INTEGRATED DATABASE “GERMINATE”

THE DATA REPOSITORY OF THE PROJECT BREEDINGVALUE

14. FEB. 2024 | ELISA SENGER

IBG-4, FORSCHUNGSZENTRUM JUELICH, GERMANY

Member of the Helmholtz Association

OUTLOOK

- **Intro: Need for data repository**
- **The Germinate database**
- **Data preparation for upload**
- **BreedingValue berry data @ Germinate**

NEED FOR DATA REPOSITORY

Introduction

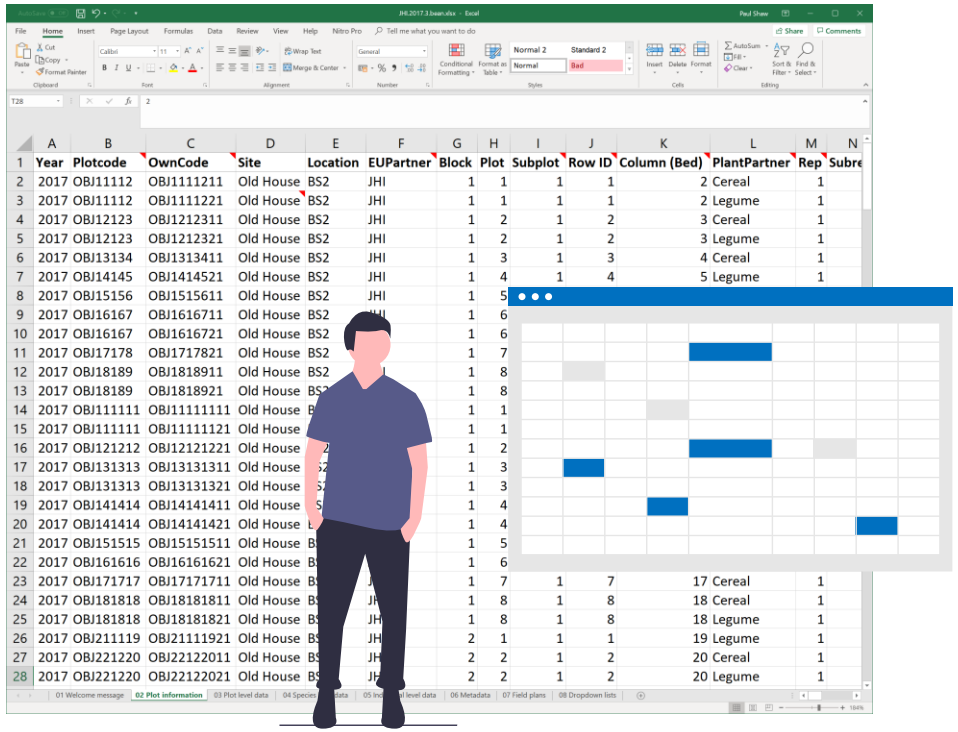
- FAIR[1] and open data
  - “desirable” or mandatory under public funding
  - increases value and utility of the data (repetitive and broader applications)
  - knowledge becomes more accessible for exploitation
- No common “plant database” present or limited for integrated[2], multi-omics data
- Databases host data (mid- to long-term)
- Databases provide tools to explore the data online

[1] Findable + Accessible + Interoperable + Re-useable
[2] data genres connected within database (e.g. genotypic, phenotypic, metabolomics, climate, passport)

GERMINATE (DEVELOPED BY JHI)

