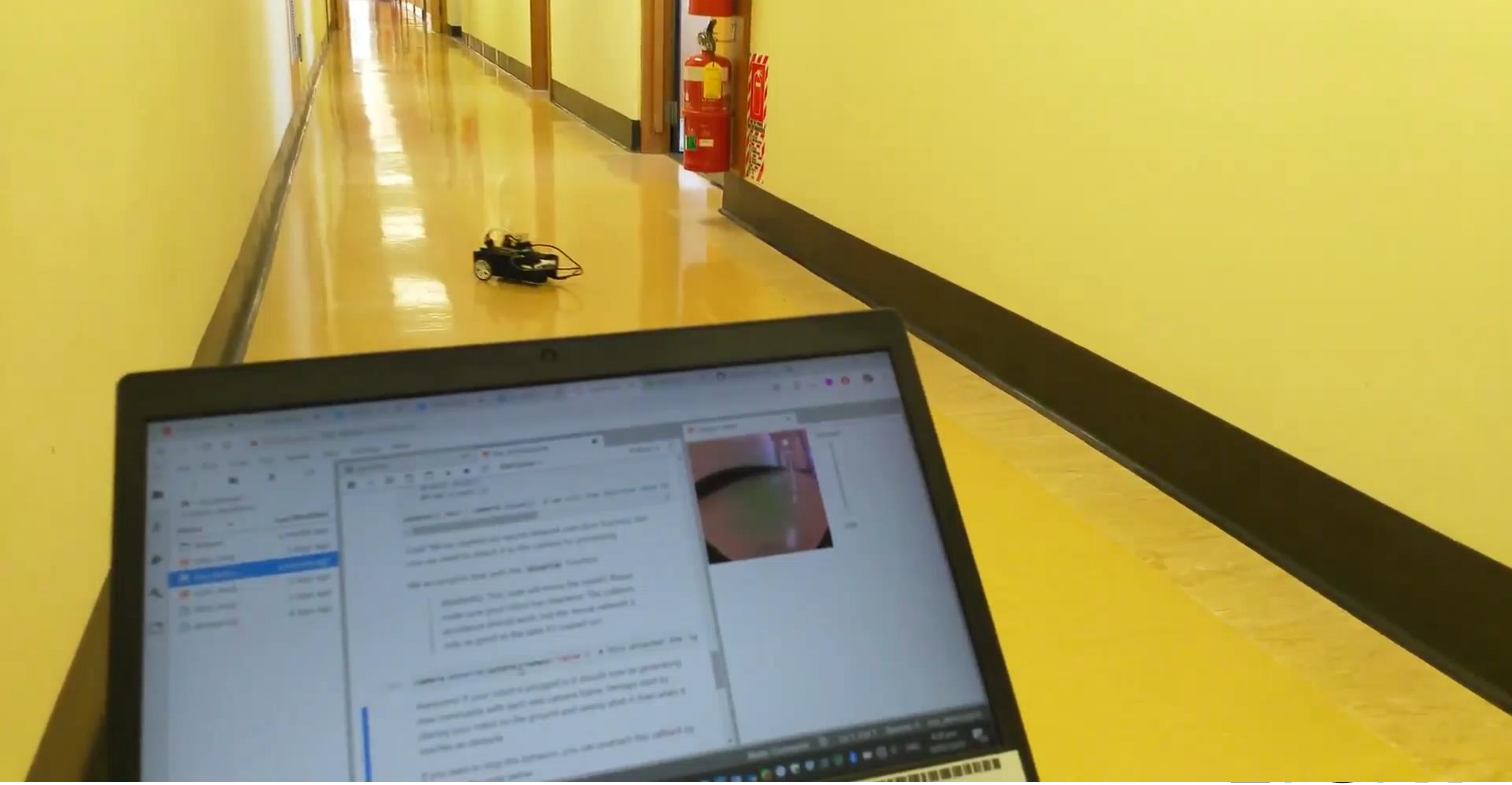

NVIDIA
E/S/R
Science for Communities

AI computing at the edge

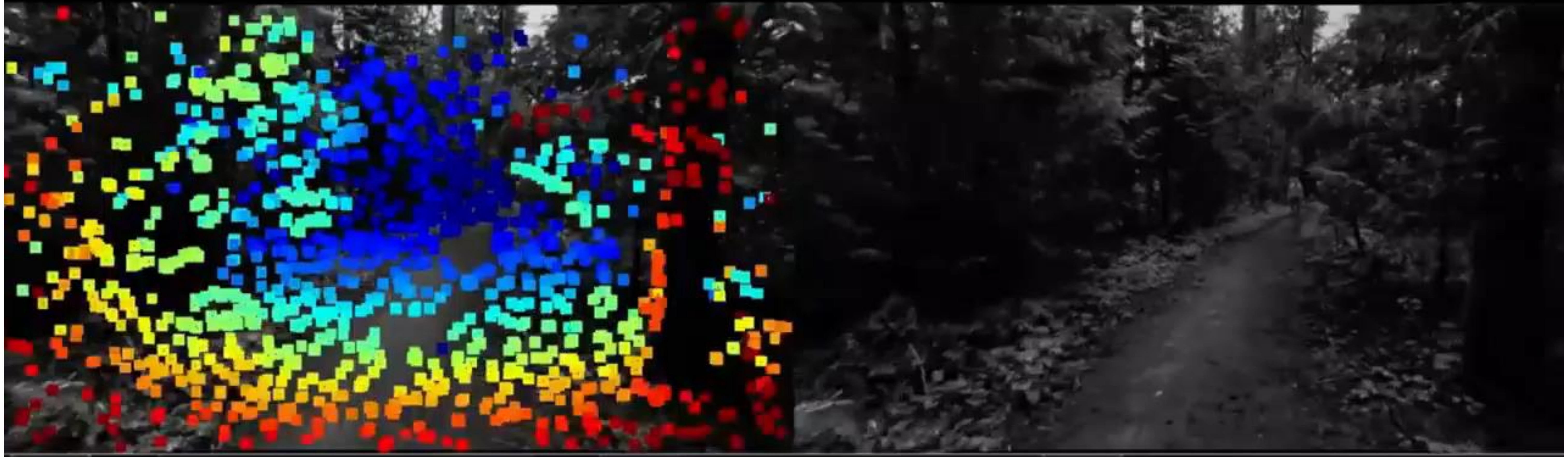


Science for Communities

AI computing at the edge



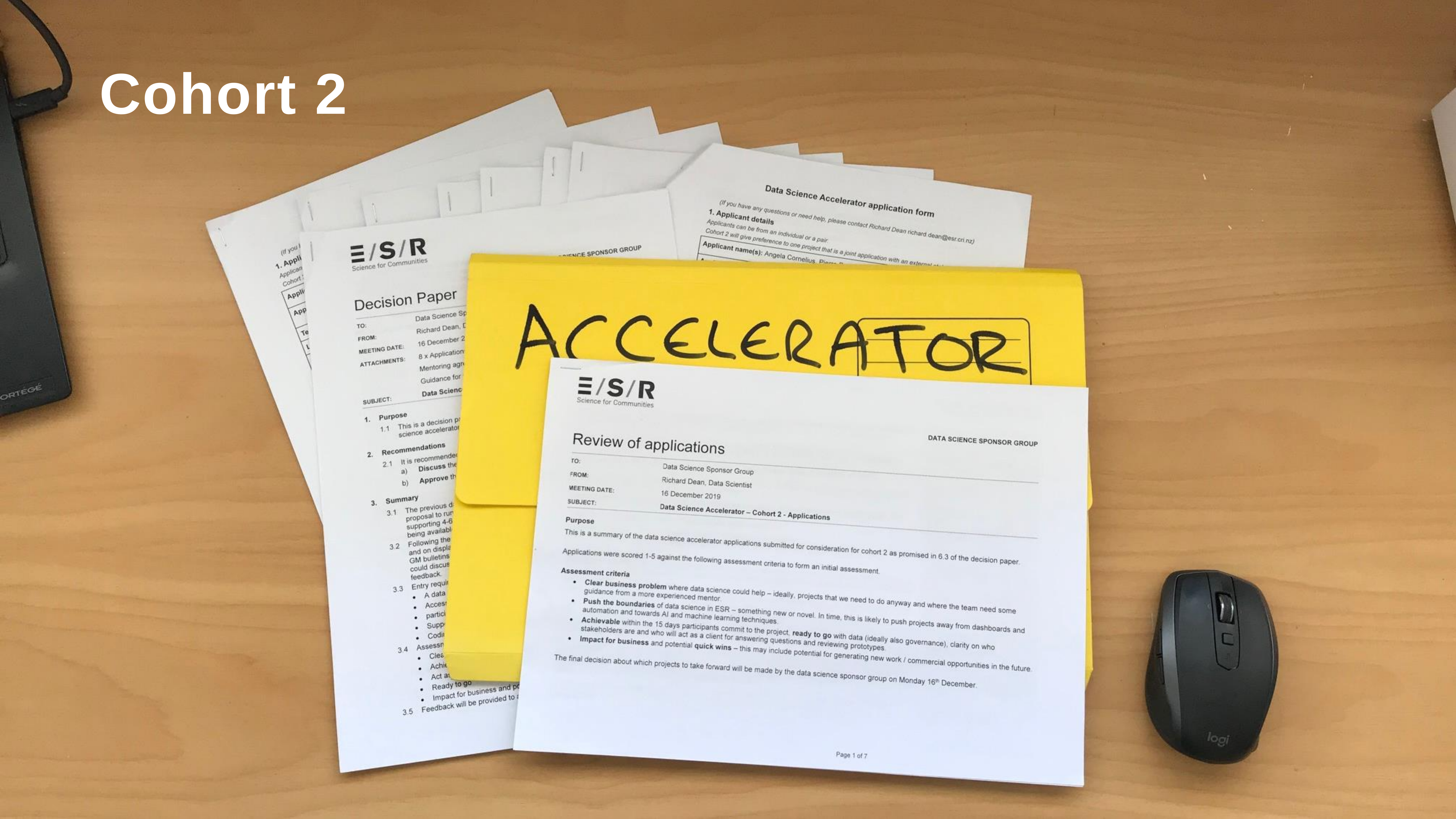
Semi-dense SLAM maps for obstacle avoidance

