WP5 Q-COLLECT INFORMATION PORTAL

Vincent Robert, Francoise Petter, Damien Griessinger, Nathalie van de Wiele

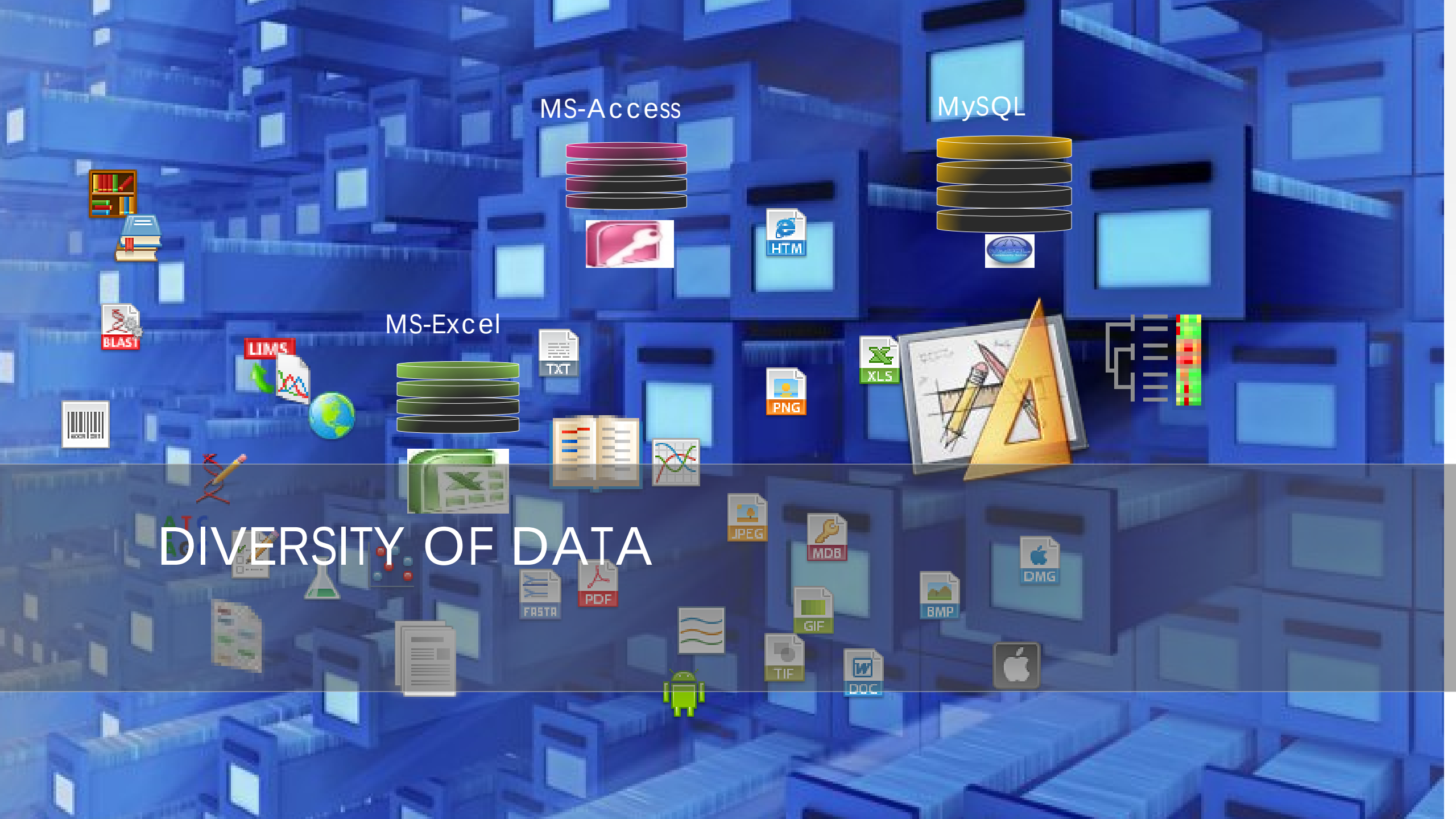
WP5 Q-COLLECT INFORMATION PORTAL

Work Package 5. Info-portal

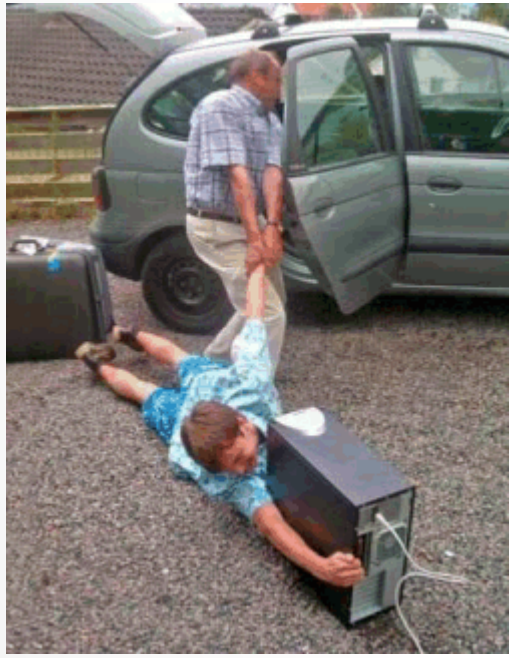
5.1. To make an inventory of all existing tools, websites, databases that might be complementary to the inventory on collections (WP2).

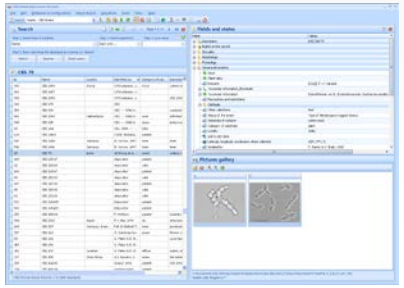
5.2. From other WPs, obtain the information on the existing systems, strengths and missing features that might need to be implemented within this project and associated