In a nutshell…

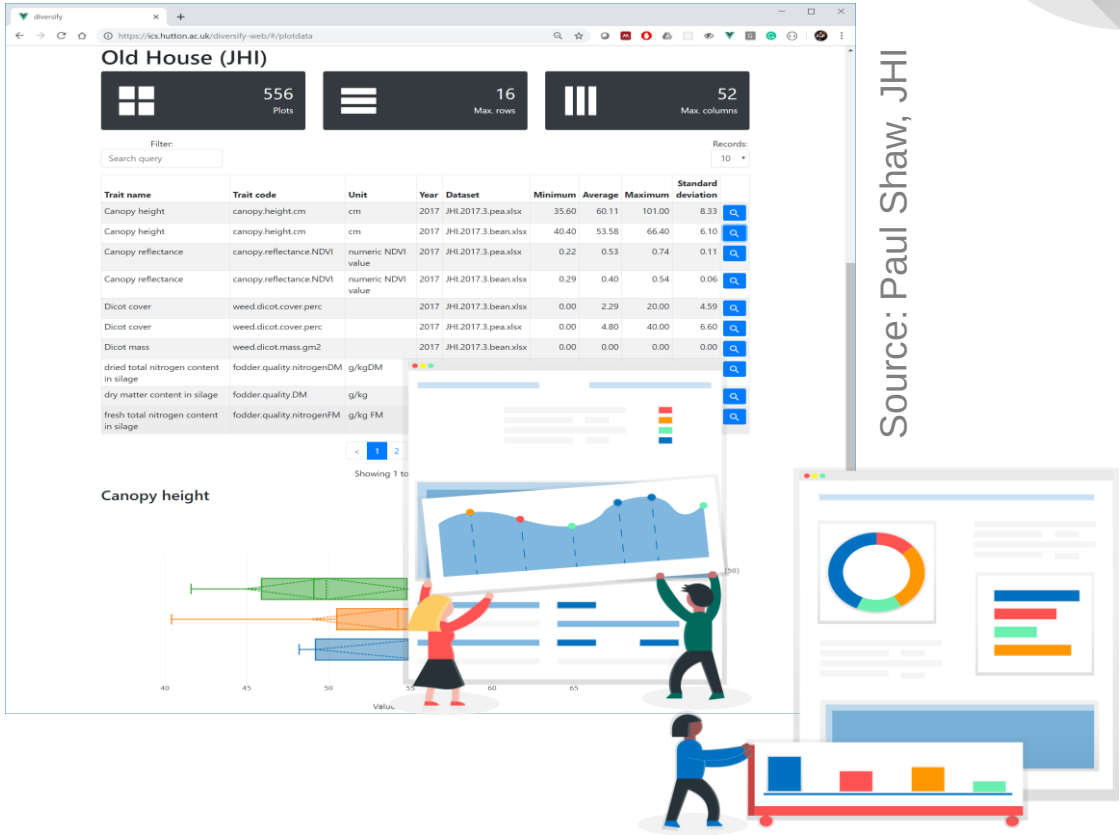
From this…



	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						
9	2017	OBJ16167	OBJ1616711	Old House	BS2		1	6						
10	2017	OBJ16167	OBJ1616721	Old House	BS2		1	6						
11	2017	OBJ17178	OBJ1717821	Old House	BS2		1	7						
12	2017	OBJ18189	OBJ1818911	Old House	BS2		1	8						
13	2017	OBJ18189	OBJ1818921	Old House	BS2		1	8						
14	2017	OBJ111111	OBJ11111111	Old House	B		1	1						
15	2017	OBJ111111	OBJ11111121	Old House			1	1						
16	2017	OBJ121212	OBJ12121221	Old House			1	2						
17	2017	OBJ131313	OBJ13131311	Old House	S2		1	3						
18	2017	OBJ131313	OBJ13131321	Old House	S2		1	3						
19	2017	OBJ141414	OBJ14141411	Old House	S		1	4						
20	2017	OBJ141414	OBJ14141421	Old House	L		1	4						
21	2017	OBJ151515	OBJ15151511	Old House	B		1	5						
22	2017	OBJ161616	OBJ16161621	Old House	B		1	6						
23	2017	OBJ171717	OBJ17171711	Old House	B		1	7	1	7	17	Cereal	1	
24	2017	OBJ181818	OBJ18181811	Old House	B		1	8	1	8	18	Cereal	1	
25	2017	OBJ181818	OBJ18181821	Old House	B		1	8	1	8	18	Legume	1	
26	2017	OBJ211119	OBJ21111921	Old House	B		2	1	1	1	19	Legume	1	
27	2017	OBJ221220	OBJ22122011	Old House	B		2	2	1	2	20	Cereal	1	
28	2017	OBJ221220	OBJ22122021	Old House	B		2	2	1	2	20	Legume	1	

Source: Paul Shaw, JHI

…to this



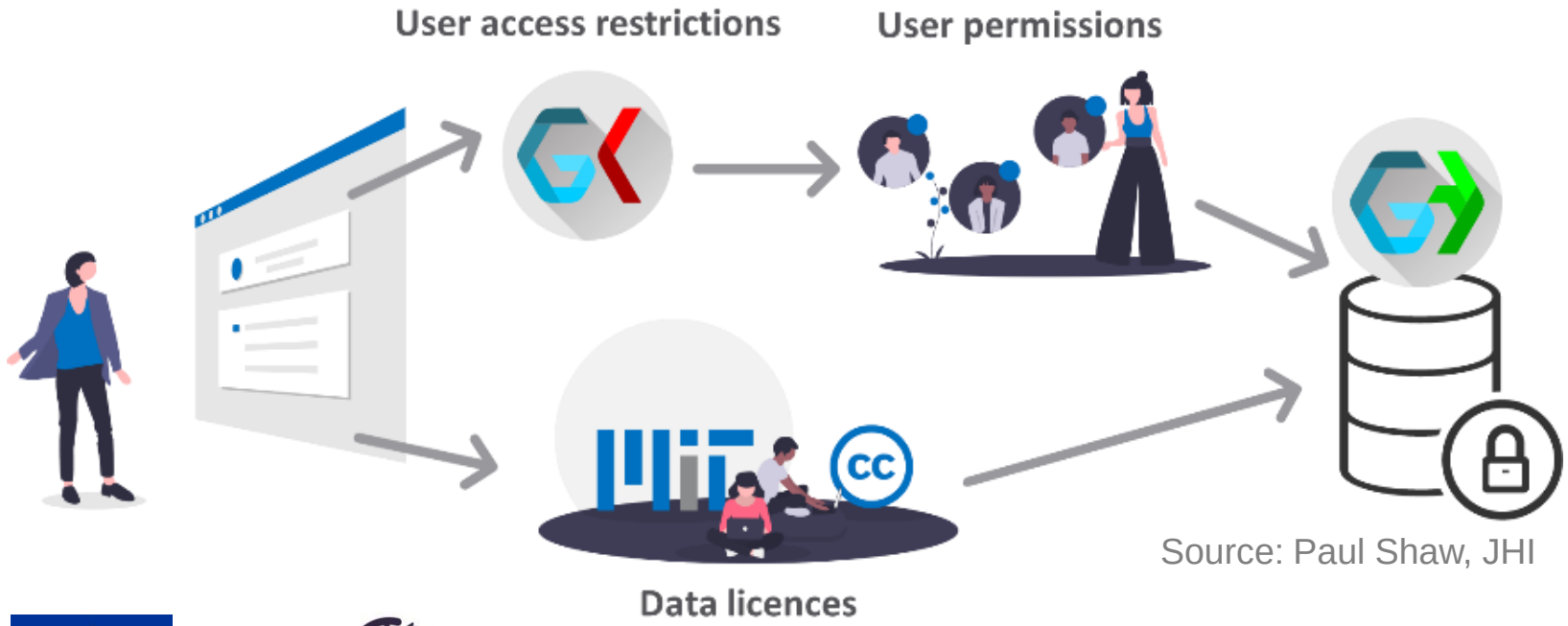
Source: Paul Shaw, JHI

GERMINATE

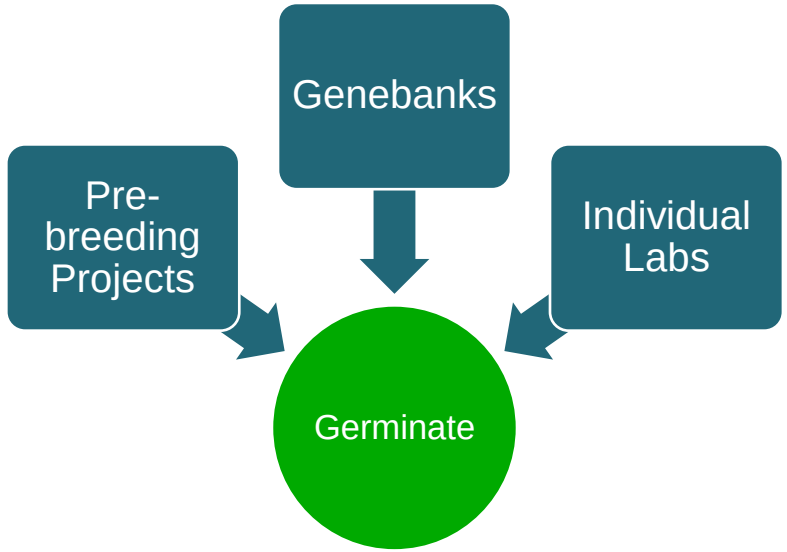
Overall objective

“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 the knowledge that is generated by each project is more accessible for exploitation.”

