From Bits to Bytes
Software Development for Genetics, Genetic Resources, Plant Breeding and Genebanks

SSCR Winter Soft Fruit Meeting
23rd February 2023 Dundee

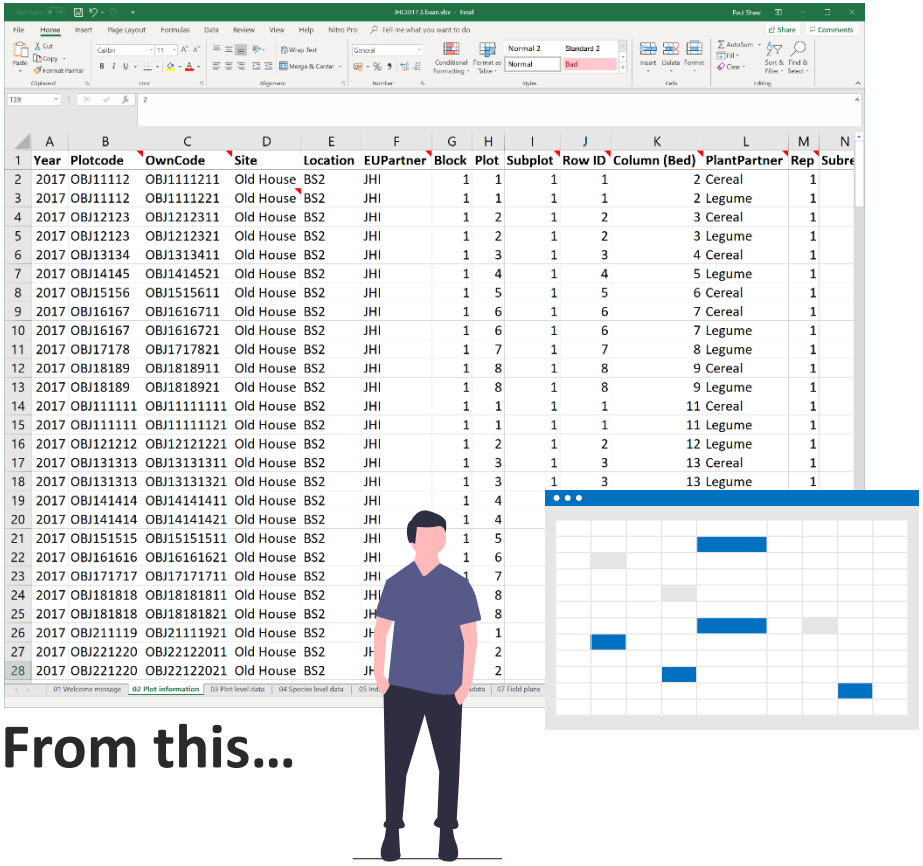
Paul Shaw

The International Barley Hub/James Hutton Institute Scotland

The challenge

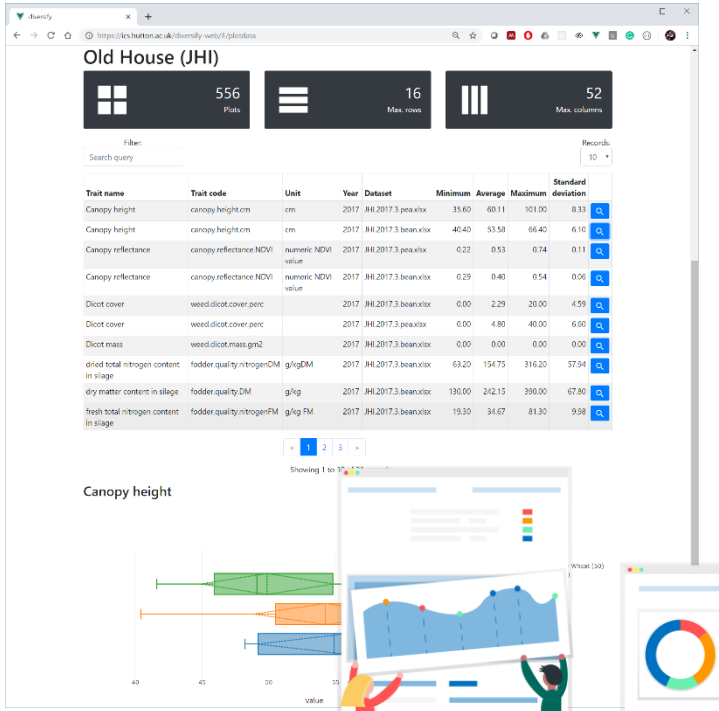
- Our science is producing **larger** and **more complex** datasets.
- We need to provide the infrastructure and tools to **store**, **explore**, **interact** and **understand** these large datasets.
- We need to develop tools to help provide **insight** into complex biological data.
- We need to provide **easy access** to huge volumes of data in **digestible** chunks.
- Show **relationships** between data that may not be visible in raw data.
- It's not about telling people what to think but giving them tools to look for **patterns** and **visual cues**.

