EURISCO progress report
(April 2014 – April 2018)

15th Steering Committee meeting, 15–17 May 2018, Thessaloniki, Greece
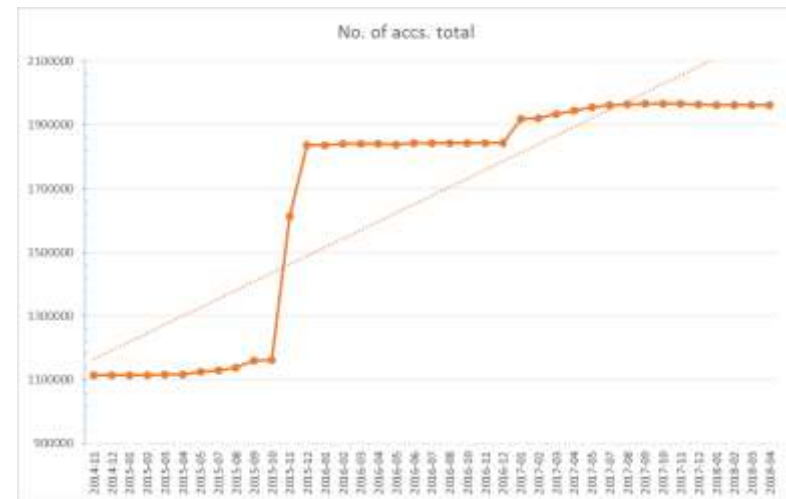
Stephan Weise

CURRENT STATUS OF EURISCO

Contents of EURISCO

- 1,961,985 accessions

- 6,317 genera
(including synonyms, spelling variants)

- 42,974 species names
(unique combinations genus + species, including synonyms)