→ But we can keep your data private as well!



Source: Paul Shaw, JHI

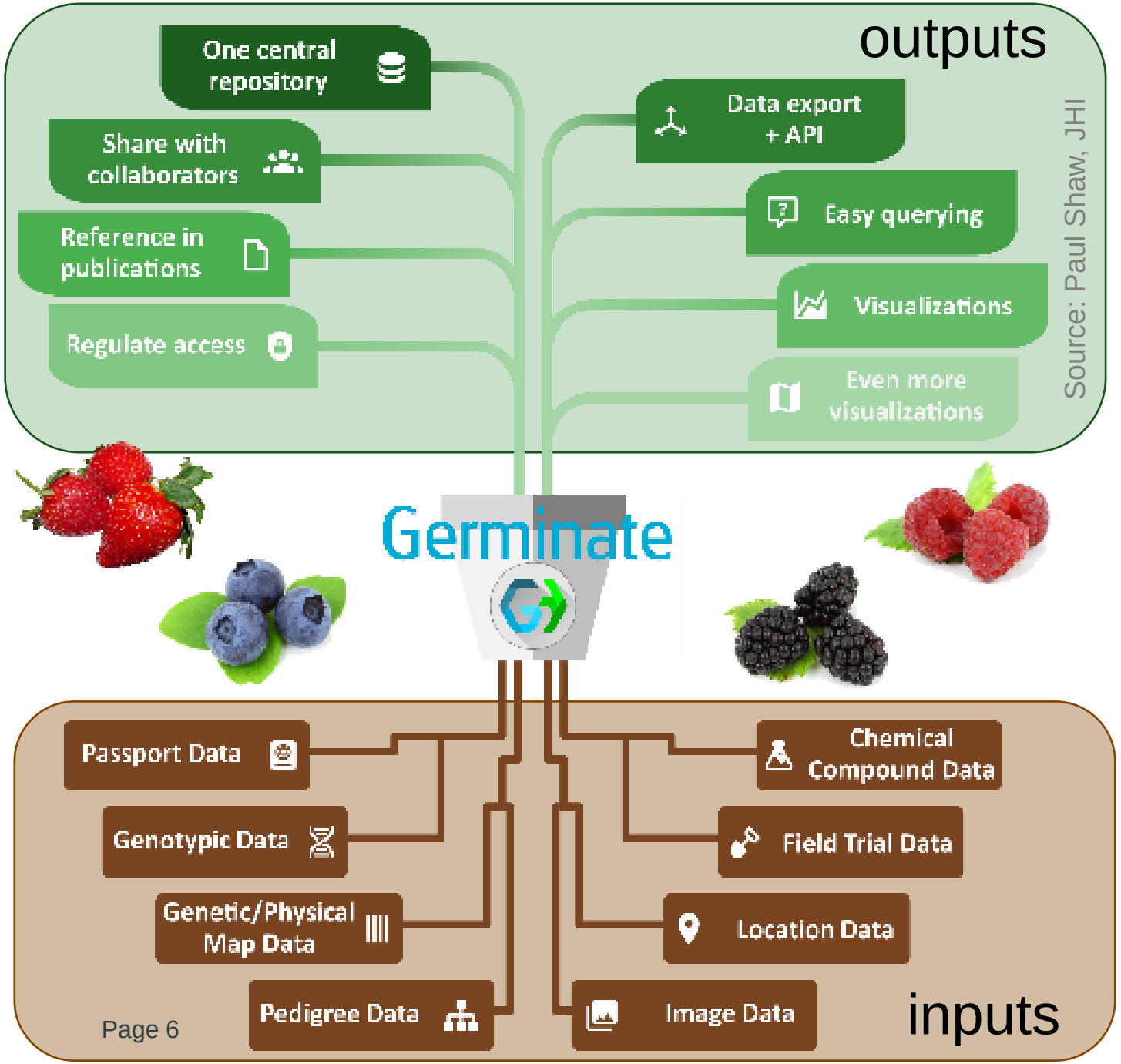


Source: Paul Shaw, JHI

GERMINATE

Data of plant genetic resources

- Connecting different data genre
- Visualizing + exploring
- Compatible with MCPD + Dublin Core
- BrAPI implementation currently being extended



Publications:

Shaw et al. (2017) Germinate 3
doi:10.2135/cropsci2016.09.0814

Raubach et al. (2020) From bits to bites
doi:10.1002/csc2.20248

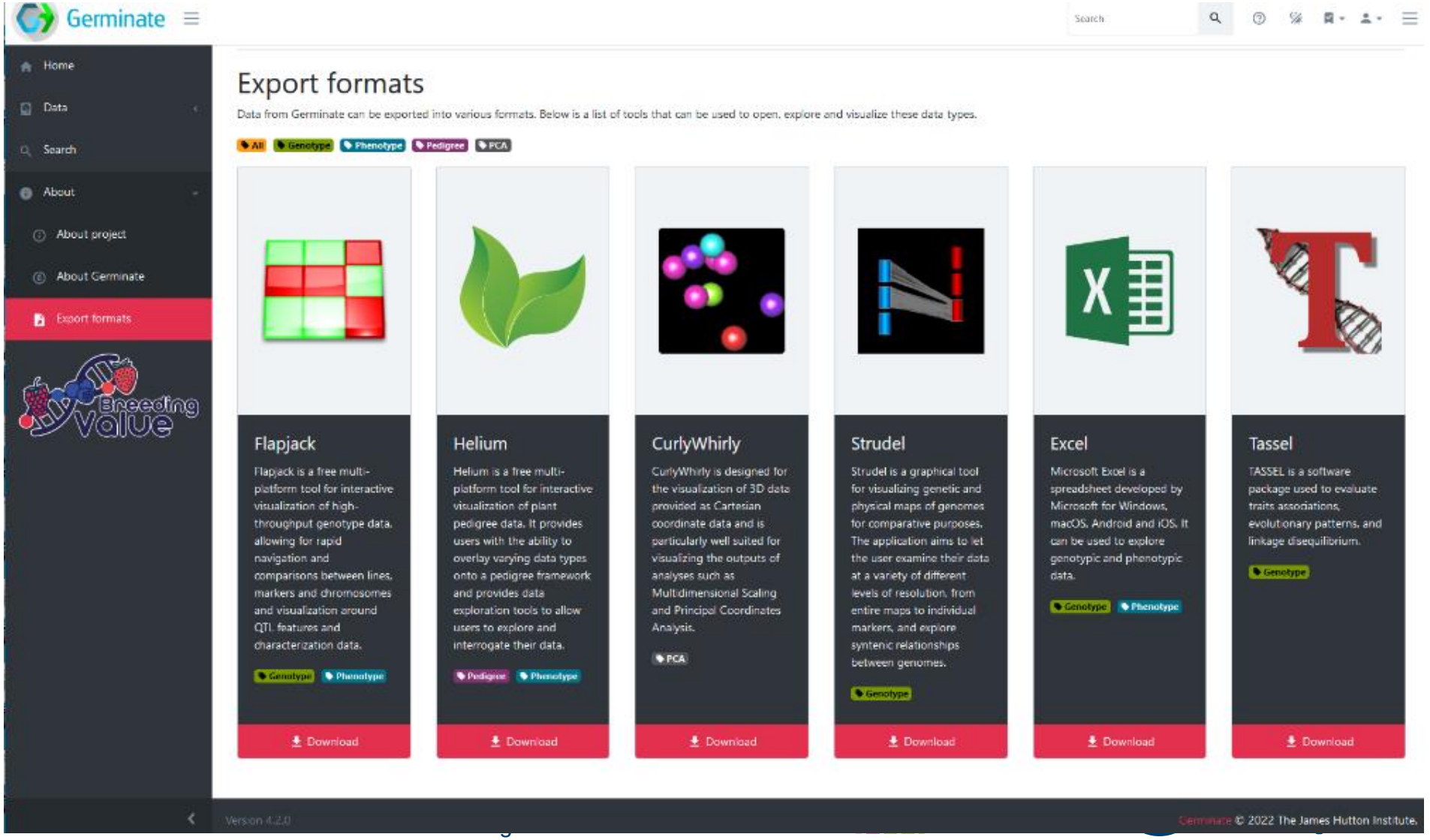
GERMINATE

Data of plant genetic resources

- Download + citation

https://germinateplatform.github.io/get-germinate/

https://ics.hutton.ac.uk/germinate-demo/#/home



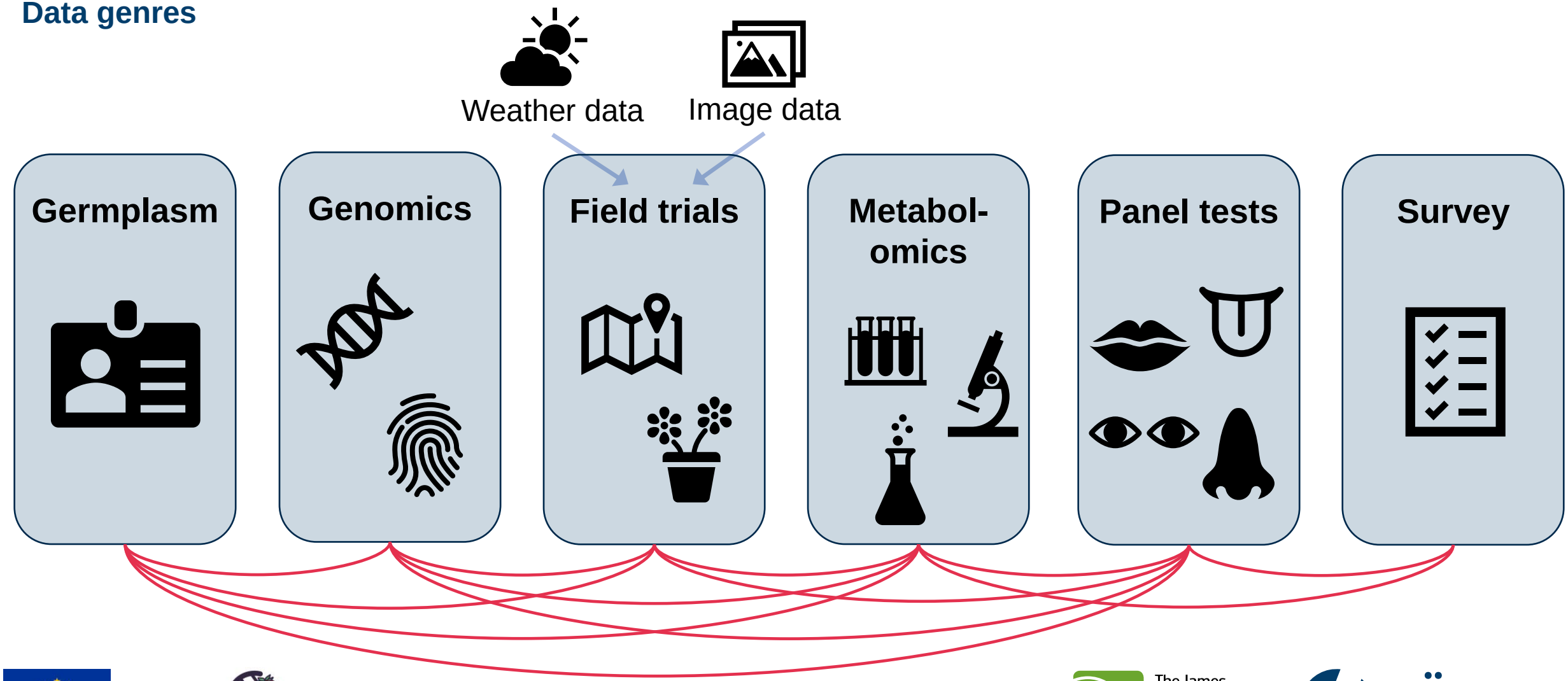
GERMINATE

Data of plant genetic resources

- Growing number or users
- Growing number of plant species (21 public databases + few private)
- Growing number of tools and functions