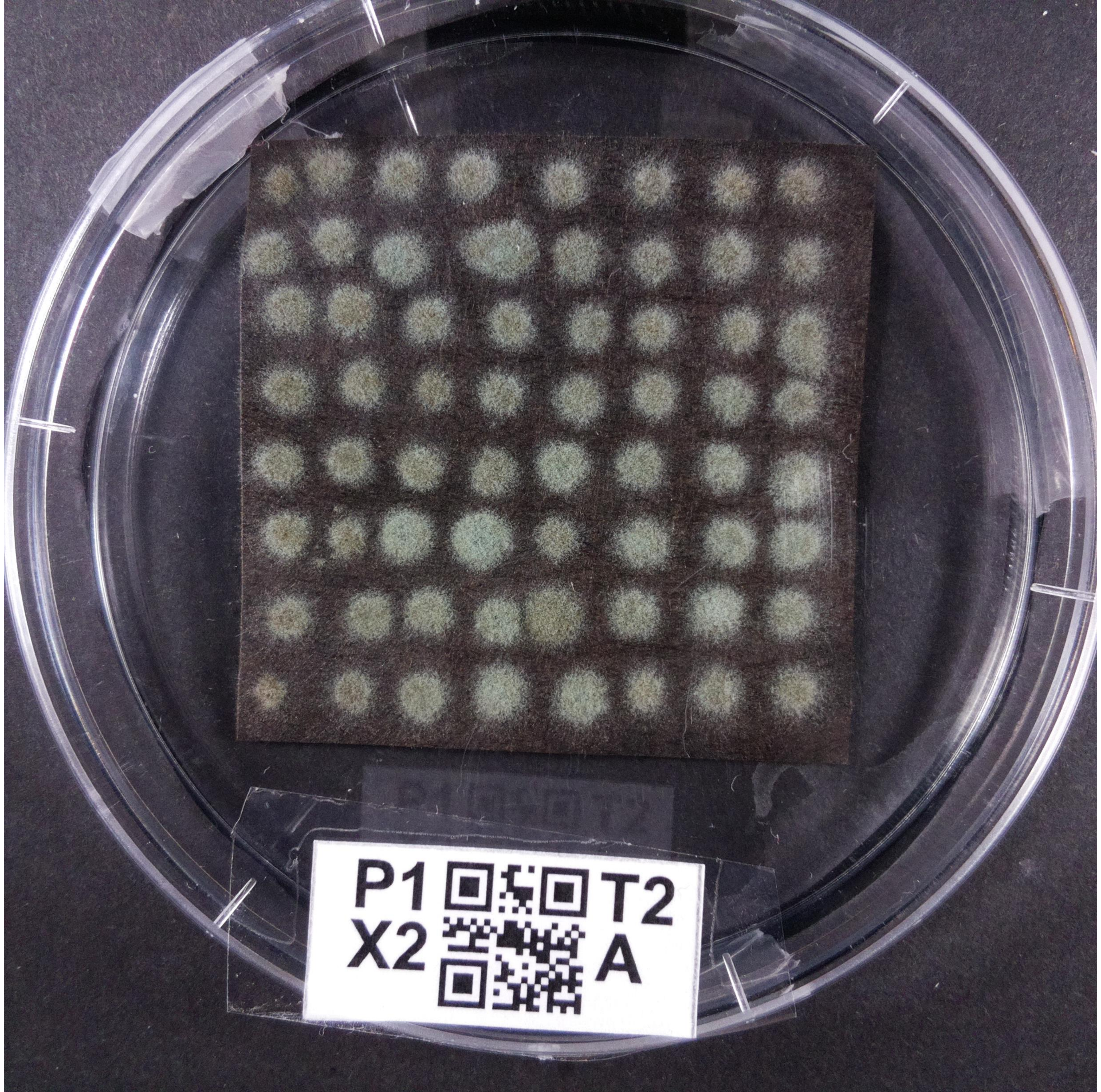


https://github.com/NVIDIA-AI-IOT/redtail

E/S/R
Science for Communities

Cohort 2


ACCELERATOR
E/S/R
Science for Communities
Decision Paper
Data Science Accelerator application form
Review of applications
DATA SCIENCE SPONSOR GROUP
TO: Data Science Sponsor Group
FROM: Richard Dean, Data Scientist
MEETING DATE: 16 December 2019
SUBJECT: Data Science Accelerator – Cohort 2 - Applications
Purpose
This is a summary of the data science accelerator applications submitted for consideration for cohort 2 as promised in 6.3 of the decision paper.
Applications were scored 1-5 against the following assessment criteria to form an initial assessment.
Assessment criteria
The final decision about which projects to take forward will be made by the data science sponsor group on Monday 16th December.
Page 1 of 7