In a nutshell…



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Year	Plotcode	OwnCode	Site	Location	EUPartner	Block	Plot	Subplot	Row ID	Column (Bed)	PlantPartner	Rep	Subre
2	2017	OBJ11112	OBJ1111211	Old House	BS2	JHI	1	1	1	1	2	Cereal	1	
3	2017	OBJ11112	OBJ1111221	Old House	BS2	JHI	1	1	1	1	2	Legume	1	
4	2017	OBJ12123	OBJ1212311	Old House	BS2	JHI	1	2	1	2	3	Cereal	1	
5	2017	OBJ12123	OBJ1212321	Old House	BS2	JHI	1	2	1	2	3	Legume	1	
6	2017	OBJ13134	OBJ1313411	Old House	BS2	JHI	1	3	1	3	4	Cereal	1	
7	2017	OBJ14145	OBJ1414521	Old House	BS2	JHI	1	4	1	4	5	Legume	1	
8	2017	OBJ15156	OBJ1515611	Old House	BS2	JHI	1	5	1	5	6	Cereal	1	
9	2017	OBJ16167	OBJ1616711	Old House	BS2	JHI	1	6	1	6	7	Cereal	1	
10	2017	OBJ16167	OBJ1616721	Old House	BS2	JHI	1	6	1	6	7	Legume	1	
11	2017	OBJ17178	OBJ1717821	Old House	BS2	JHI	1	7	1	7	8	Legume	1	
12	2017	OBJ18189	OBJ1818911	Old House	BS2	JHI	1	8	1	8	9	Cereal	1	
13	2017	OBJ18189	OBJ1818921	Old House	BS2	JHI	1	8	1	8	9	Legume	1	
14	2017	OBJ111111	OBJ11111111	Old House	BS2	JHI	1	1	1	1	11	Cereal	1	
15	2017	OBJ111111	OBJ11111121	Old House	BS2	JHI	1	1	1	1	11	Legume	1	
16	2017	OBJ121212	OBJ12121221	Old House	BS2	JHI	1	2	1	2	12	Legume	1	
17	2017	OBJ131313	OBJ13131311	Old House	BS2	JHI	1	3	1	3	13	Cereal	1	
18	2017	OBJ131313	OBJ13131321	Old House	BS2	JHI	1	3	1	3	13	Legume	1	
19	2017	OBJ141414	OBJ14141411	Old House	BS2	JHI	1	4						
20	2017	OBJ141414	OBJ14141421	Old House	BS2	JHI	1	4						
21	2017	OBJ151515	OBJ15151511	Old House	BS2	JHI	1	5						
22	2017	OBJ161616	OBJ16161621	Old House	BS2	JHI	1	6						
23	2017	OBJ171717	OBJ17171711	Old House	BS2	JHI	1	7						
24	2017	OBJ181818	OBJ18181811	Old House	BS2	JHI		8						
25	2017	OBJ181818	OBJ18181821	Old House	BS2	JH		8						
26	2017	OBJ211119	OBJ21111921	Old House	BS2	JH		1						
27	2017	OBJ221220	OBJ22122011	Old House	BS2	JH		2						
28	2017	OBJ221220	OBJ22122021	Old House	BS2	JH		2						