BREEDINGVALUE DATA

Data genres

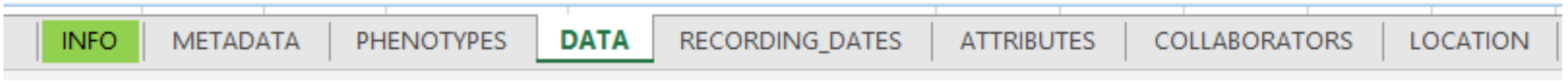


DATA PREPARATION FOR UPLOAD

Introduce strategy and template files to data providers

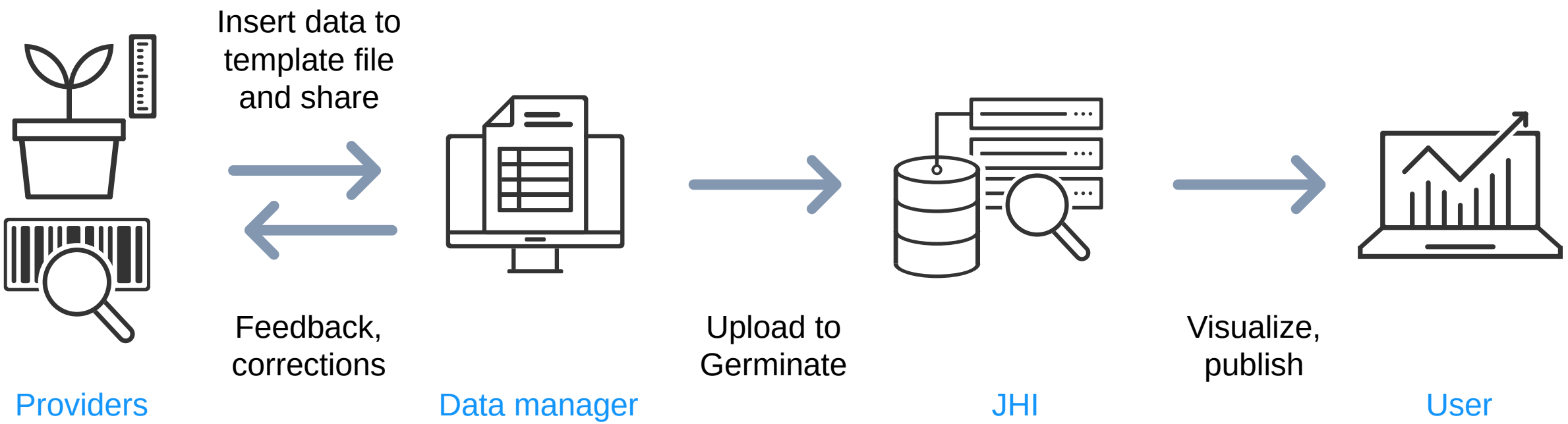
- Motivate, encourage, justify, clarify
- FAIR and open data (but requirements can be lowered and adapted)
- Share data with consortium and (ideally) with public

Name	Short Name	Description	Data Type	Unit Name	Unit Abbrev	Unit De	Trait categories (comma	Min (only fc	Max (only fc
Marketable Yield	MY	Marketable yie	numeric	grams per plar	g/plant			0	1000
Discarded Yield	DY	Discarded yielc	numeric	grams per plar	g/plant			0	1000
Average single fruit weight	AvSFW	Average fruit w	numeric	grams per frui	g/fruit			1	100
Balanced average single fruit weight	BalSFW	Calculated fron	numeric	grams per frui	g/fruit			1	100
Fruit size	FSI	Fruit size using	categorical				1: extremely small; 3: small; 5: medium; 7: large; 9: extrem		
Fruit size uniformity, Blueberry	FSIU_Vac	Fruit size unifo	categorical				1: extremely low; 3: low; 5: moderate; 7: high; 9: extremel		
Surface Color (CPVO 1-7), Strawberry	SCol_17	Surface color u	categorical				1: whitish yellow; 2: light orange; 3: medium orange; 4: ora		
Surface Color (CPVO 1-8), Raspberry	SCol_18	Surface color u	categorical				1: yellow; 2: orange; 3: light red; 4: medium red; 5: dark red		
Surface Color (Ctifl)	SCol_Ctifl	Surface color u	categorical				1; 2; 3; 4; 5; 6; 7; 8; 9		