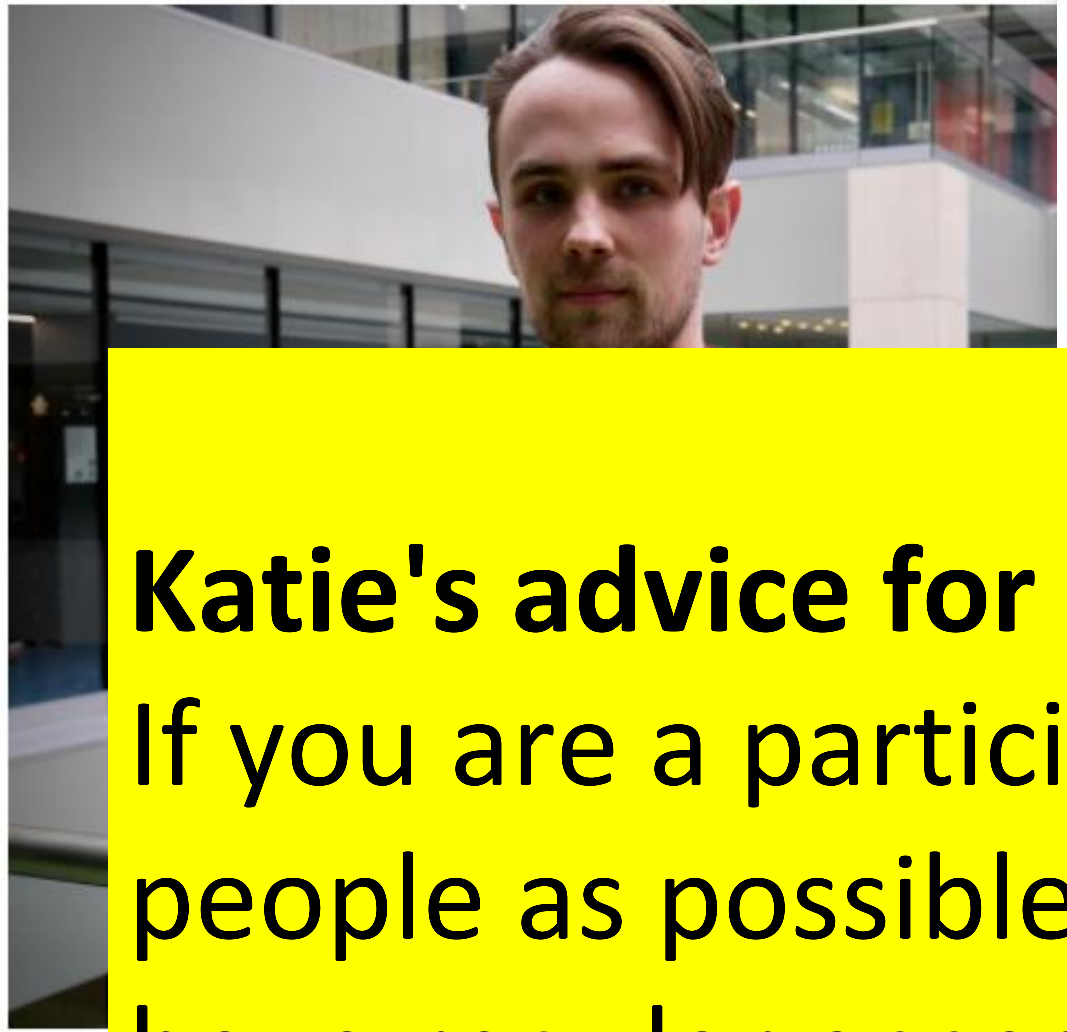


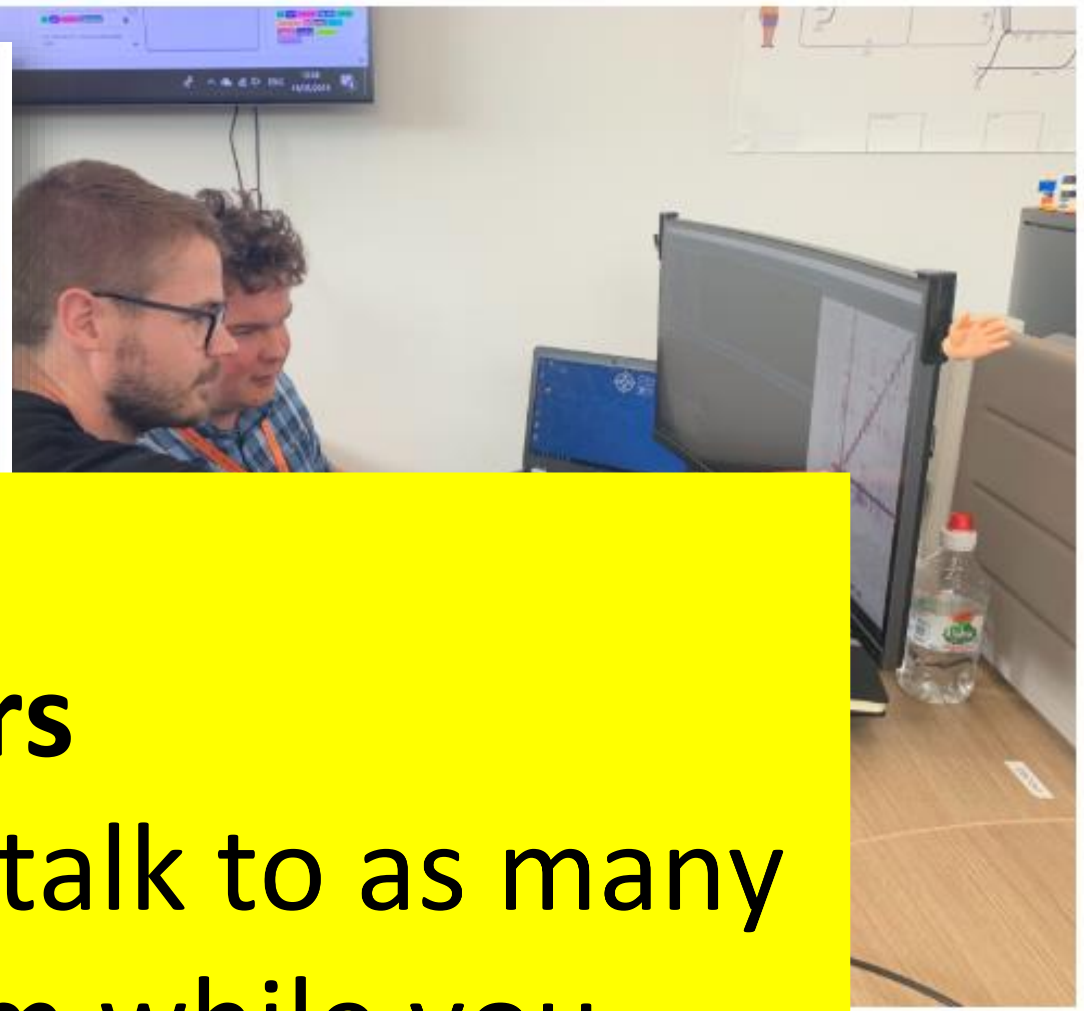
E/S/R
Science for Communities

The future…

Cockle Bay, Auckland, July 2018

What do the participants say?





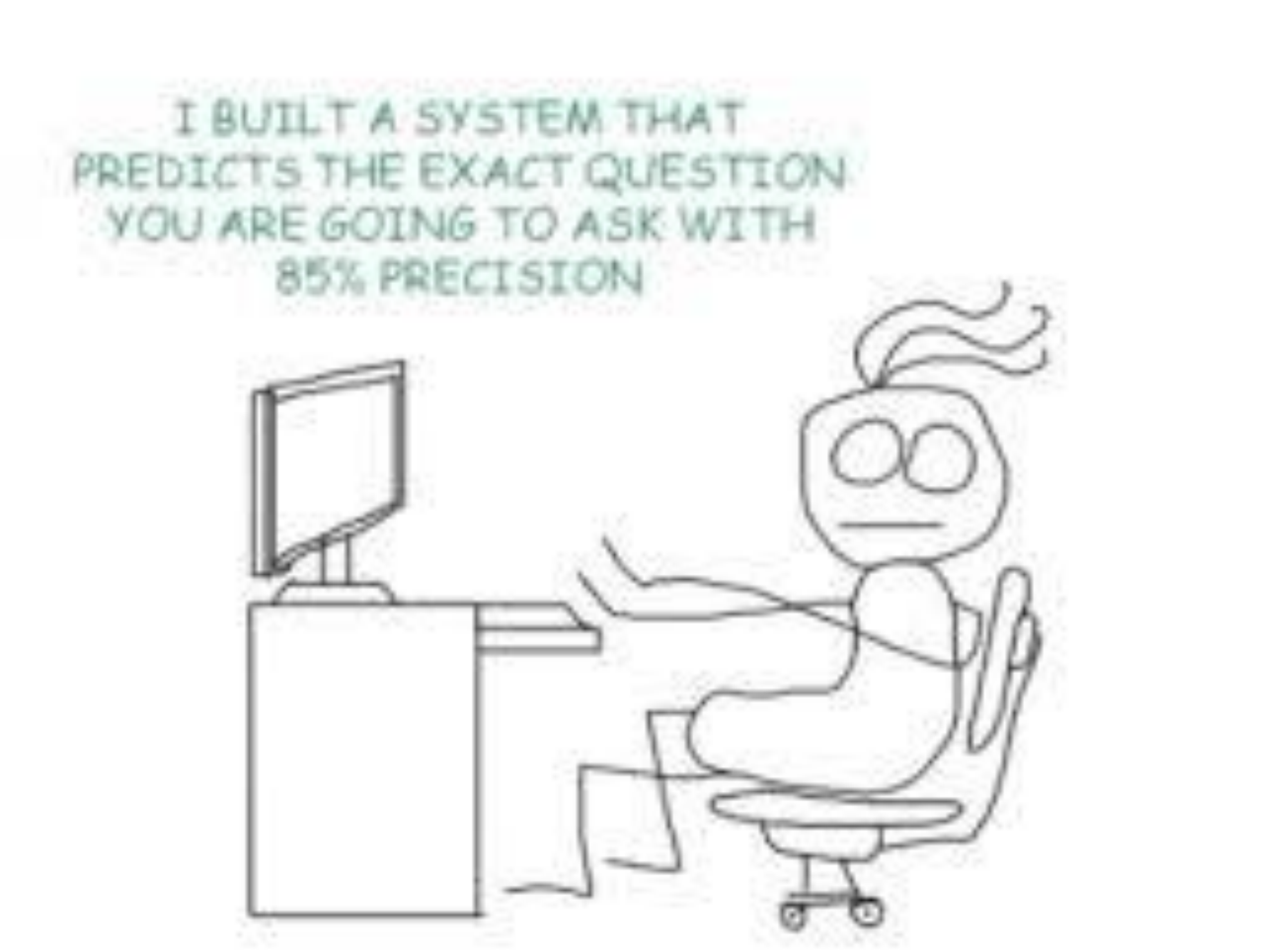
Katie's advice for participants and mentors

If you are a participant, ask questions and talk to as many people as possible - make the most of them while you have regular access to them. For mentors, do not underestimate your own abilities. We are always learning and data science is always evolving. So far, this experience has been incredibly rewarding as I get to help people do amazing things and learn more myself.

https://datasciencecampus.ons.gov.uk/data-science-accelerator-programme/

https://dataingovernment.blog.gov.uk/2019/05/13/going-from-mentee-to-mentor-with-the-data-science-accelerator/