From this…



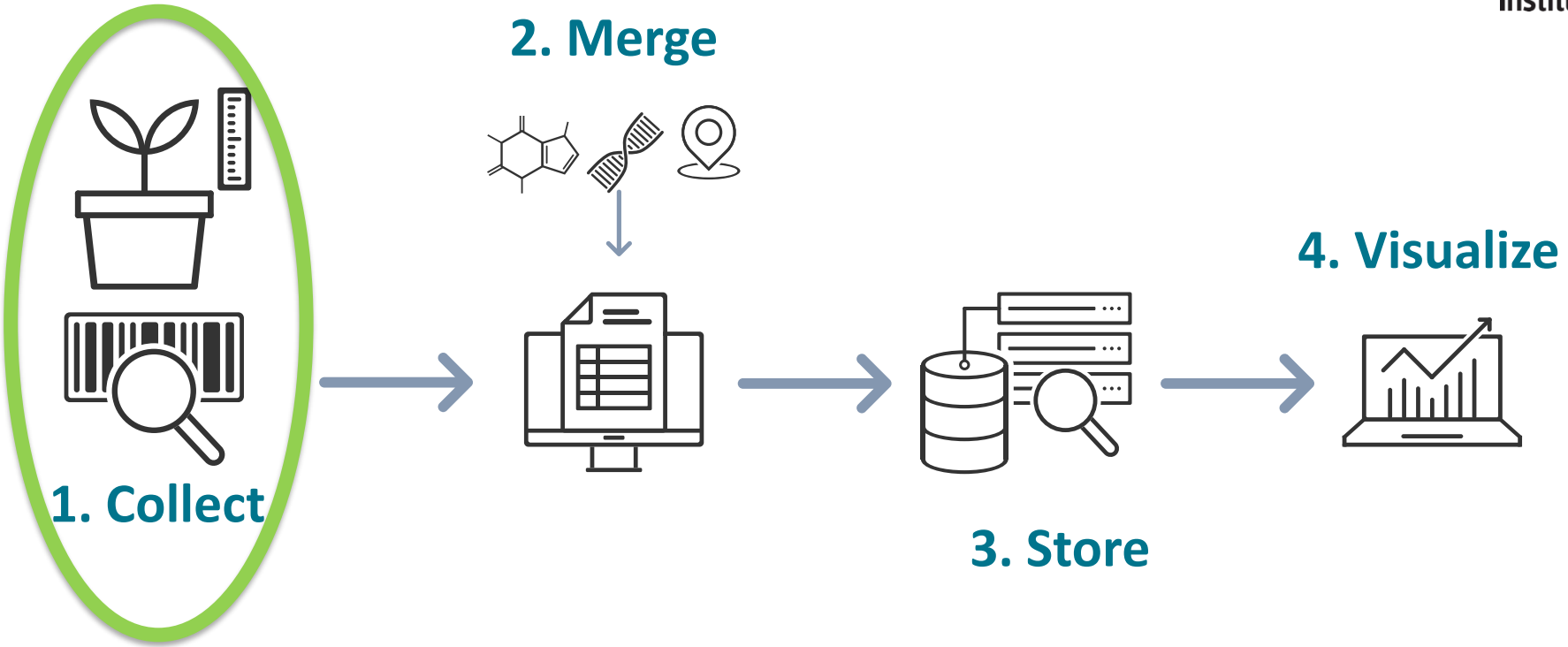
To this…

The James Hutton Institute

“Making experimental data available to breeders, scientific researchers and farmers in approachable and tailored ways will increase the utility and value of the data and knowledge the project will generate ensuring diversity is more accessible for exploitation.”

But how to help try to achieve this?

The James Hutton Institute

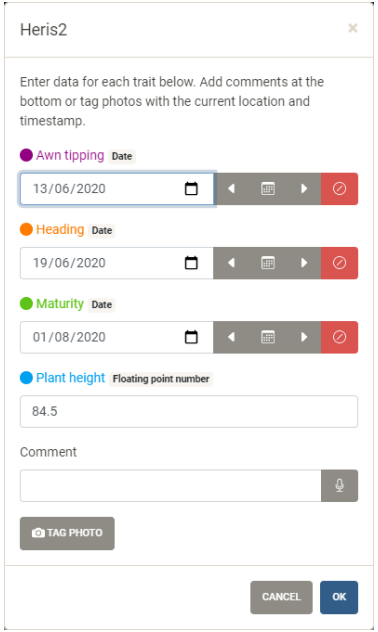


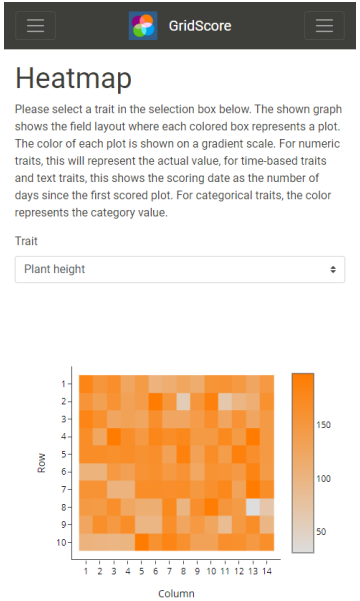
GridScore

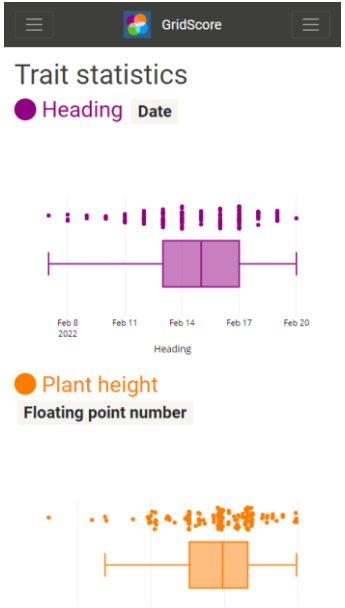
Raubach, S., Schreiber, M. & Shaw, P.D. GridScore: a tool for accurate, cross-platform phenotypic data collection and visualization. *BMC Bioinformatics* 23, 214 (2022). https://doi.org/10.1186/s12859-022-04755-2

Stop writing stuff down!

- ✓ Less error-prone data entry.
- ✓ Voice feedback using text-to-speech.
- ✓ Voice recognition for audio notes.
- ✓ Automatic tagging with timestamp, identifier and GPS.