databases. Summarize the wishes of the different WPs and make a proposal how future information should be made accessible for the end-users.

5.3. To create an integrated system (info-portal) hosting the information and end-results generated by all other WPs.



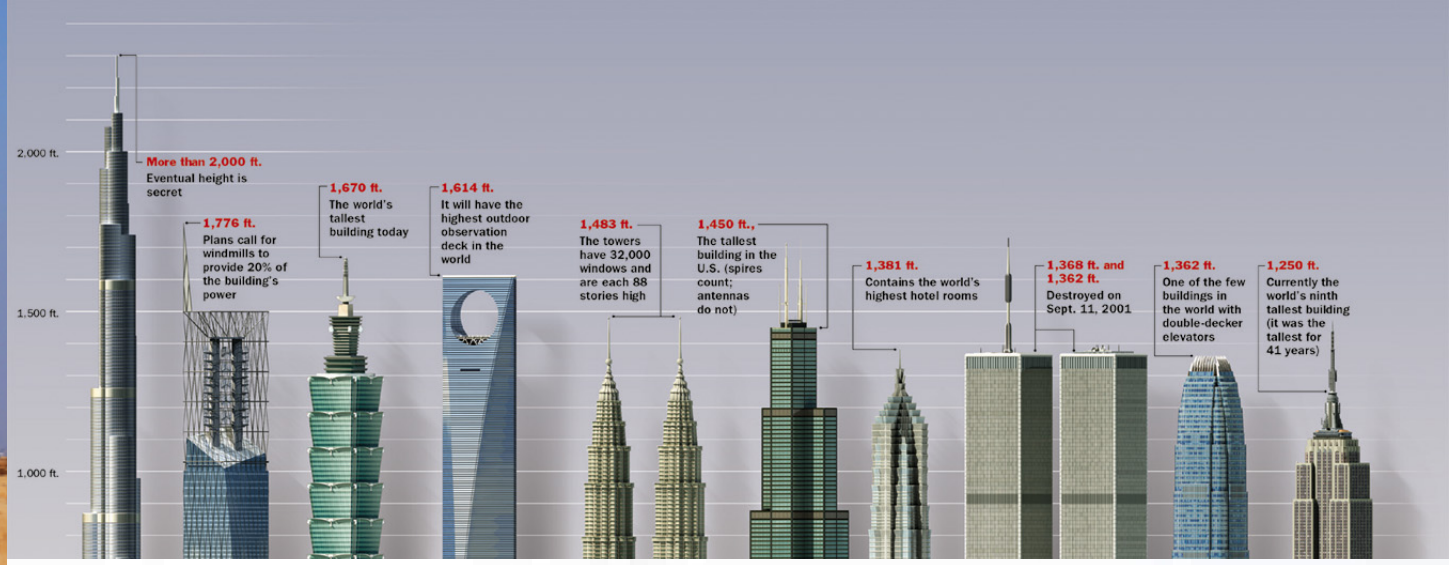
DIVERSITY OF DATA





DIVERSITY OF PEOPLE

DIVERSITY OF RESOURCES & DIGITALIZATION LEVELS



USE THE RIGHT FOUNDATIONS






STEP BY STEP WITH A PLAN/VISION

WORLD IS CHANGING

COLLECTIONS ISLANDS

Disconnected databases
Only1/10000 strains in scientific publications are in collections