DATA PREPARATION FOR UPLOAD

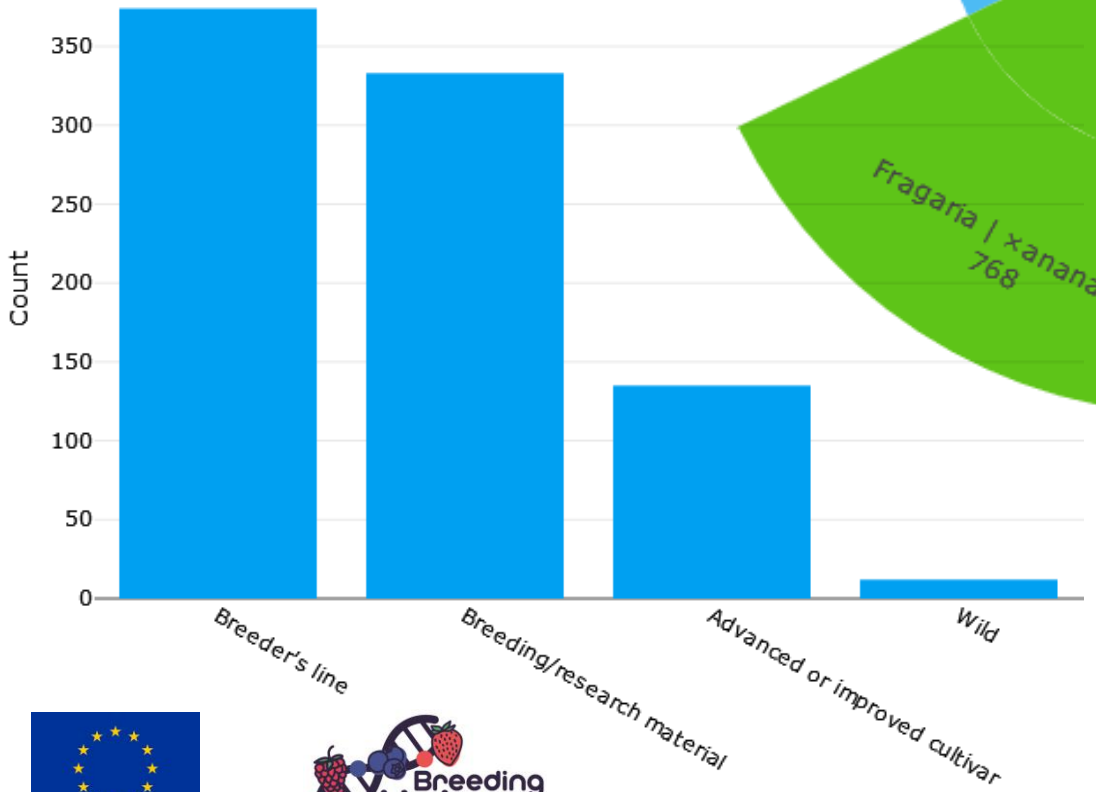
Data revision and upload

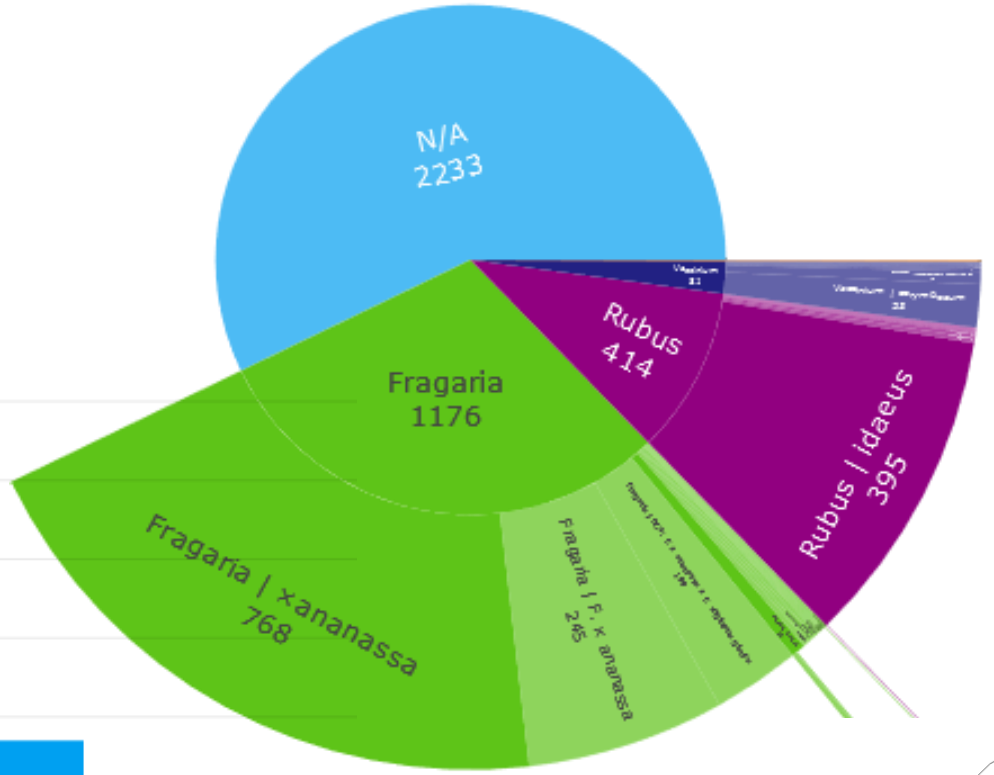


Source: Paul Shaw, JHI

BREEDINGVALUE BERRY DATA @ GERMINATE

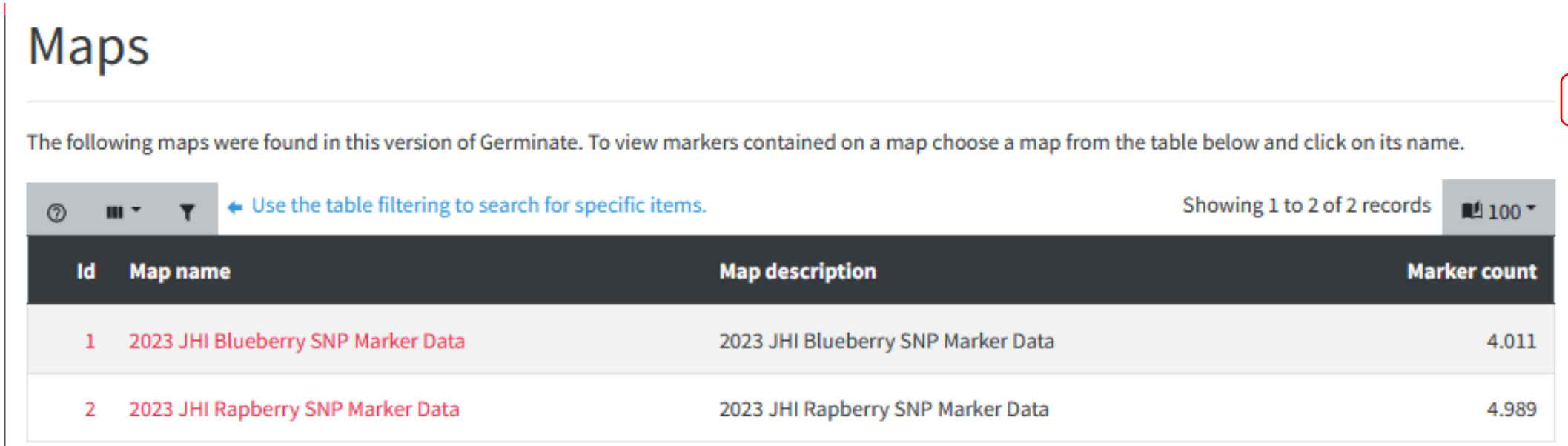
Passport data

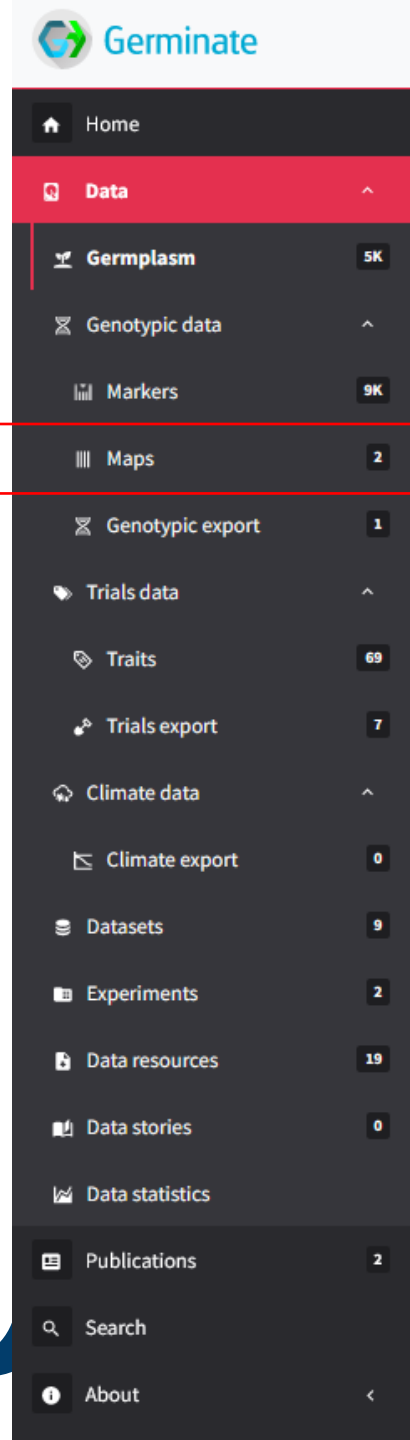




BREEDINGVALUE BERRY DATA @ GERMINATE

Genotypic data





BREEDINGVALUE BERRY DATA @ GERMINATE

Phenotypic data

Trials

Use the table filtering to search for specific items.