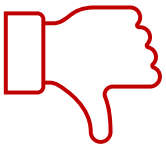




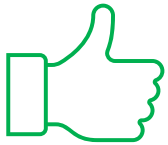


P1046371.JPG

P1046374.JPG

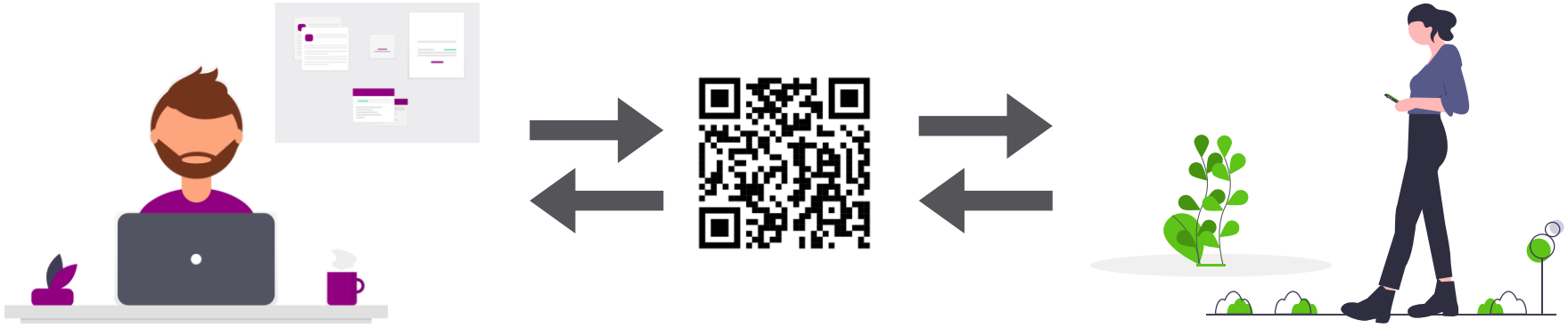


2021-03-18_16-19-11-Birka1.jpg



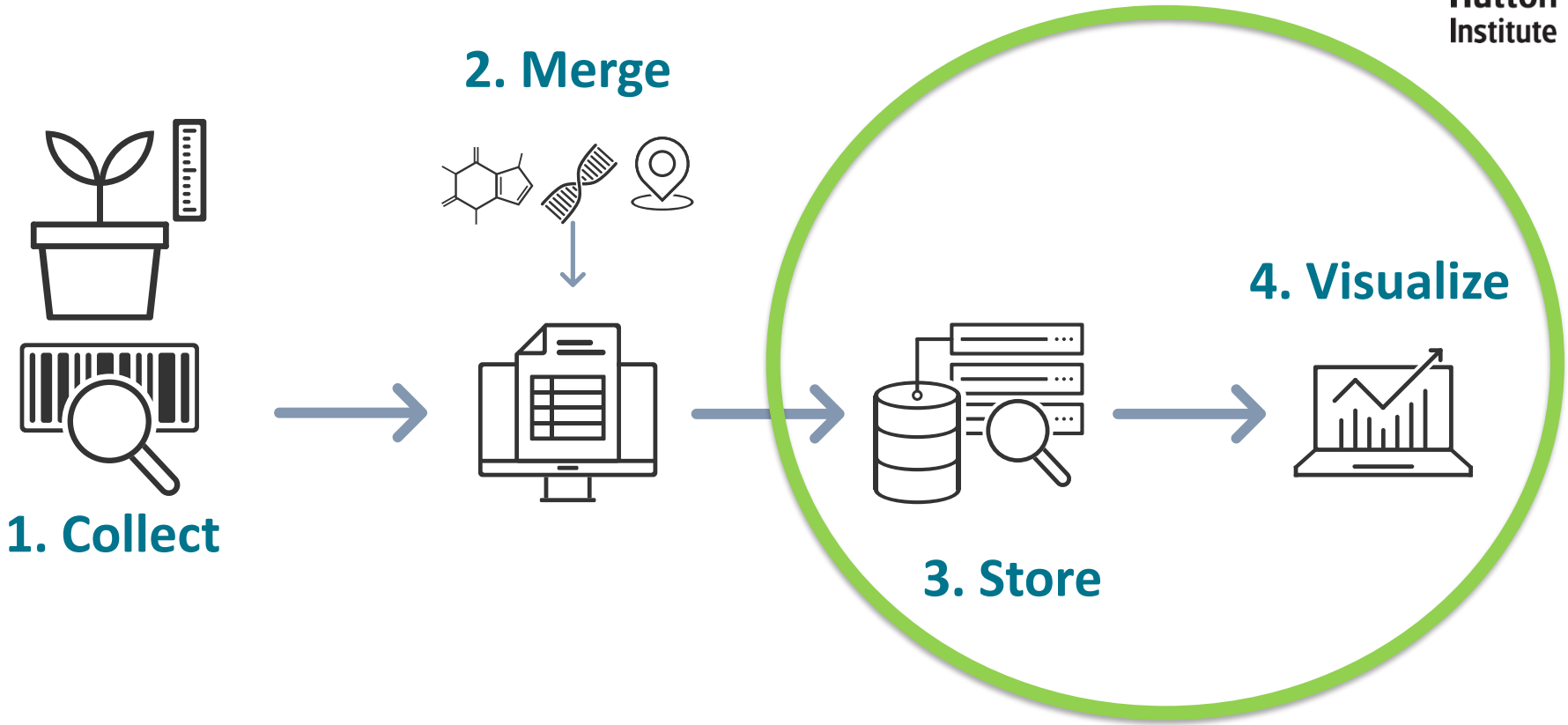
QR-based data sharing

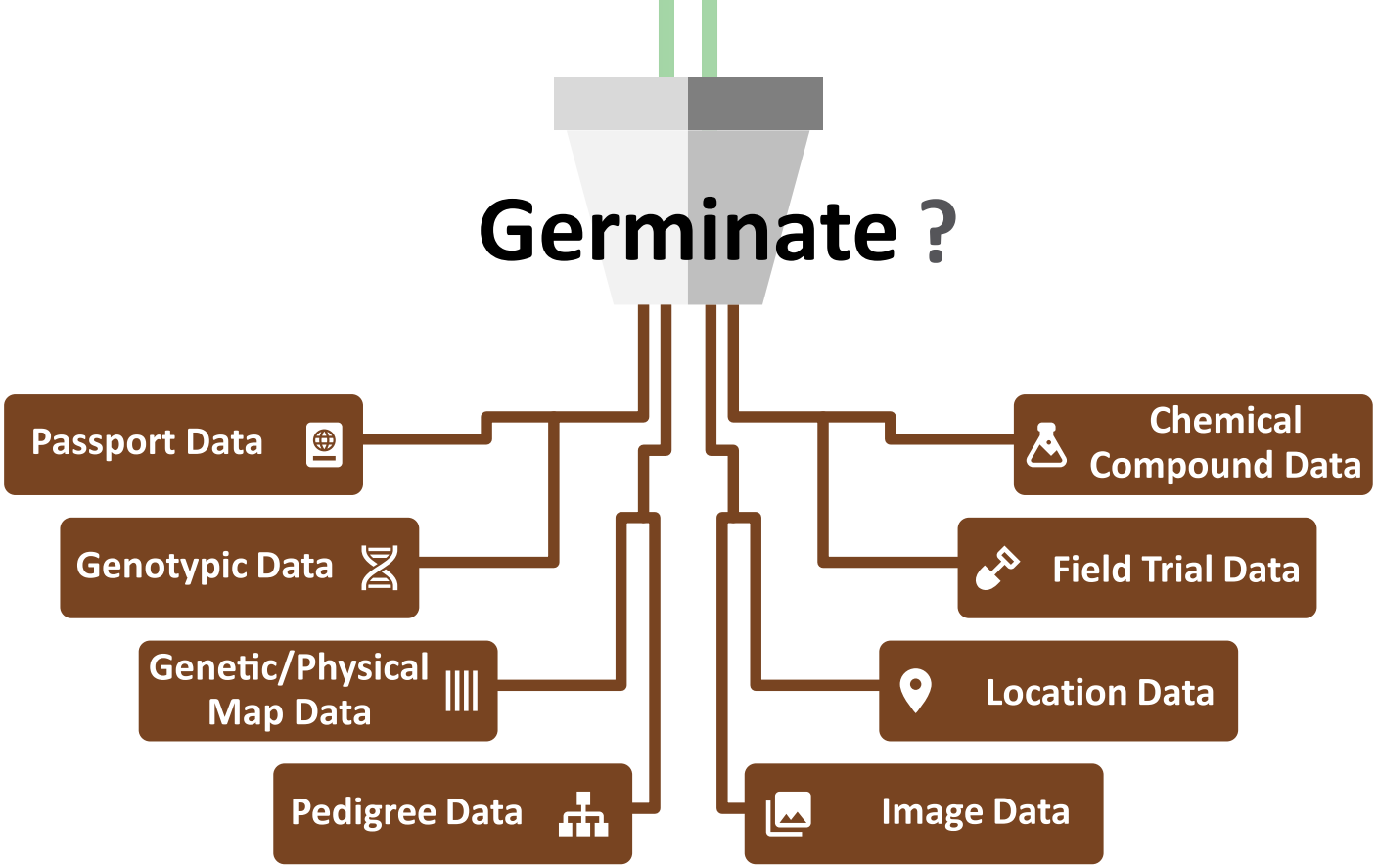
- Share trial setup and data in seconds using QR codes.
- No mucking about – same setup for all!



But how to help try to achieve this?

The James Hutton Institute




Germinate ?
Passport Data
Genotypic Data
Genetic/Physical Map Data
Pedigree Data
Chemical Compound Data
Field Trial Data
Location Data
Image Data

Germinate

Search

Home | MCPD | Institution | Datasets | Pedigree data | Location data | Images | Groups | Entity data | Additional attributes | User comments

- Data
  - Germplasm
  - Trials data
  - Geographic data
  - Datasets
  - Experiments
  - Data statistics
- Groups
- Images
- Search
- About

Images

This section shows all the images associated with this germplasm.

CF_T3_B2_SWEET

04/09/2019 12:36 pm

CF_T3_B2_SWEET

CF_T3_B2_SWEET

05/08/2019 11:51 am

CF_T3_B2_SWEET

CF_T3_B2_SWEET

22/07/2019 11:26 am

CF_T3_B2_SWEET

CF_T3_B2_W1_SWEET

20/06/2019 12:20 pm

CF_T3_B2_SWEET

CF_T3_B2_W2_SWEET

20/06/2019 12:20 pm

CF_T3_B2_SWEET

CF_T3_B2_W3_SWEET

20/06/2019 12:20 pm

CF_T3_B2_SWEET

CF_T3_B2_SWEET

20/06/2019 12:05 pm

CF_T3_B2_SWEET

CF_T3_B2_W1_SWEET

10/06/2019 11:47 am

CF_T3_B2_SWEET

« ‹ 1 2 3 4 › »

Download images | Upload

Groups

Open Data?
What people say....