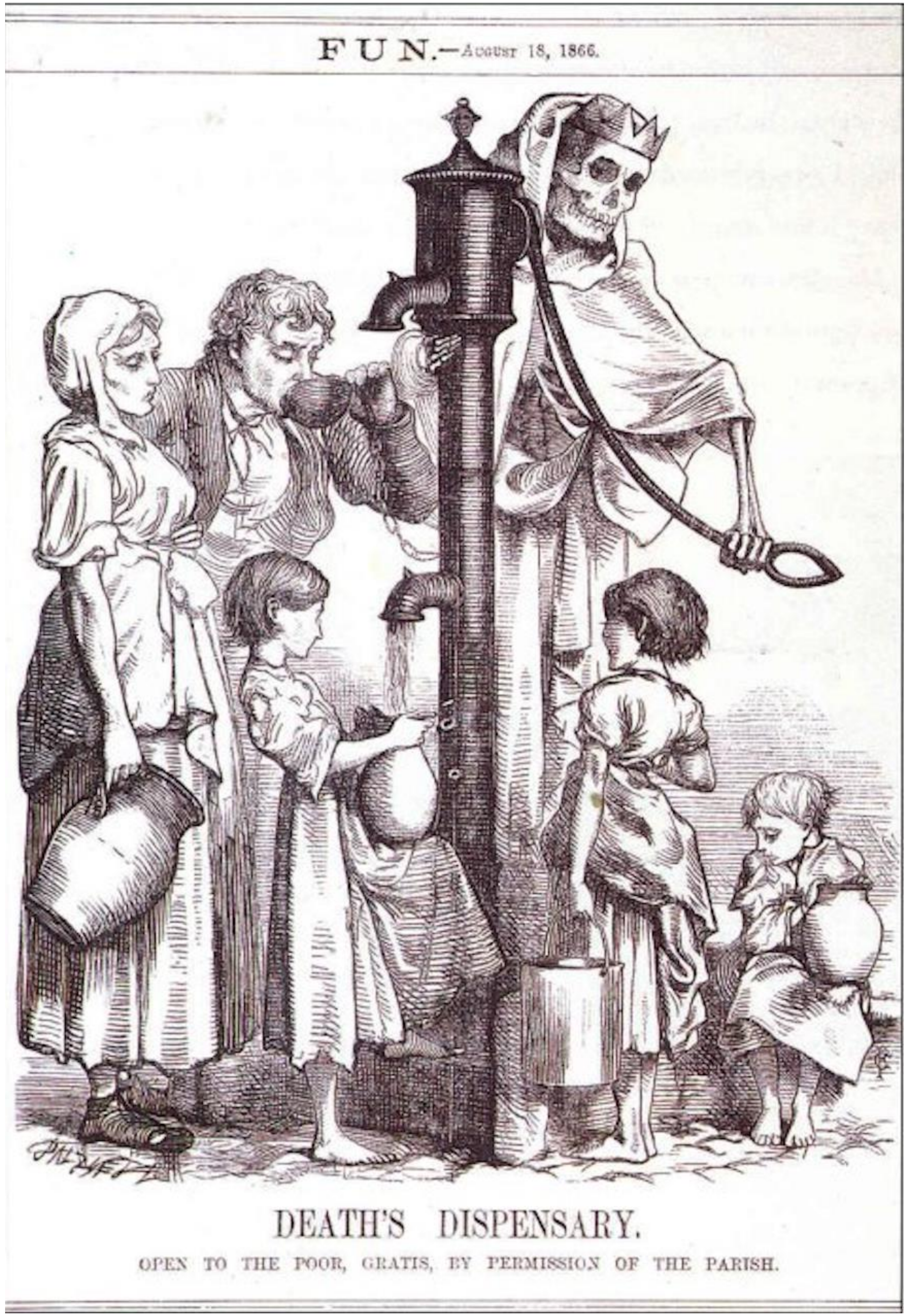
Science for Communities


I BUILT A SYSTEM THAT
PREDICTS THE EXACT QUESTION
YOU ARE GOING TO ASK WITH
85% PRECISION

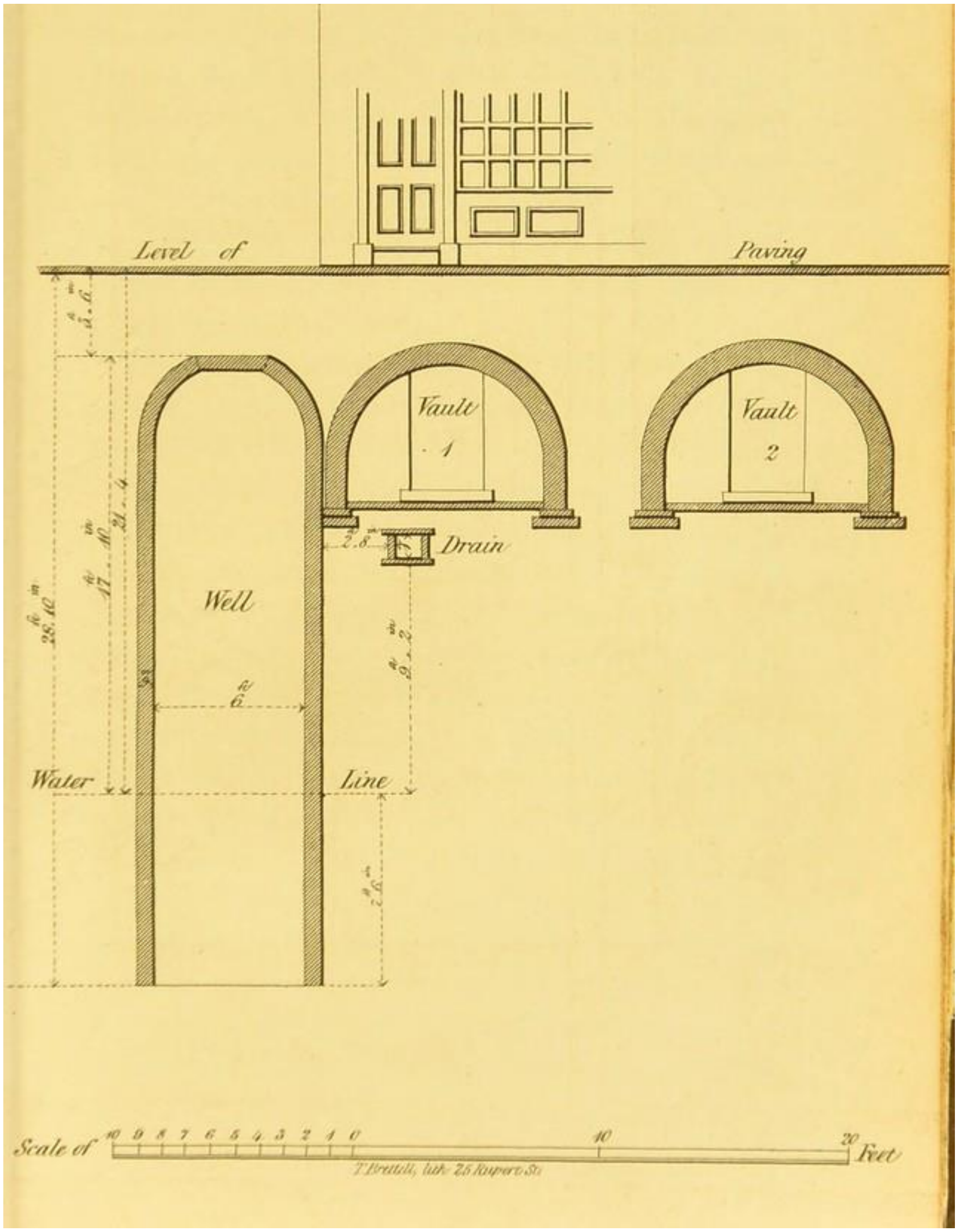
Plan Showing the Ascertained Deaths from Cholera. Credit: [Wellcome Collection](). [CC BY]()

The main drain of the house was opened in the front vault under the street, and was found to be constructed on the old fashioned plan of a flat bottom, 12 inches wide, with brick sides, rising about twelve inches high and covered with old stone. As this drain had but a small fall, or inclination outwards to the main sewer, the bottom was covered with an accumulation of soil deposit about two inches thick, and upon clearing this soil away the mortar joints of the old stone bottom were found to be perished, as was also all the jointing of the brick sides, which brought the brick work into the condition of a sieve, and through which the house drainage water must have percolated for a considerable period. Into this drain in the middle of the vault an intersecting smaller drain ran from the front stack pipe at the south-west angle of the front area, bringing the rain water from off the roof of the house, and also forming a communication with the drains of the adjoining house westward (No. 39.)