Showing 1 to 7 of 7 records

10

	Id	Dataset name	Dataset description	Experiment id	Experiment name	Dataset type	Data type	License	Data objects	Data points	Linked data resources	Ext.
☐	2	2020 JHI Blueberry LAB field mapping population Tunnel J8A	A blueberry cross used for genetic mapping	2	Generic Experiment	Trials		CC BY-SA	301	3.58K	Show	
☐	3	2021 JHI Blueberry LAB field mapping population Tunnel J8A	A blueberry cross used for genetic mapping	2	Generic Experiment	Trials		CC BY-SA	299	3.45K	Show	
☐	4	2022 JHI Blueberry LAB field mapping population Tunnel J8A	A blueberry cross used for genetic mapping	2	Generic Experiment	Trials		CC BY-SA	298	3.85K	Show	
☐	17	2022 JHI Blueberry Bullionfield cultivars	Replicated commercial blueberry cultivars	2	Generic Experiment	Trials		CC BY-SA	29	992	Show	
☐	18	2022 JHI Blueberry replicated mapping population	Mapping progeny blueberry cultivars	2	Generic Experiment	Trials		CC BY-SA	272	3.41K	Show	
☐	19	2023 JHI Blueberry Bullionfield cultivars	Replicated commercial blueberry cultivars	2	Generic Experiment	Trials		CC BY-SA	29	937	Show	
☐	21	2023 JHI Blueberry LAB field mapping population Tunnel J8A	A blueberry cross used for genetic mapping	2	Generic Experiment	Trials		CC BY-SA	293	3.78K	Show	

Use the checkboxes to select one or more rows.