BRIDGING SYSTEMS & COLLECTIONS

Bridging Sciences


The Future
NEXT EXIT

1. Manage collection’s data using web based applications
2. Manage collection’s data using desktop applications
3. Operating system for desktop application: Windows 82%, Linux 5%, Apple 9.6% but Mobiles are rising fast
4. Create management software using in-house resources
5. Use existing open-source or free software
6. Use existing commercial software
7. Databases
8. Centrally hosted infrastructure

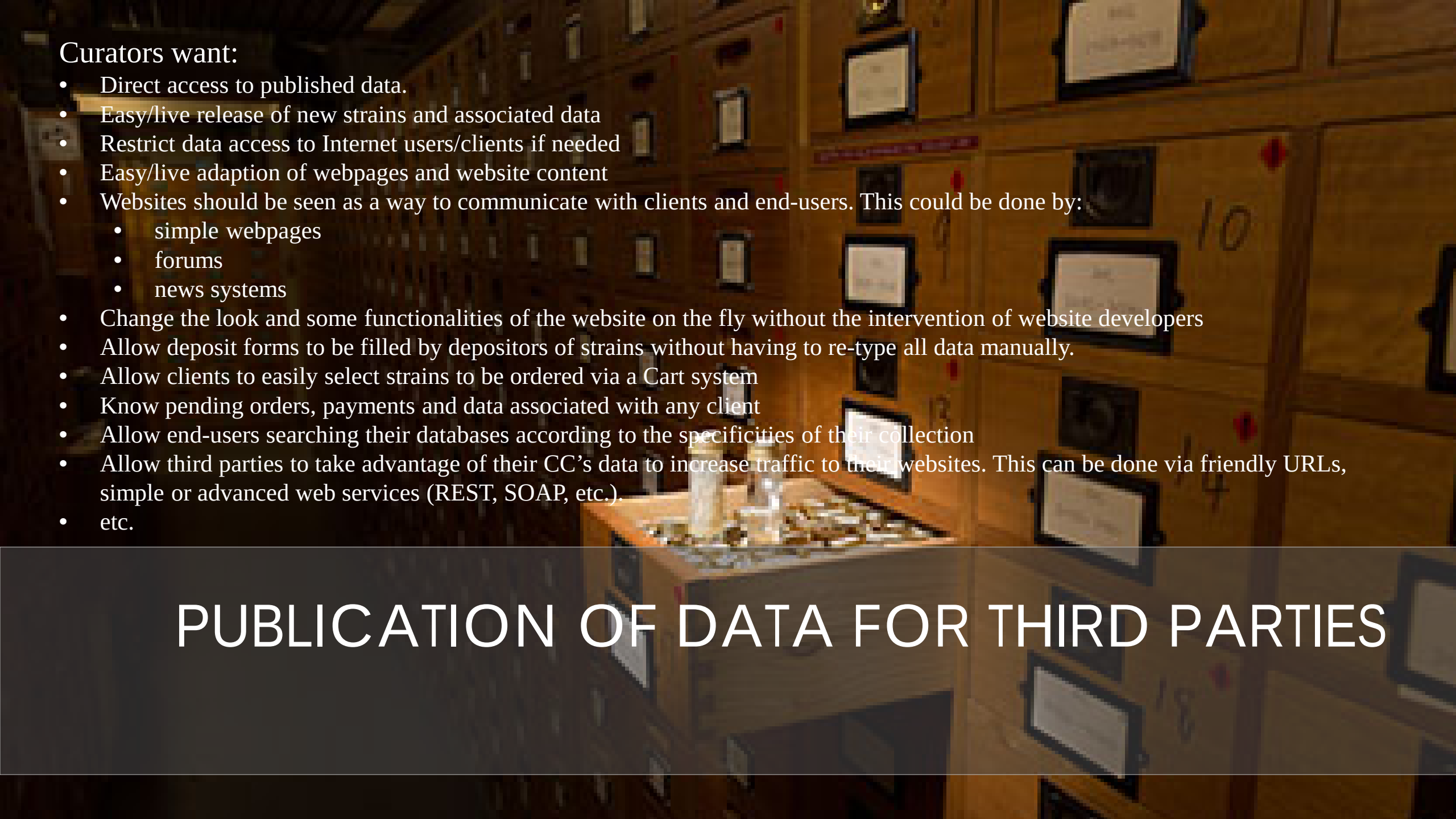
TOPICS UNDER INVESTIGATION

We love:

- Basic collection maintenance/management
- Biological material distribution
- Research
- Screening
- Dynamic System (curators/researcher can change the system without the need for IT or developers)
- Advanced security and access management
- Tracking of database modifications by each user
- Ability to import and export data as text, images, DNA trace files, microplate reader data, MS-Excel, HTML, XML, FASTA, NCBI and more
- Linking or exportation of data to other websites such as Q-bank, GBIF, StrainInfo, NCBI, etc.
- Ability to create custom layouts such as invoices, catalogs, sample labels
- Stock management
- etc

MANAGEMENT SYSTEMS FOR CURATORS

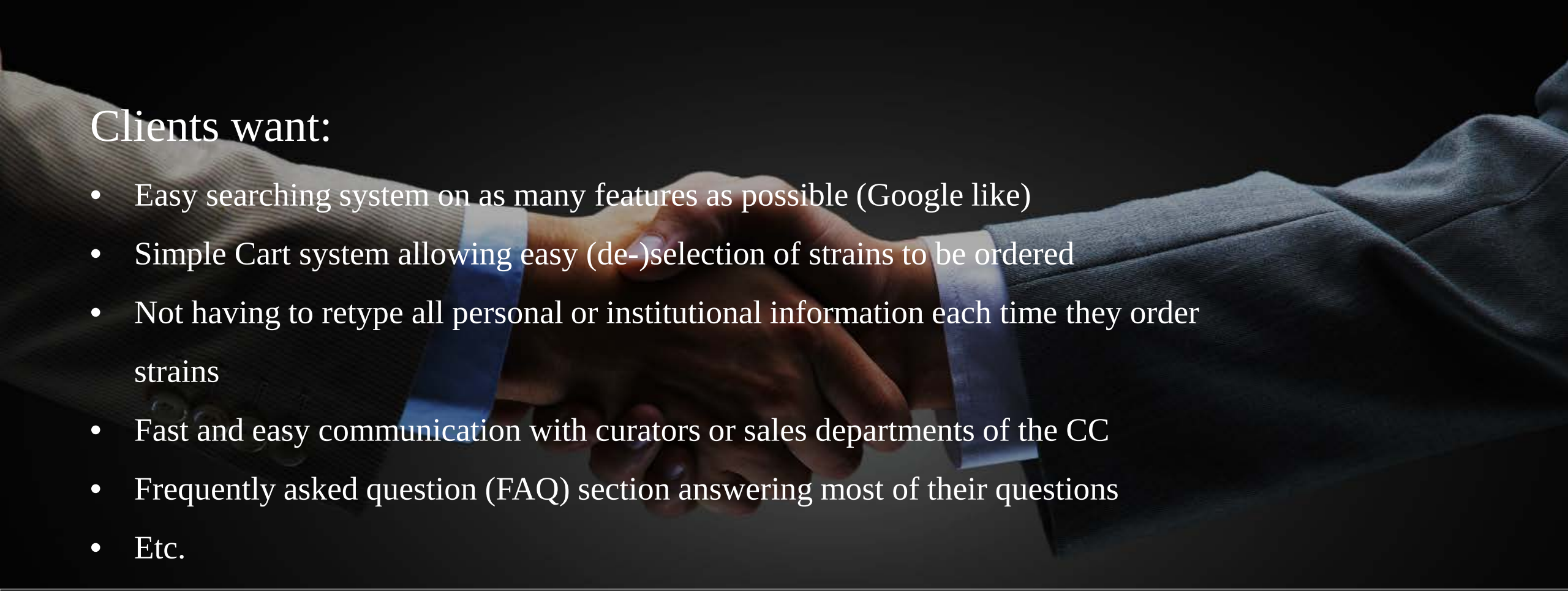
MOST WANTED FUNCTIONALITIES



Curators want:

- Direct access to published data.
- Easy/live release of new strains and associated data
- Restrict data access to Internet users/clients if needed
- Easy/live adaption of webpages and website content
- Websites should be seen as a way to communicate with clients and end-users. This could be done by:
  - simple webpages
  - forums
  - news systems
- Change the look and some functionalities of the website on the fly without the intervention of website developers
- Allow deposit forms to be filled by depositors of strains without having to re-type all data manually.
- Allow clients to easily select strains to be ordered via a Cart system
- Know pending orders, payments and data associated with any client
- Allow end-users searching their databases according to the specificities of their collection
- Allow third parties to take advantage of their CC's data to increase traffic to their websites. This can be done via friendly URLs, simple or advanced web services (REST, SOAP, etc.).
- etc.

PUBLICATION OF DATA FOR THIRD PARTIES



Clients want:

- Easy searching system on as many features as possible (Google like)
- Simple Cart system allowing easy (de-)selection of strains to be ordered
- Not having to retype all personal or institutional information each time they order strains
- Fast and easy communication with curators or sales departments of the CC
- Frequently asked question (FAQ) section answering most of their questions
- Etc.

PUBLICATION OF DATA FOR THIRD PARTIES

DATA STANDARDS AND PROTOCOLS

- BioSharing (http://biosharing.org/)
- Biodiversity Information Standards (TDWG; http://www.tdwg.org/)
- Genomic Standards Consortium (GSC; http://en.wikipedia.org/wiki/Genomic_Standards_Consortium)
- Semantic web

LINKS TO EXISTING RESOURCES

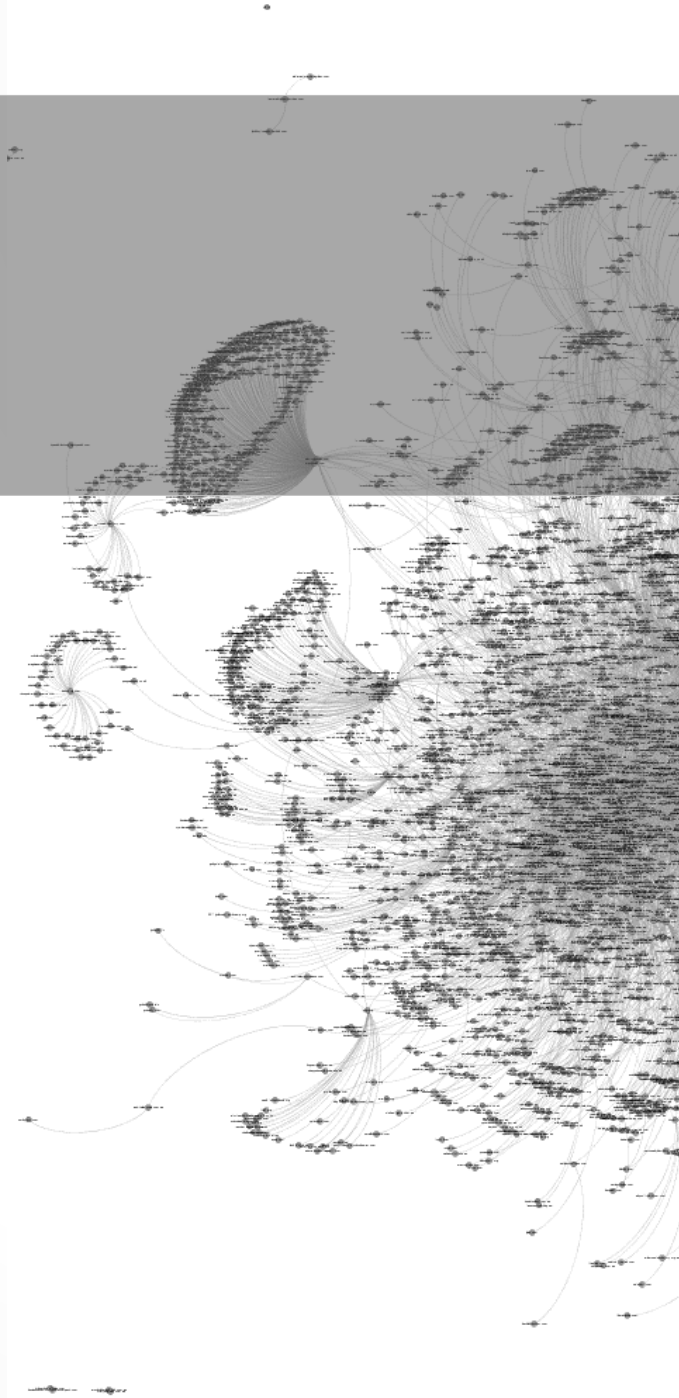
- Q-BANK
- STRAININFO, WDCM
- TAXONOMIC DATABASES (MycoBank, DSMZ, etc), GBIF
- INSDC (NCBI, ENBL, DDBJ, etc), BOLD
- LifeWatch, BioVel, Vibrant, LifeLink, Elixir, EU-Openscreen, etc
- Many more …