What that really means…

Germinate species diversity

The James Hutton Institute

Barley

Berrybase

Demo

CPC

Grasses

CWR Cowpea

CWR Finger Millet

CWR Grass Pea

CWR Pigeon Pea

CWR Chickpea

CWR Alfalfa

CWR Lentil

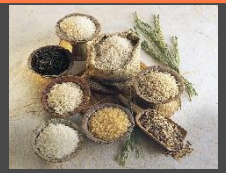
CWR Pearl Millet

CWR Barley

CWR Wheat

CWR Sorghum

CWR Eggplant


CWR Rice

CWR Sunflower

CIMMYT Maize

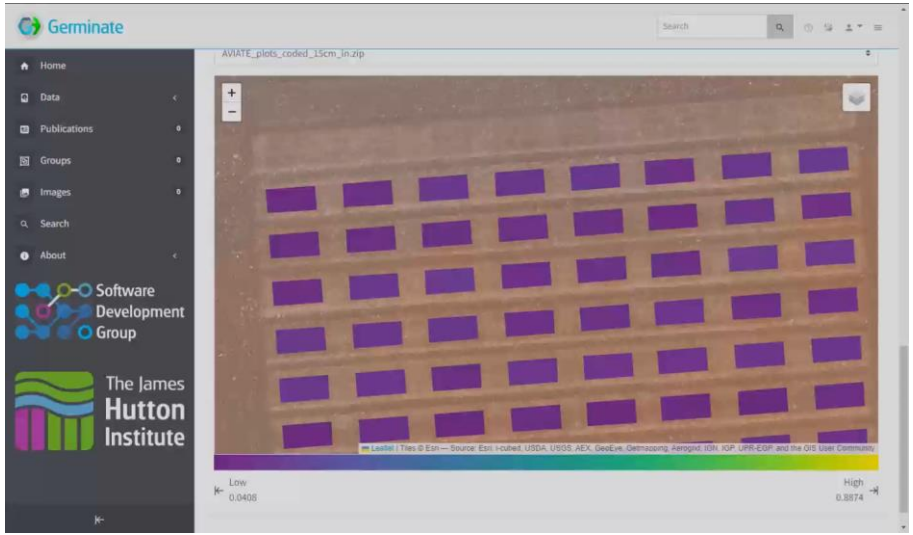
CIMMYT Wheat

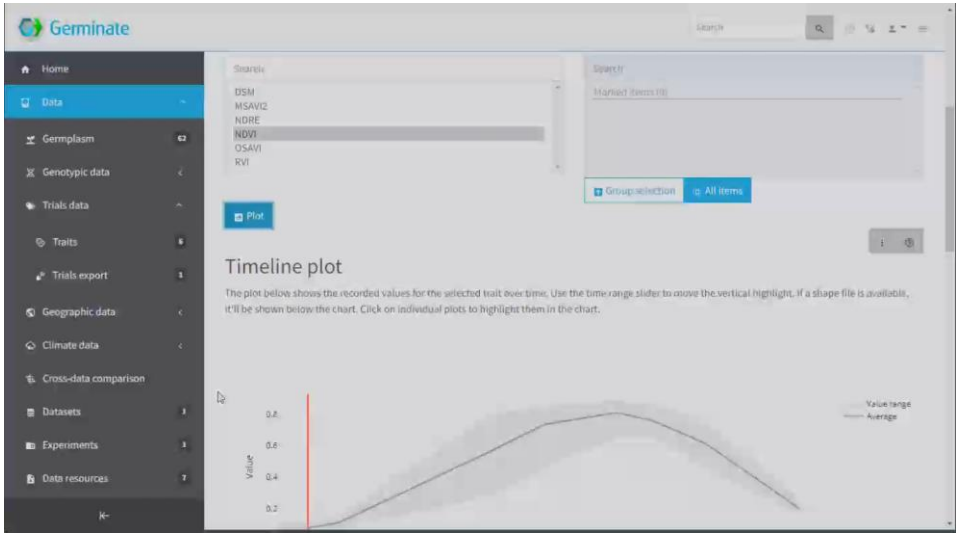
Rye

It’s not a competition – exploit other resources where you can!