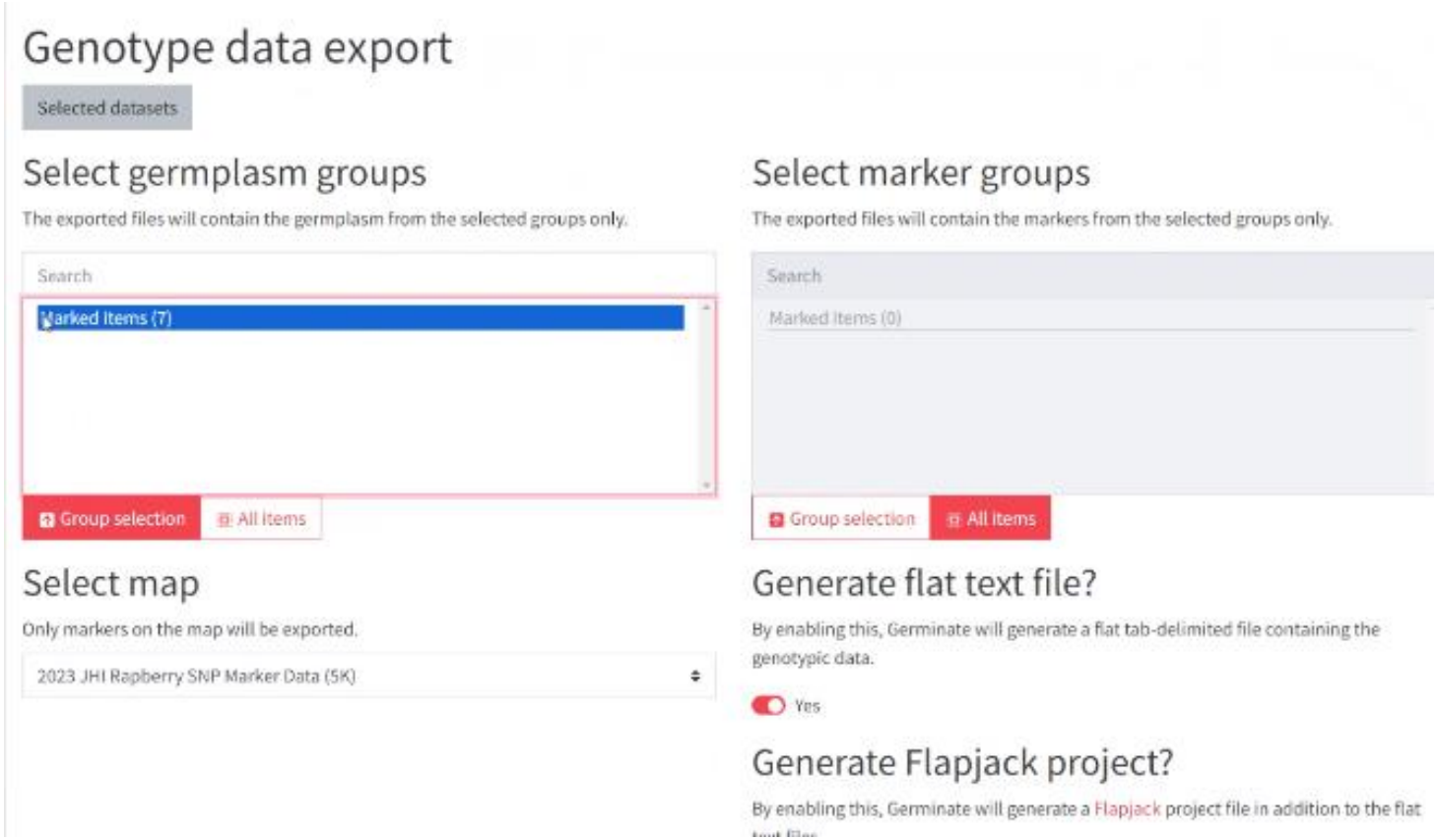
Next

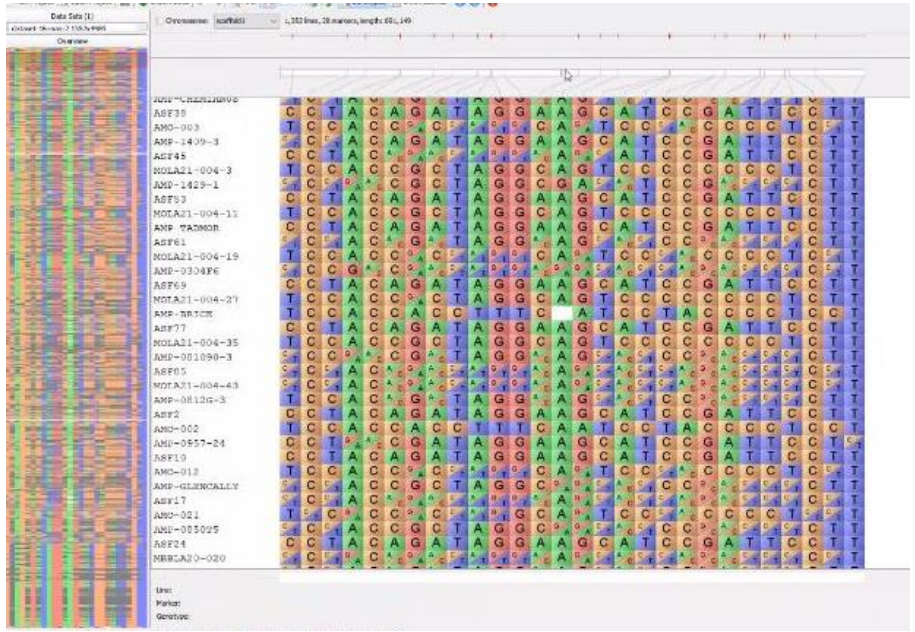
Germinate

- Home
- Data
  - Germplasm 5K
  - Genotypic data
    - Markers 9K
    - Maps 2
    - Genotypic export 1
  - Trials data
    - Traits 69
    - Trials export 7
  - Climate data
    - Climate export 0
  - Datasets 9
  - Experiments 2
  - Data resources 19
  - Data stories 0
  - Data statistics
- Publications 2
- Search
- About

BREEDINGVALUE BERRY DATA @ GERMINATE

Example for data export

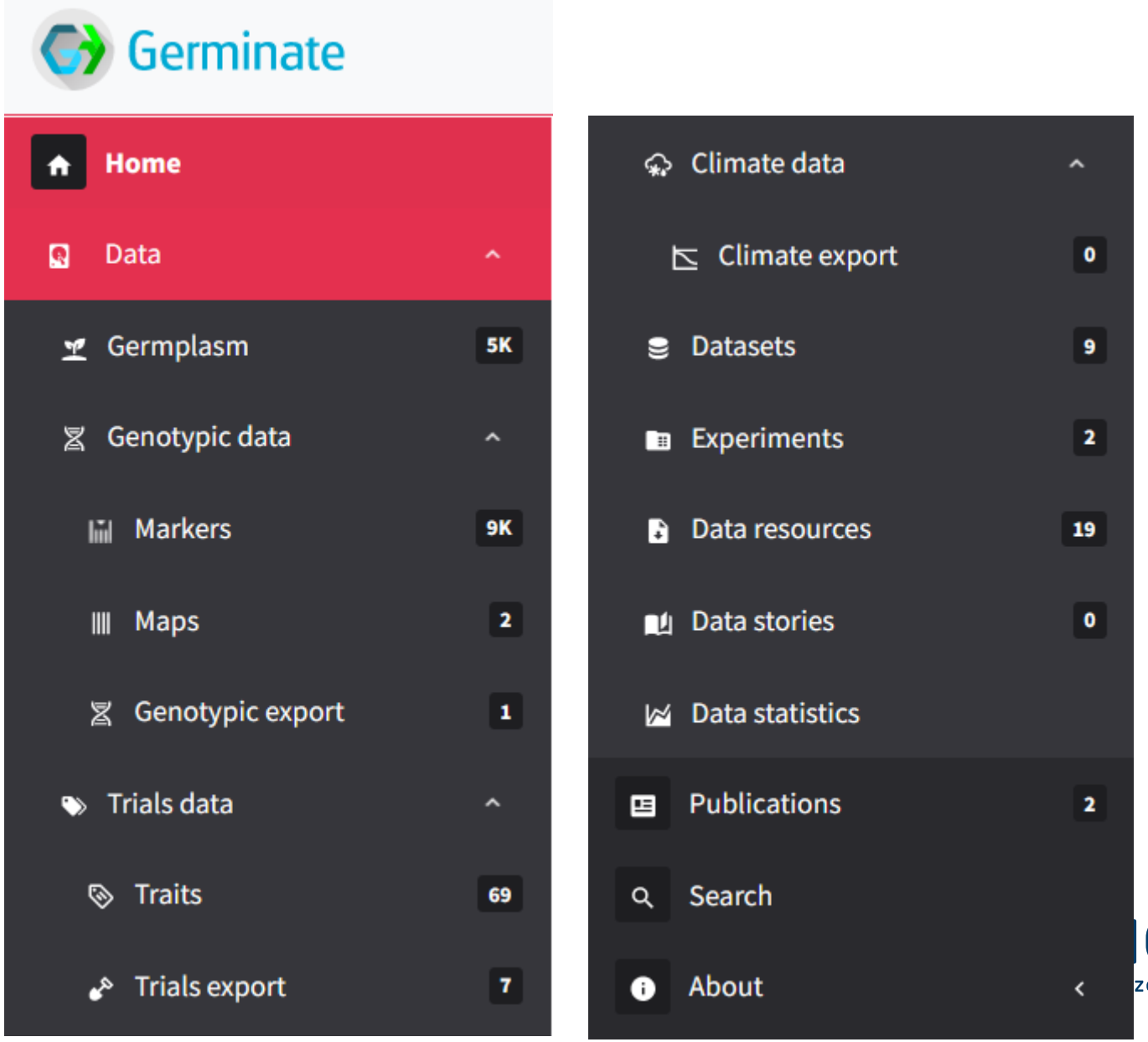




BREEDINGVALUE BERRY DATA @ GERMINATE

Further data to come

- Data, data, data…
- Photographs of plants and fruits
- Metabolic profiles
- Location
- Climate



THANKS TO

→ **YOU !** ...FOR YOUR ATTENTION

→ **JAMES-HUTTON INSTITUTE** ...FOR PROVIDING THE PLATFORM AND SOME SLIDES (PAUL SHAW, SEBASTIAN RAUBACH, DEPARTMENT OF INFORMATION AND COMPUTATIONAL SCIENCES, SUSAN MCCALLUM, DEPARTMENT OF CELL AND MOLECULAR SCIENCES, INVERGROWIE, SCTOLAND, UK

→ **BREEDINGVALUE PARTNERS** ...FOR COLLABORATION AND PROVIDING DATA

→ **EU** ...FOR FUNDING

→ **IBG-4 AT FZJ** ...FOR TRAINING AND COMFORTABLE WORKING SPACE