- 418,679 MLS accessions

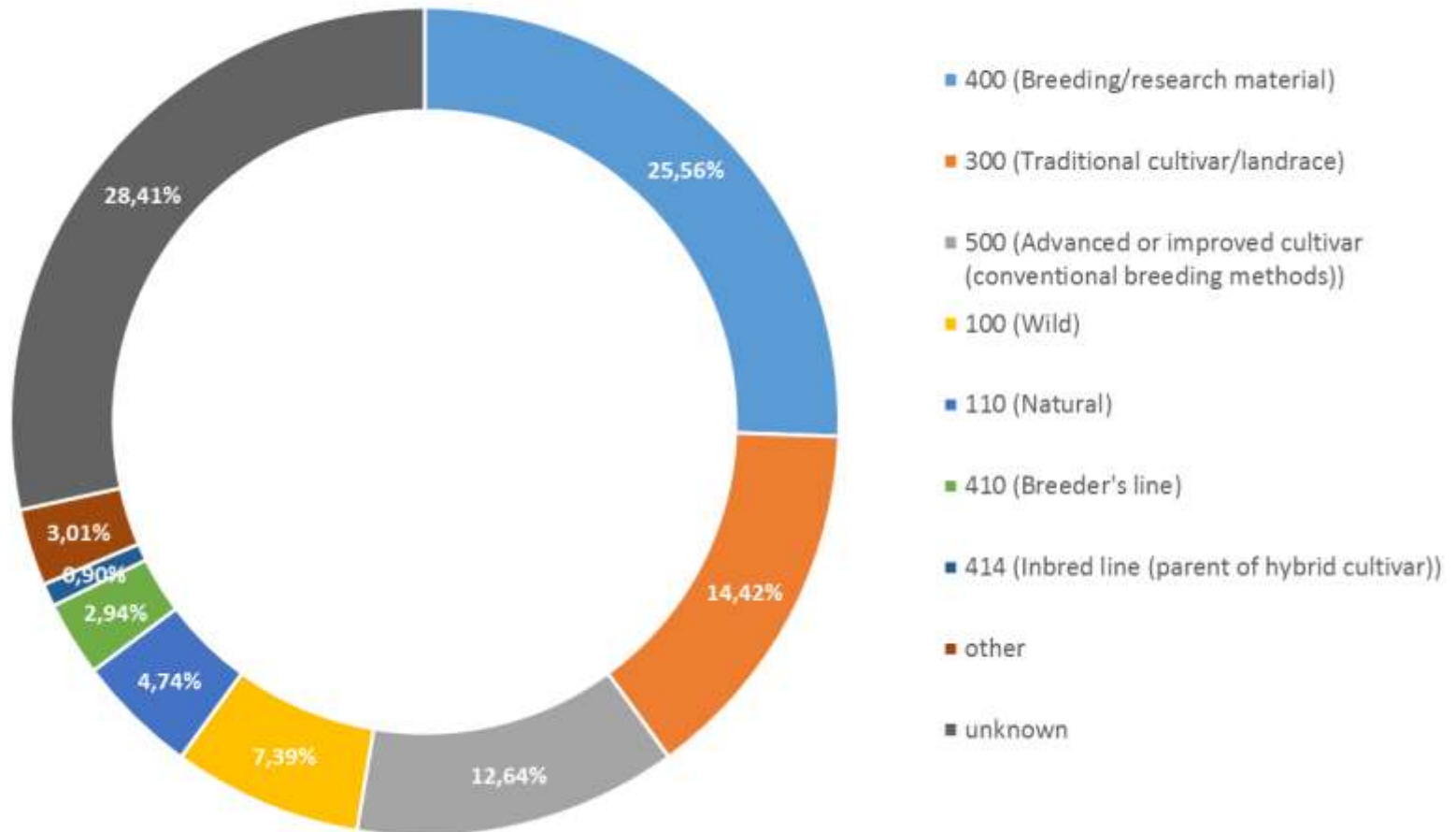- 34,364 AEGIS accessions

- 22,906 DOIs

as at 2018-04-25

Taxonomic composition

Genus	Species	No. accs.	Total
Arabidopsis	*thaliana*	681,983	682,191
	others	208	
Triticum (wheat)	*aestivum*	136,778	189,104
	durum	16,312	
	turgidum	9,601	
	monococcum	3,483	
	spelta	3,202	
	others	19,728	
Hordeum (barley)	*vulgare*	104,374	121,589
	spontaneum	5,897	
	others	11,318	
Zea (maize)	*mays*	61,715	61,849
	others	134	
Phaseolus (garden bean)	*vulgaris*	46,621	52,179
	coccineus	3,064	
	others	2,494	

Genus	Species	No. accs.	Total
Solanum (tomato, potato, eggplant, etc.)	*lycopersicum*	19,152	51,265
	tuberosum	14,591	
	andigenum	2,814	
	melongena	2,119	
	others	12,589	
Avena (oat)	*sativa*	33,274	41,299
	sterilis	2,152	
	byzantina	1,968	
	others	3,905	
Malus (apple)	*domestica*	24,759	31,875
	others	7,116	
Pisum (pea)	*sativum*	27,417	30,455
	others	3,038	
Vitis (grape)	*vinifera*	26,096	30,049
	others	3,953	
others			670,130
		Total	1,961,985

as at 2018-04-25

Biological status

- 400 (Breeding/research material)
- 300 (Traditional cultivar/landrace)
- 500 (Advanced or improved cultivar (conventional breeding methods))
- 100 (Wild)
- 110 (Natural)
- 410 (Breeder's line)
- 414 (Inbred line (parent of hybrid cultivar))
- other
- unknown