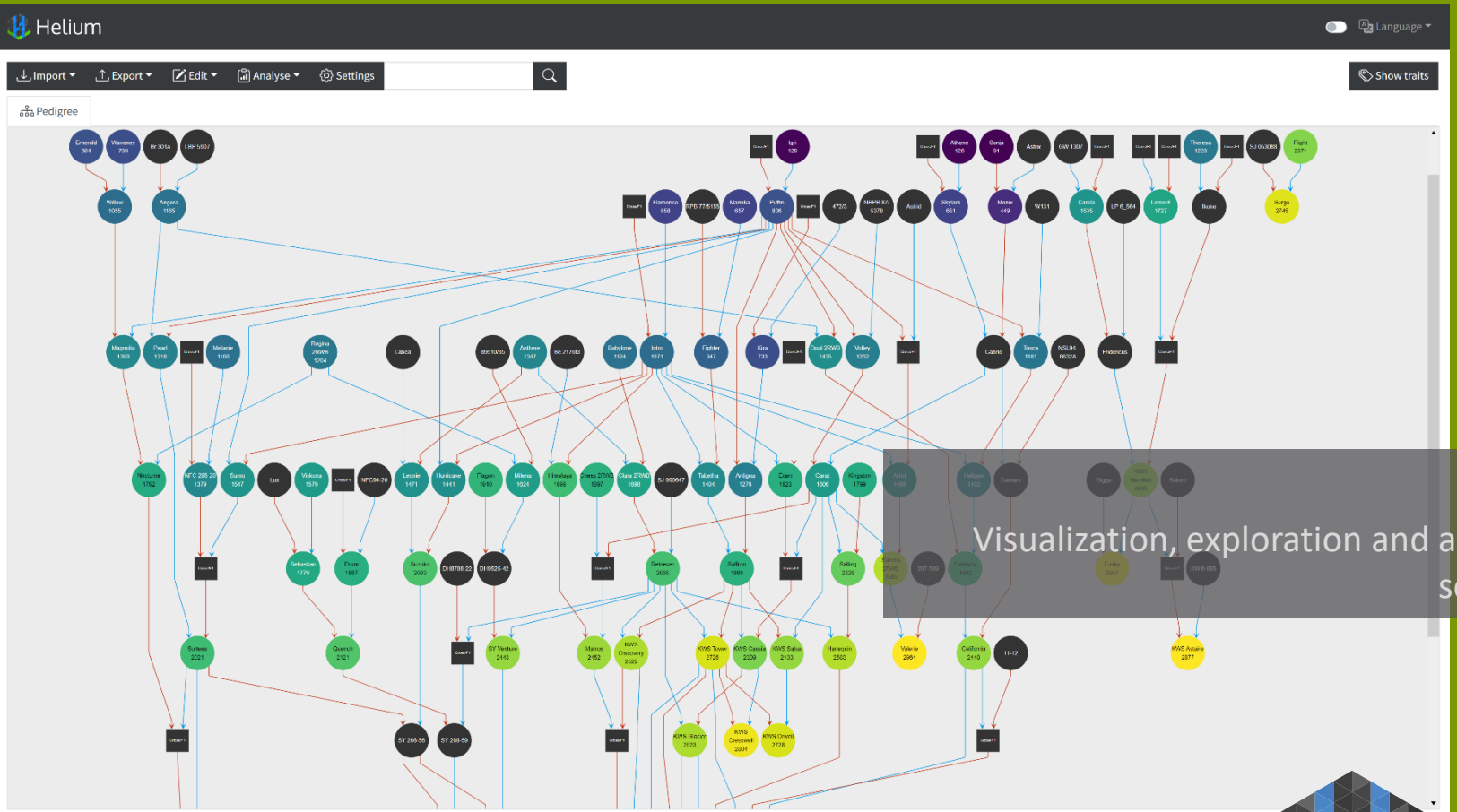
Germinate – RESAS JHI-B1-2 BARGAIN

The James Hutton Institute





UAV, ionomics and environmental sensor data.



The James Hutton Institute

Shaw, P.D. et al. Helium: visualization of large scale plant pedigrees. *BMC Bioinformatics*. 2014

Helium

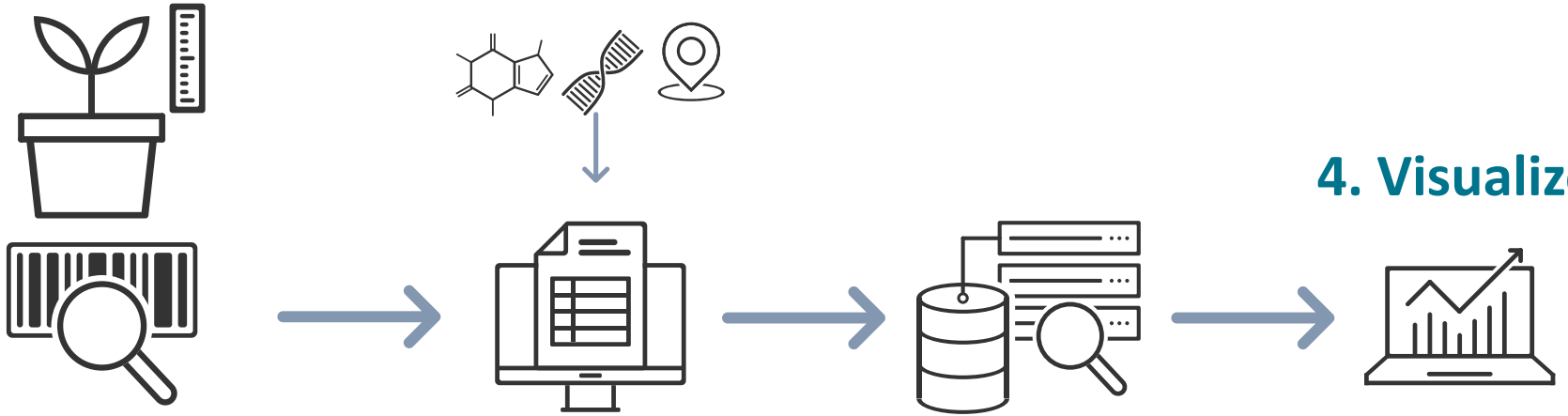
Visualization, exploration and analysis of large scale pedigrees.

https://helium.hutton.ac.uk

Coming soon…

But how to help try to achieve this?