Upon opening back the main drain, a cesspool intended for a trap, but misconstructed, was found in the area 3ft. 8in. long, by 2ft. 6in. wide, and 3ft. deep, and upon, and over a part of this cesspool a common open privy (without water supply) for the use of the house was erected, the cesspool being fully charged with soil. This privy is formed across the east end of the area, and upon removing the soil the brickwork of the cesspool was found to be in the same decayed condition as the drain, and which may be better comprehended by stating that the bricks were easily lifted from their beds without any, the least force; so that any fluid could readily pass through the work, or as was the case when first opened, over the top course of bricks of the trap into the earth or made ground immediately under and adjoining the end wall eastward, this surface drainage being caused by the accumulation of soil



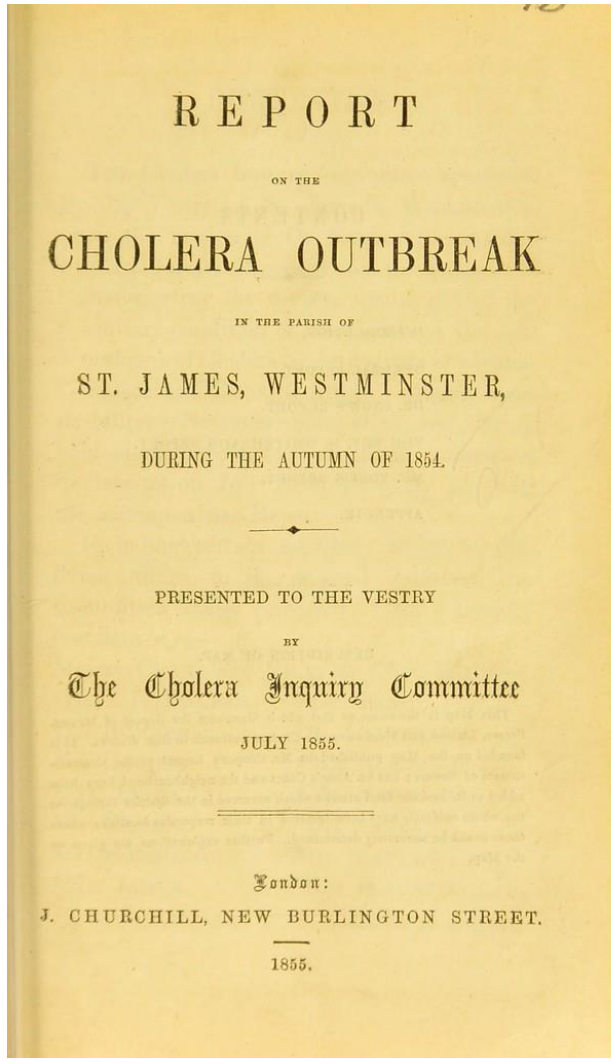
"Death's Dispensary" by George Pinwell



Science for Communities

"A Woman Dropping Her Tea-cup in Horror upon Discovering the Monstrous Contents of a Magnified Drop of Thames Water Revealing the Impurity of London Drinking Water"

Courtesy of the World Digital Library: https://www.wdl.org/en/item/3956/


REPORT

ON THE

CHOLERA OUTBREAK

IN THE PARISH OF

ST. JAMES, WESTMINSTER,

DURING THE AUTUMN OF 1854.

PRESENTED TO THE VESTRY

BY

The Cholera Inquiry Committee

JULY 1855.

London:

J. CHURCHILL, NEW BURLINGTON STREET.

1855.