25,56%
14,42%
12,64%
7,39%
4,74%
2,94%
0,90%
3,01%
28,41%

as at 2018-04-25

C&E data

- Currently, 1,652,895 records of data from seven countries

  – Czech Republic

  – Germany

  – Latvia

  – The Netherlands

  – Poland

  – Romania

  – United Kingdom

- 68,821 accs. with C&E data

Legend:
- Triticum — 26,07%
- Solanum — 11,09%
- Hordeum — 7,76%
- Pisum — 6,71%
- Lactuca — 6,48%
- Capsicum — 3,98%
- Brassica — 3,27%
- Lolium — 2,62%
- Avena — 2,55%
- Phaseolus — 2,45%
- other — 27,00%
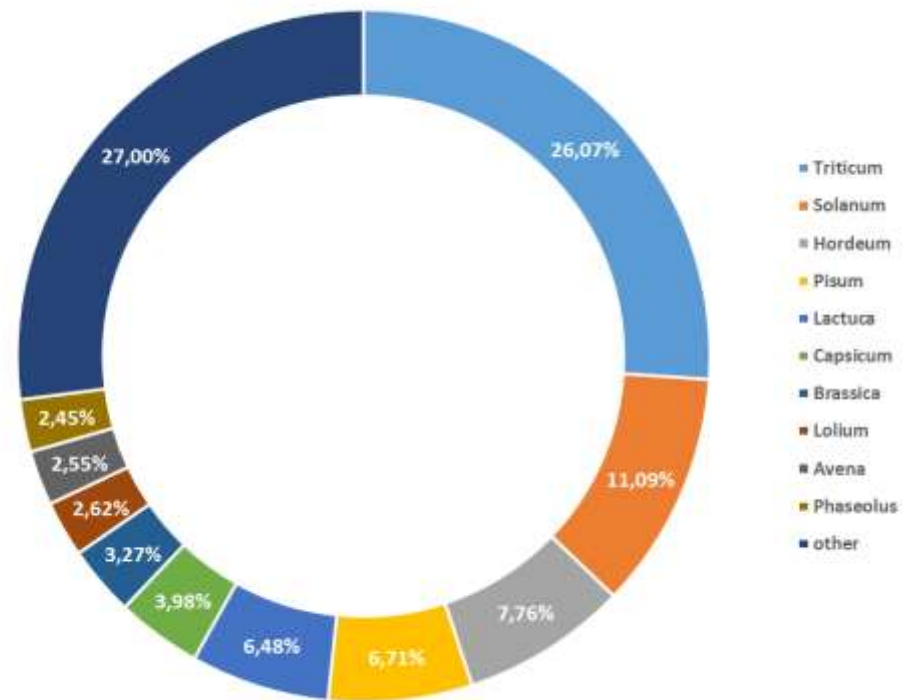
as at 2018-04-25

Architecture: Overview

EURISCO intranet

EURISCO web

EURISCO format

Staging area

Integrity checks

EURISCO

EURISCO

Synchronisation
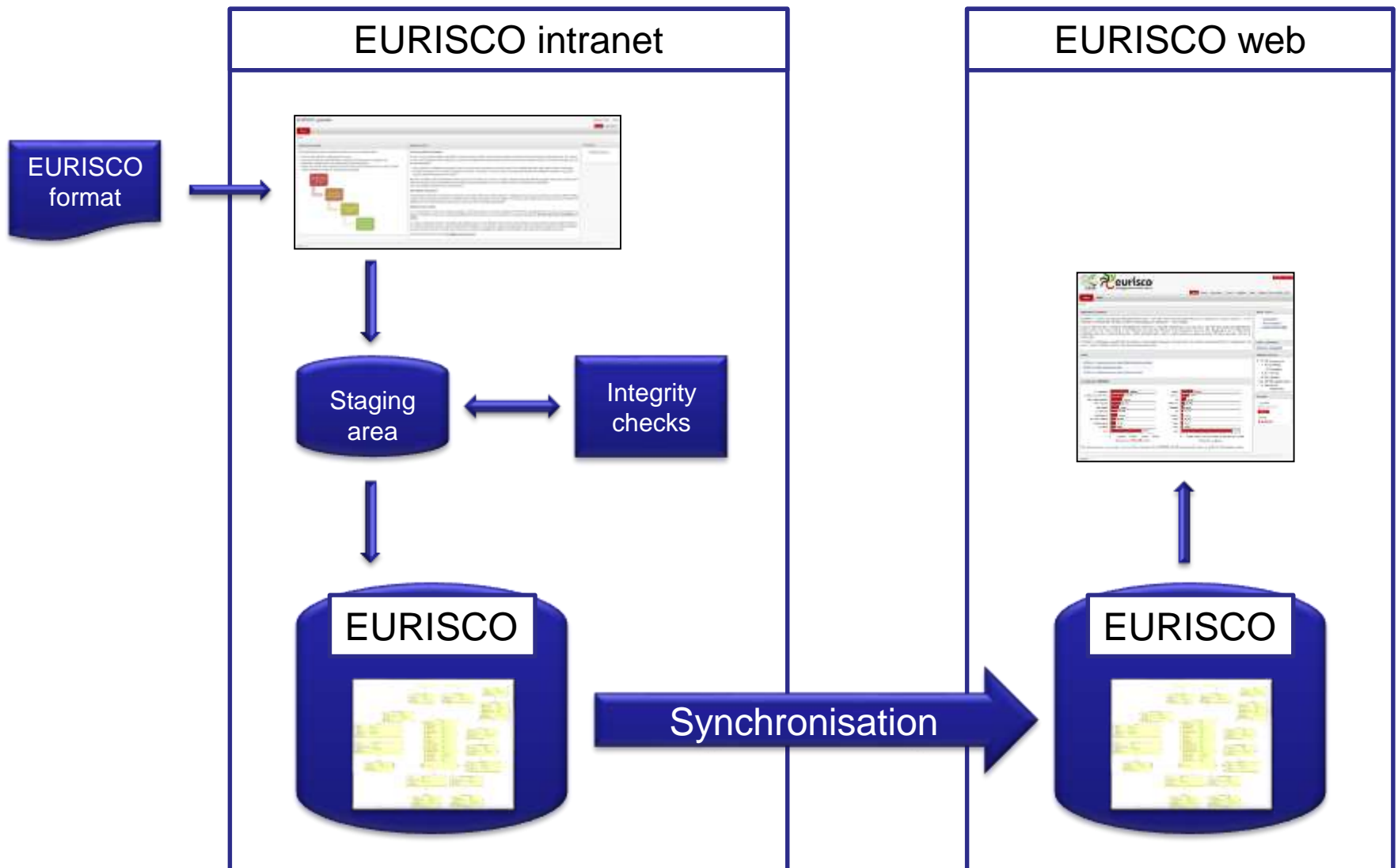
Architecture: DB schema features