The James Hutton Institute



GridScore

Germinate

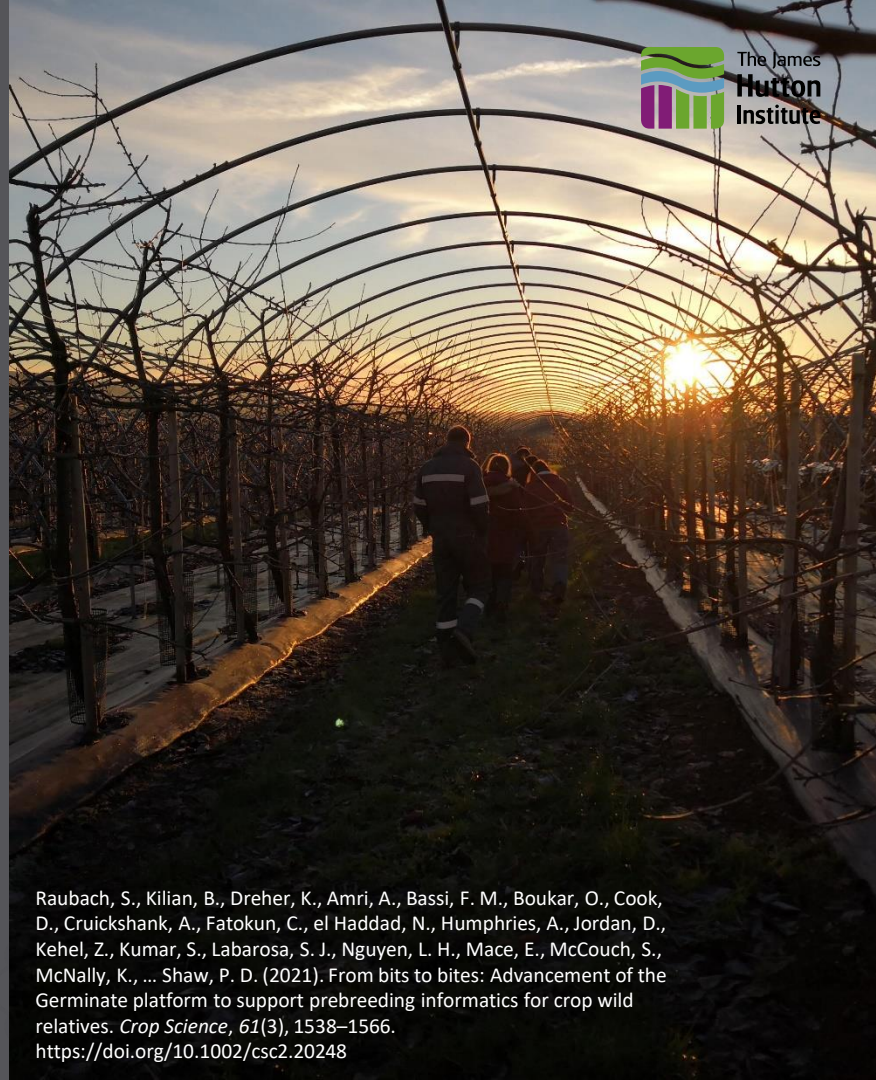
Humbug

Germinate Daim / Data templates / BrAPI

Helium

Summary

- All our software is free to use, most is open source.
- We will be adding a lot of soft fruit data under the EU H2020 BreedingValue Project
- Germinate is under active development and we welcome contributions, feedback and ideas on how we can improve.
- We are looking to develop better links to tools like Genesys, GGCE/GrinGlobal, EURISCO, Ensembl & GIGWA.
- Please just give us a shout if you want to try Germinate or want an extended demo.
- Contact myself or Sebastian Raubach for more information!



Raubach, S., Kilian, B., Dreher, K., Amri, A., Bassi, F. M., Boukar, O., Cook, D., Cruickshank, A., Fatokun, C., el Haddad, N., Humphries, A., Jordan, D., Kehel, Z., Kumar, S., Labarosa, S. J., Nguyen, L. H., Mace, E., McCouch, S., McNally, K., … Shaw, P. D. (2021). From bits to bites: Advancement of the Germinate platform to support prebreeding informatics for crop wild relatives. *Crop Science*, *61*(3), 1538–1566. https://doi.org/10.1002/csc2.20248

Acknowledgements

Hutton Barley, Soft Fruit and Potato Groups

- **Sebastian Raubach**
- Gaynor McKenzie
- Luke Ramsay
- Robbie Waugh
- Joanne Russell
- Malcolm Macaulay
- Ali Karley
- Niki McCallum
- Glenn Bryan
- Micha Bayer
- Miriam Schreiber
- Pauline Smith
- Derek Stewart
- Iain Milne
- Kelly Houston
- Susan McCallum
- Kaye Smith
- Julie Graham
- Amy Learmonth
- All our funders and partners!
- + Apologies to those who I may have missed.

@GerminateHub

https://ics.hutton.ac.uk/get-germinate

germinate@hutton.ac.uk