Science for Communities

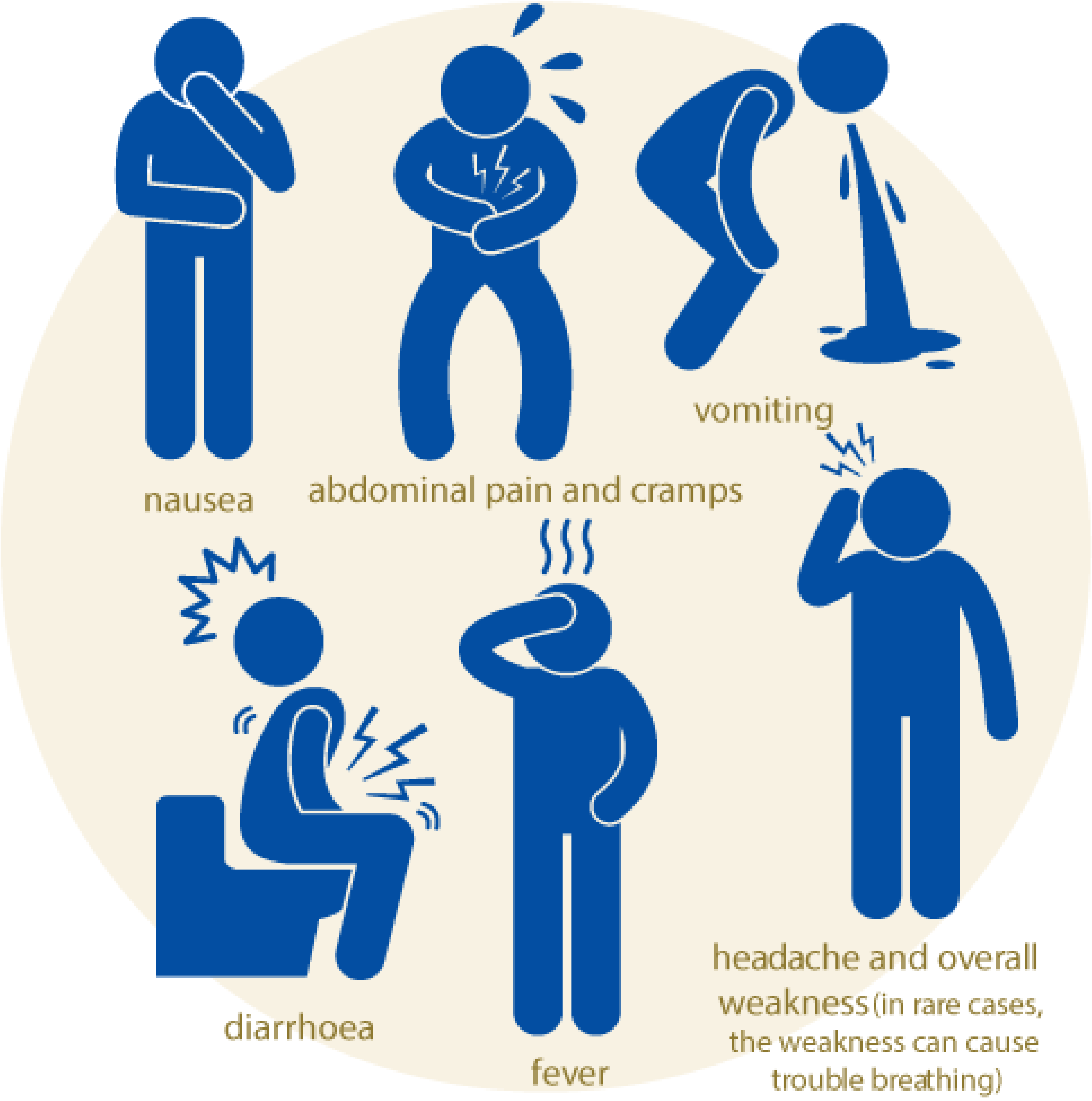
Science for Communities

Tuesday 16th August 2016

ESR
Science for Communities

Campylobacteriosis

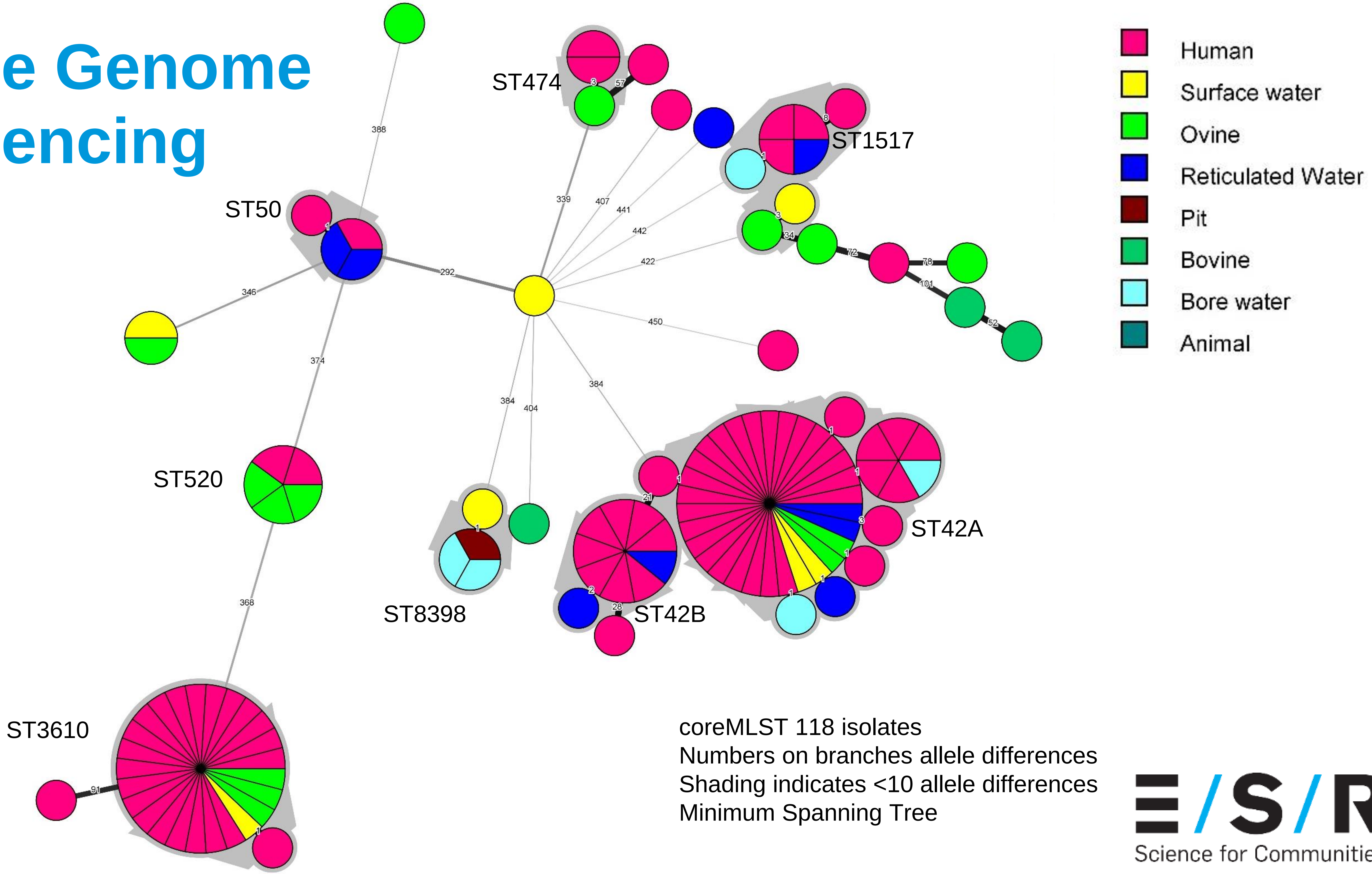
Campylobacter jejuni



Saturday Monday Wednesday Friday

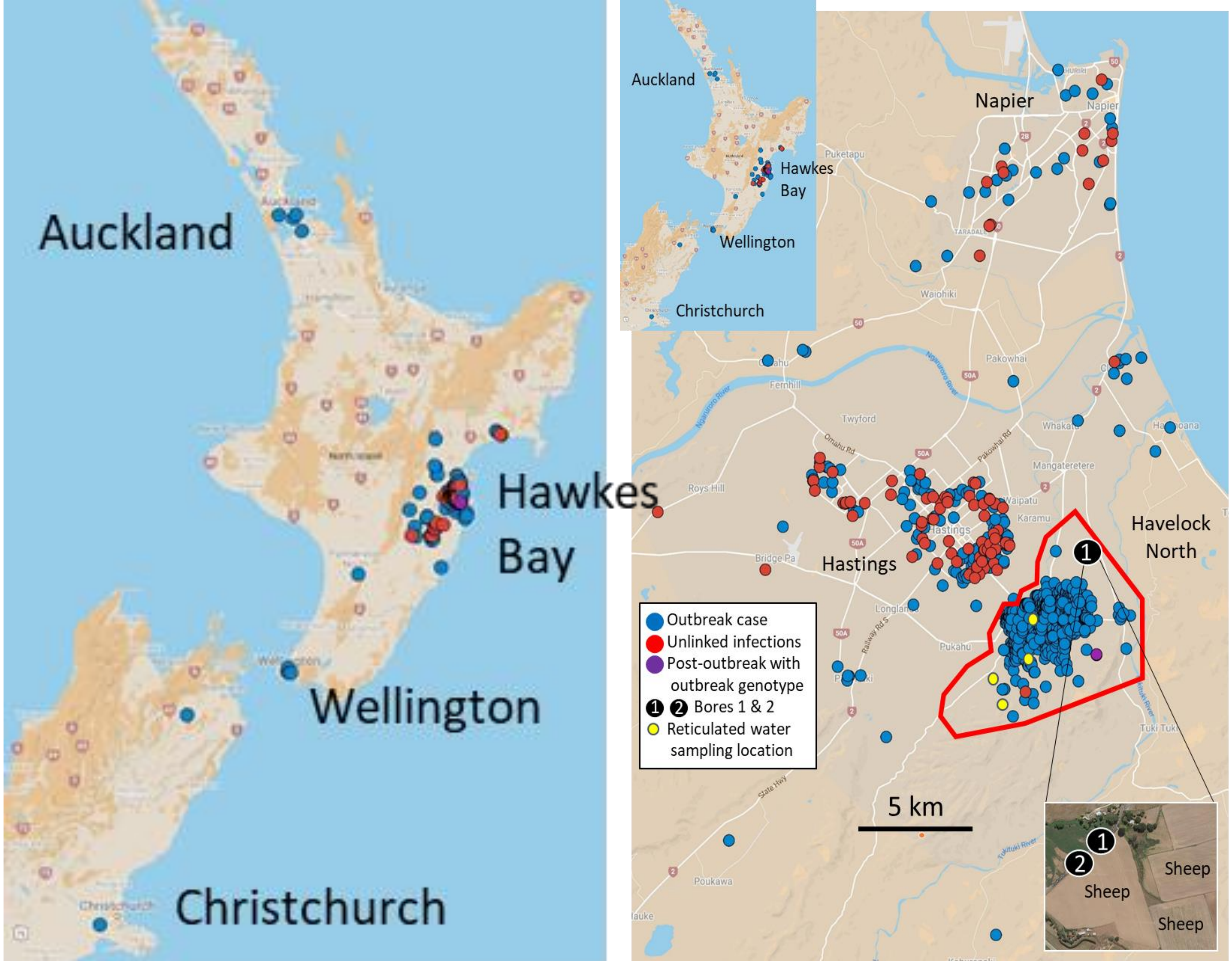
ESR
Science for Communities

Whole Genome Sequencing



coreMLST 118 isolates
Numbers on branches allele differences
Shading indicates <10 allele differences
Minimum Spanning Tree