- EURISCO intranet

  - 59 tables

  - 502 indexes

  - 108 triggers

  - 14 PL/SQL packages

  - 177 function and procedures
    - data upload and import
    - integrity checks
    - updates (passport and C&E)
    - taxonomy support

  - 27 Java classes

- EURISCO web

  - 46 tables

  - 29 materialised views

  - 698 indexes

  - 9 PL/SQL packages

  - 38 function and procedures
    - download
    - newsletter
    - statistics
    - C&E data visualisation
    - AEGIS status auditing
    - taxonomy support

  - In-memory features

  - 4 Java classes

Web interface

39 (sub)versions since 2014

LEIBNIZ INSTITUTE OF PLANT GENETICS AND CROP PLANT RESEARCH

Passport data in EURISCO

- Four standard searches:

  - Taxonomy

  - Accession

  - Biological status

  - Collecting site

- Advanced search
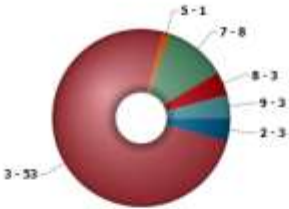
- Different user-specific export features

C&E data in EURISCO

Wizard-based searches for
- Genus
- Species and trait
- Experiment
- Trait

Version history of the web interface

v1.0.0 (Oct. 2014)
- First public version
- v1.0.1 – v1.0.6 continuous improvements

v1.1.0 (Nov. 2014)
- New export functionality + download of full dump
- v1.1.1 – v1.1.17 (2015 – 2016) continuous improvements

v1.2.0 (Jun. 2016)
- C&E data extension; new export functionalities; new advanced search; lots of small improvements
- v1.2.1 – v1.2.7 (2016 – 2017) continuous improvements

v1.3.0 (Dec. 2017)
- Migration to MCPD2; increased usability; lots of small improvements
- v1.3.1 – v1.3.5 (2018) continuous improvements

v1.4.0 (planned for Jun. 2018)
- Taxonomy search completely reworked

NEW HOSTING OF EURISCO

EURISCO transfer

- Preparation of transfer (2013)
  - Preparatory meetings
    - High costs for transfer of as-is status
      → New development from scratch
  - Analysis of the former web application

- System reengineering and data migration
  - Frontend reengineering (public EURISCO application)
    - Database structures and public web application
    - PL/SQL for functionality; APEX for rendering
  - Backend reengineering (EURISCO intranet)
    - Database structures and intranet web application
    - PL/SQL for functionality; APEX for rendering
  - Import (and cleansing) of current data set
    - Migration path MySQL → Oracle RDBMS
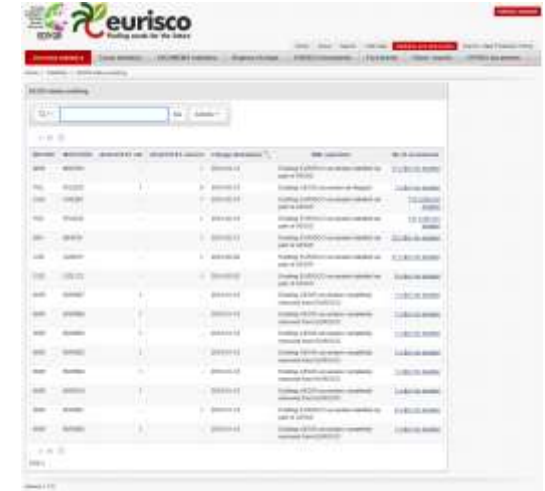
Outcomes 2014

- EURISCO intranet

    – EURISCO transfer (main task)

    – Switch to new upload mechanism (incremental updates)

- Public EURISCO application

    – EURISCO transfer (main task)

    – New export feature with customised export files in CSV format

Outcomes 2015

- EURISCO intranet
  - Java-based passport importer developed
  - Data quality/integrity checks improved
  - AEGIS status tracking

- Public EURISCO application
  - Statistics section with various reports
  - Improved search functionality (additional fields/filters)
  - Adjustments "under the hood"

- Extension for C&E data
  - Data exchange format
  - Database schema
  - Upload and check mechanism
  - Public web interface (prototype; continued in 2016)

Outcomes 2016

- EURISCO intranet
  - Interface extended for importing C&E data
  - AEGIS support as well as quality/integrity checks improved

- Public EURISCO application
  - Additional export functionalities implemented
  - Advanced search completely reworked
  - Extended performance tuning

- Extension for C&E data
  - Public web interface extended for C&E and tested
  - Additional features implemented (based on users' feedback)
  - Productive data imported
  - Official release of C&E extension

Outcomes 2017

- EURISCO intranet

  - AEGIS support improved further

  - Improvement of taxonomy support started (continued in 2018)

    - Database structures and mapping procedures

  - How-to-documents (handbooks) compiled

  - Implementation of MCPD2 for EURISCO

    - Database structures, upload and check procedures

- Public EURISCO application

  - User survey conducted and requirements for improvement collected → addressed all NCs/NFPs; received useful suggestions

- Extension for C&E data

  - Additional flexible procedure for uploading C&E data compiled and implemented → delegation and final approval by NFPs

Outcomes 2018 (ongoing)

- EURISCO intranet

  – Improvement of taxonomy support (started in 2017)

- Public EURISCO application

  – Implementation of requirements from user survey

  – Extension for SSR data (under evaluation)

- Extension for C&E data

  – Improvement of search and download features

Support of ECPGR Central Crop Databases

- Cooperation with various crop WGs

- Example: European *Poa* Database (Forage WG)

- poa.ipk-gatersleben.de (hosted at IPK since the 1990s)