Science for Communities

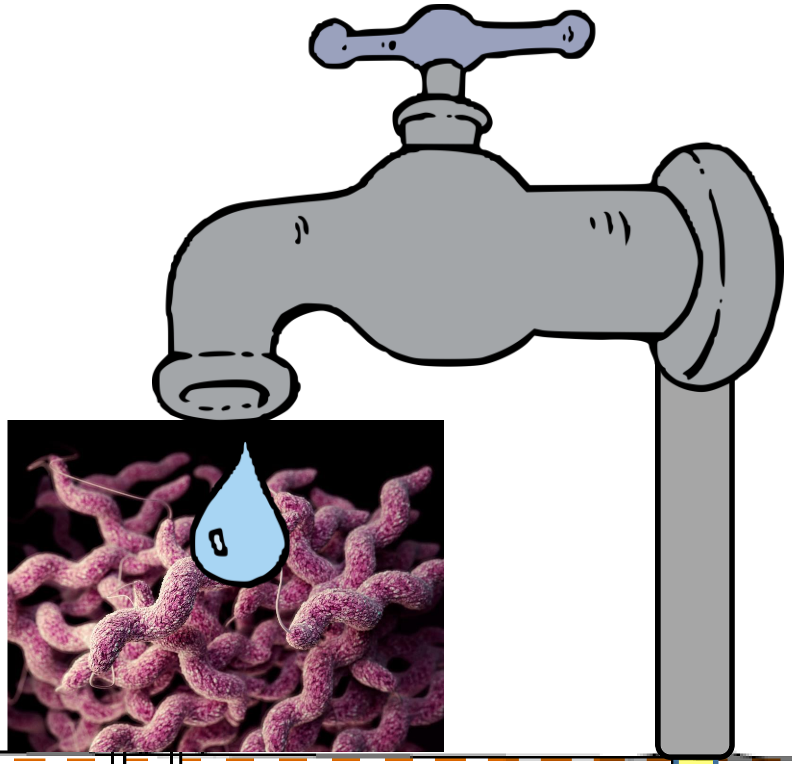


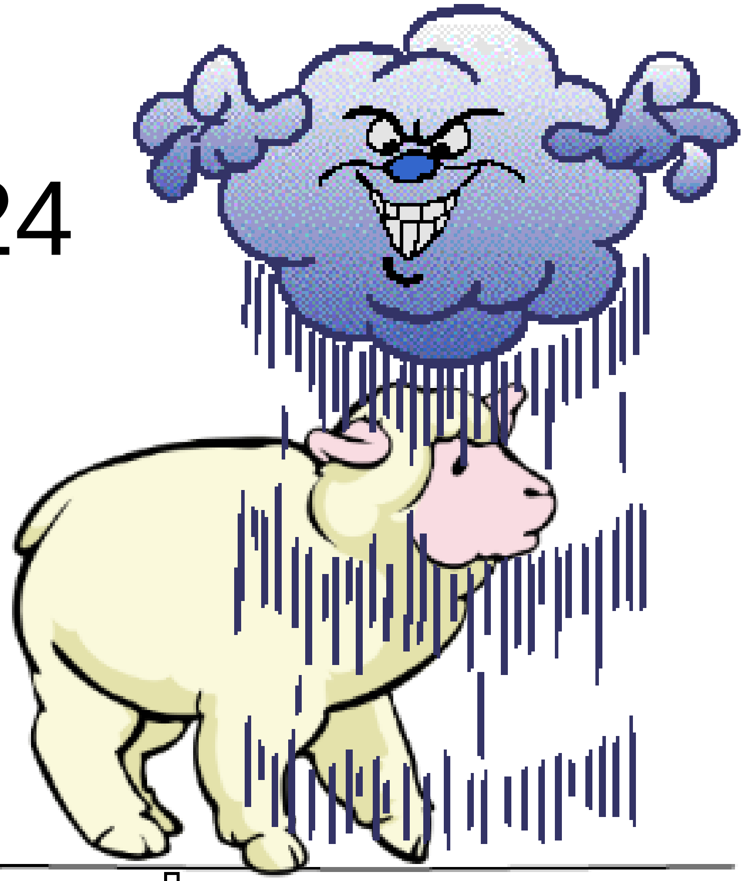
953 notified cases

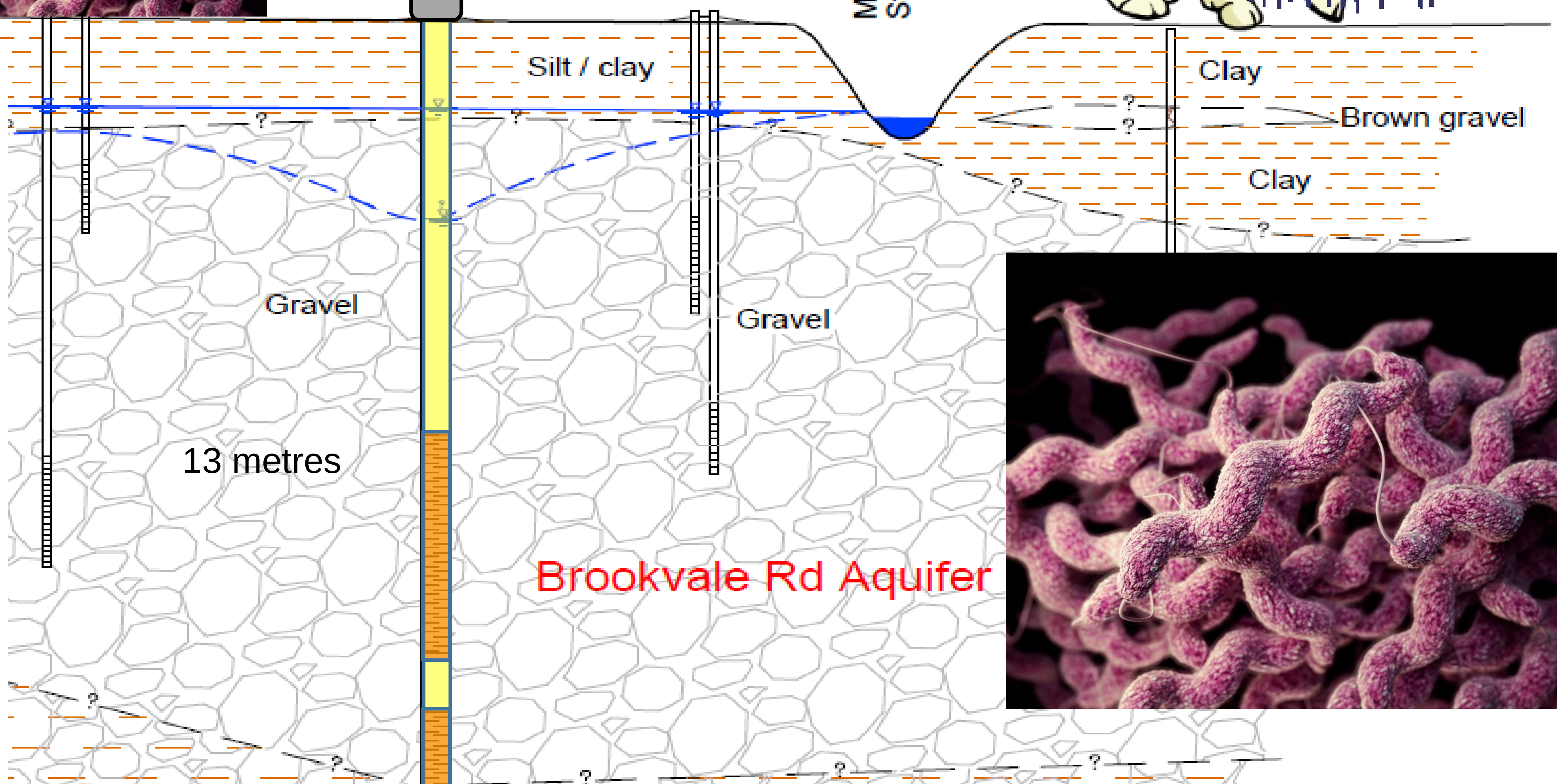
764 Havelock North reticulated water zone,
96 Hastings,
21 in Napier,
50 in other areas of the Hawkes Bay, 23 outside of the region

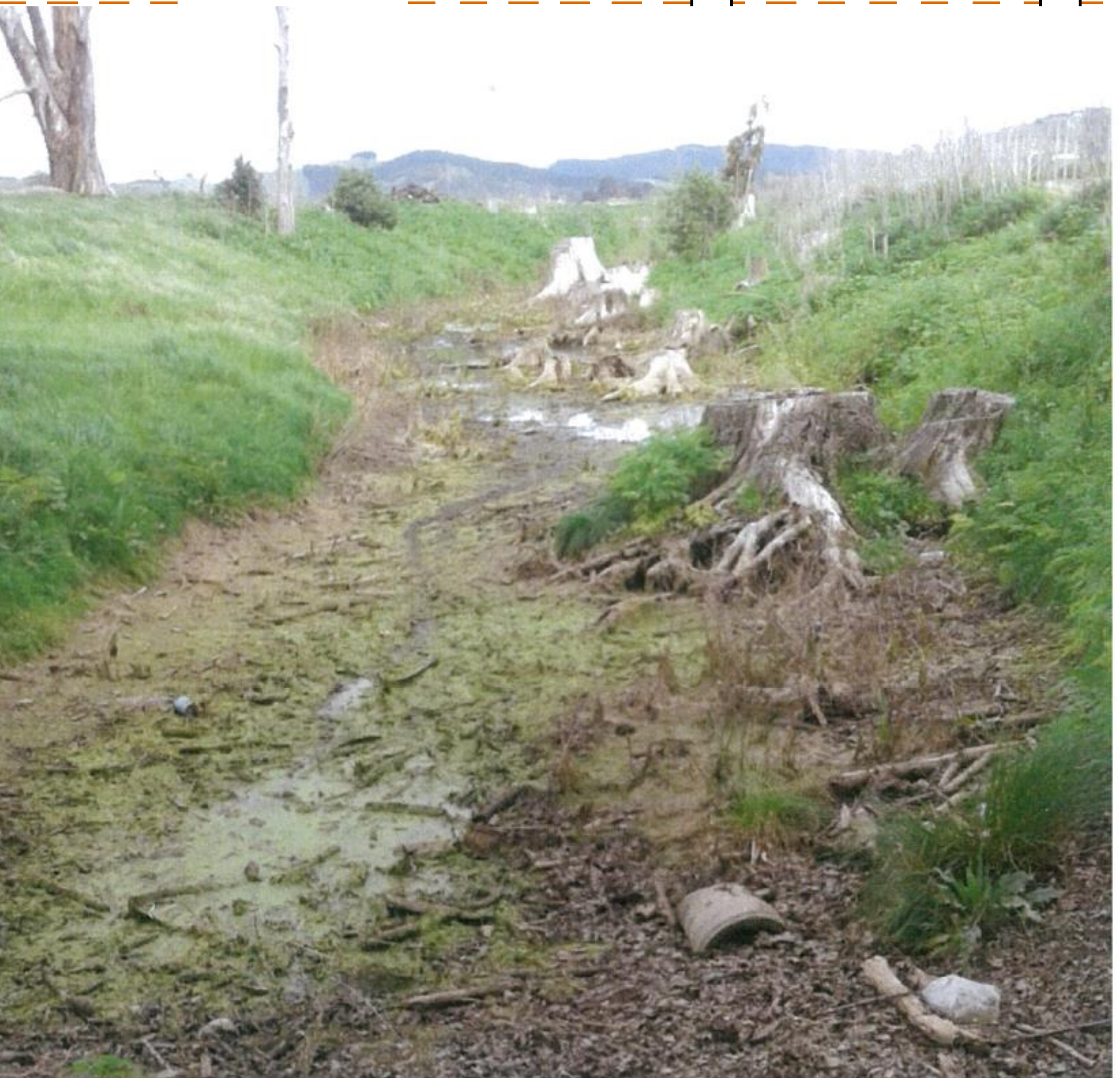
Science for Communities

165mm (6.5 inches) rain in 24 hours



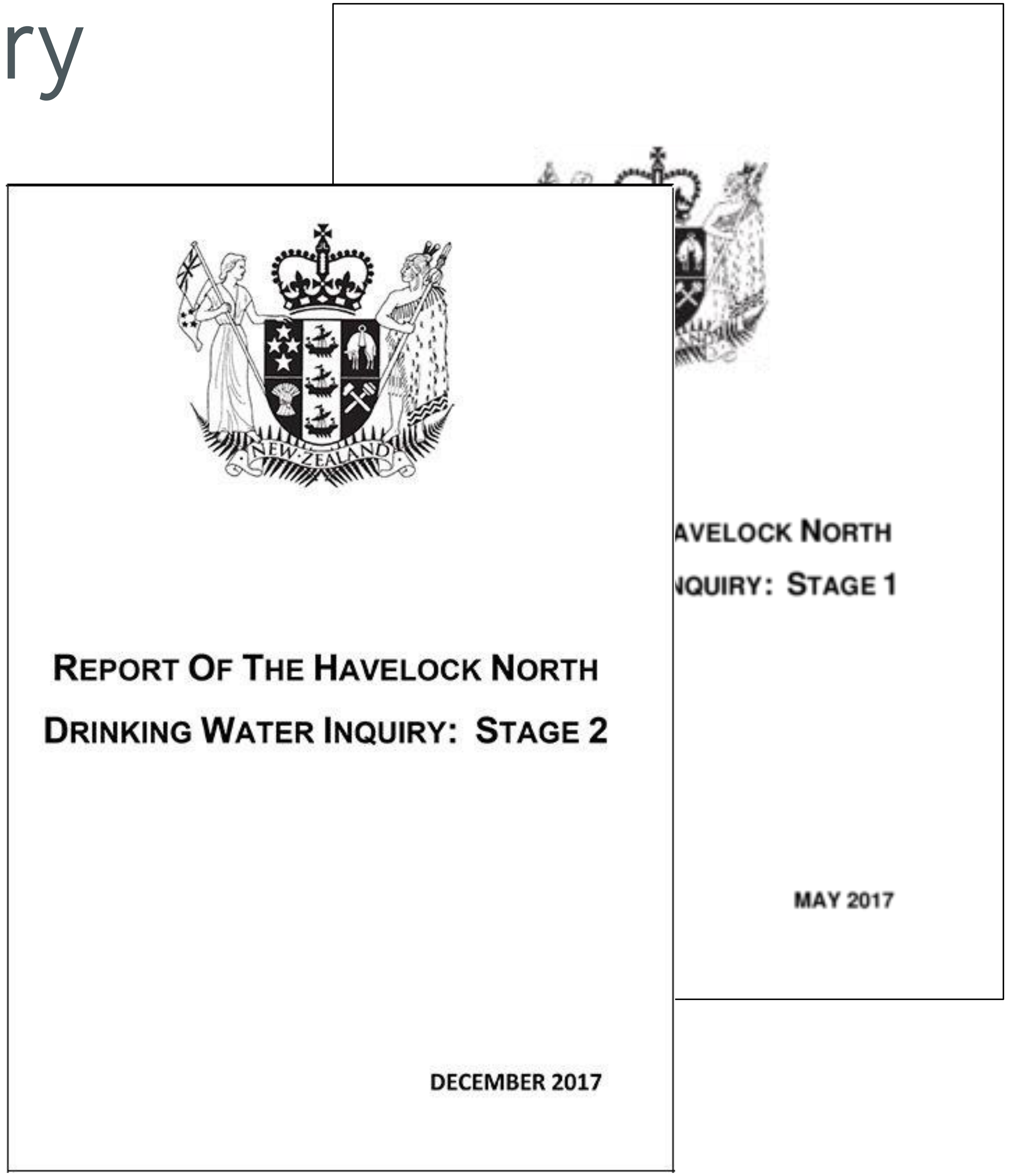






The Havelock North Water Inquiry
Concluding Comments

- *E. coli* O157 and *Campylobacter*
- 2,300 + illnesses
- 6 deaths
- 22 children permanent kidney damage



The Havelock North water inquiry panel Dr Karen Poutasi, left, Lyn Stevens QC and Anthony Wilson Wilson

E/S/R Science for Communities

We’ve covered…

- Data -> Information -> Knowledge -> Wisdom
- We see this as important …
- We’ve developed our capability...
- It would be awesome if you join…
- Our journey continues…

ESR
Science for Communities

Key messages

- Team with **super powers**
- Allow them to **evolve**
- Make **friends**
- … be like **Ash**

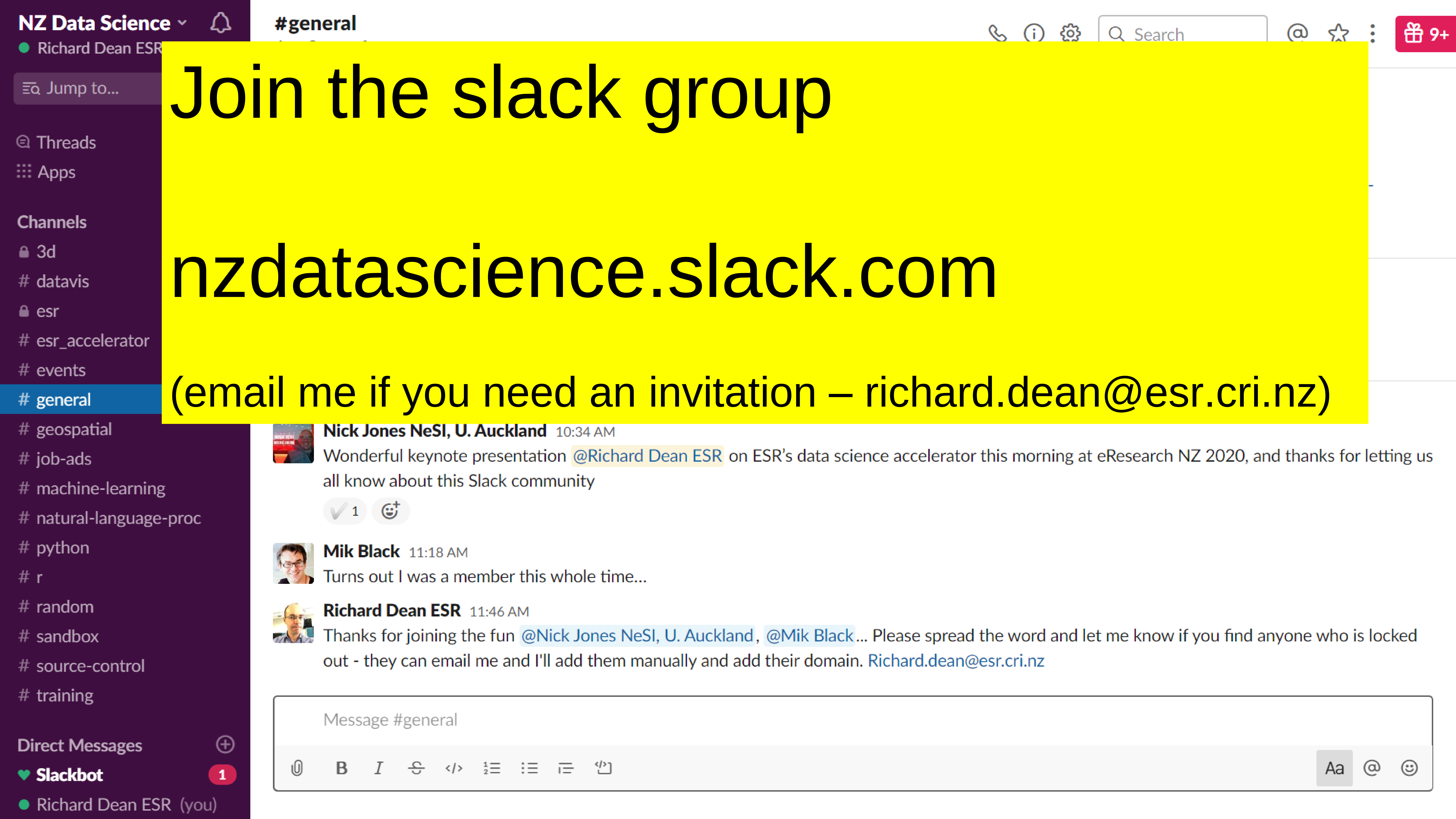
Credit - How To Draw Ash Ketchum From Pokémon! - YourDrawingLessons.com

E/S/R Science for Communities

Join the slack group

nzdatascience.slack.com

(email me if you need an invitation – richard.dean@esr.cri.nz)



richard.dean@esr.cri.nz

nzdatascience.slack.com

linkedin.com/in/ricdean