- Blueprint for a crop portal automatically updated from EURISCO
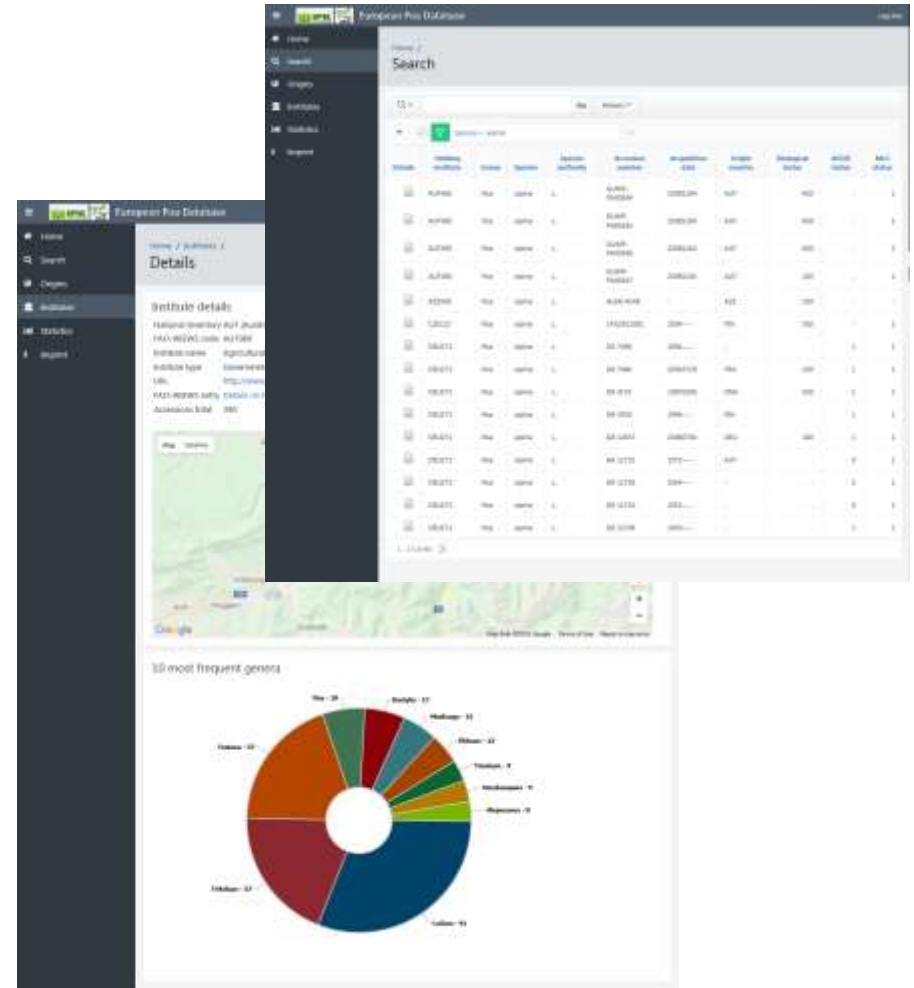
EURISCO COORDINATION

Network maintenance + development

- Contact with EURISCO stakeholders

- Definition of new services, e.g. with regard to DOIs

- Advancement of current and definition of new standards

- Coordination with initiatives such as Genesys and GLIS

- Bilateral communication with regard to the coverage of EURISCO

- Cooperation with ECPGR Working Groups

- Preparation of work plans and reports

- Helpdesk behind the scenes (should not be underestimated ;-))

Passport data updates

Year	No. of updates	Accs. total
2014	1	1,114,995
2015	28	1,837,368
2016	25	1,842,539
2017	55	1,964,062
2018 (as at 2018-04-25)	12	1,961,985

EURISCO training workshops

- Regular training for data providers (esp. for C&E data and in case of staff changes)

- Training materials + test environment developed + meeting reports

- Annual workshops:
  - 2015 Tirana, Albania (Agricultural University of Tirana)
    - 20 participants
  - 2016 Angers, France (INRA)
    - 13 participants
    - Meeting of the EURISCO Advisory Committee
  - 2017 Gatersleben, Germany (IPK)
    - 14 participants

- 2018 training under preparation

Participation in project proposals

- Involved in various ECPGR Grant Scheme Activities

- EURISCO Taxonomy (ECPGR, BLE), 2017–2018
  - Improvement of taxonomic backbone of EURISCO

- EUCLEG (Horizon 2020), 2017–2021
  - Coordination: INRA, France
  - Aim: reduction of protein import dependencies of both European and Chinese partners
  - Leader of work package for data management
    - Find data gaps in EURISCO (and try to close them)
    - Manage project data (passport, phenotypic, genetic)

- Farmer's Pride (Horizon 2020), 2018–2020
  - Coordination: University of Birmingham, U.K.
  - Aim: Development of network of *in situ* sites and stakeholders
  - Task leader: Preparation of a concept to extend EURISCO for *in situ* data

- Participation in the proposal "GenRes Bridge – Joining forces for genetic resources and biodiversity management"

Dissemination

- Regularly short information in ECPGR bulletin

- EURISCO newsletter twice a year

- Journal article

- EURISCO talks and posters at several conferences

- Presentations at several ECPGR workshops

- ECPGR Grant Scheme Activities

FUTURE OF EURISCO

EURISCO: Plans & challenges

- ECPGR Phase X (2019–2023)

  - Further development in close collaboration with ECPGR bodies

  - Specific focus on:

    - Phenotypic data

      - Harmonisation of traits

      - Implementation of additional use-cases

    - *In situ* data

      - Extension of EURISCO

    - Data quality

      - Completeness

      - Reliability, e.g. geographic coordinates

  - Continuous improvement of functions and services

THANK YOU FOR YOUR ATTENTION