INTEROPERABILITY


WE NEED DATA AND SHARING!!

SEMANTIC WEB AND TRIPLETS

A triplet indicates the connection and direction



For example:

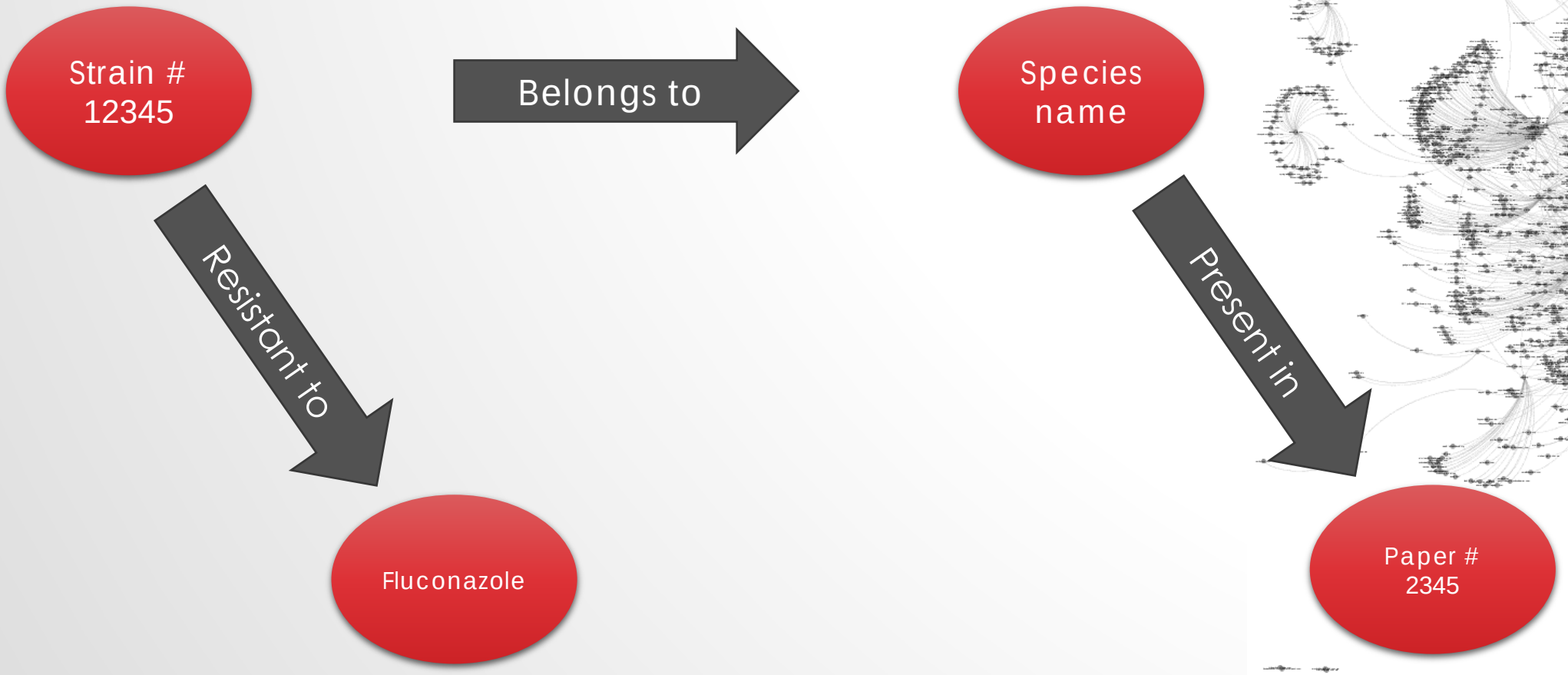
Subject	Predicate	Object
Exserohilum rostratum	causes	Phaeohyphomycosis of arm

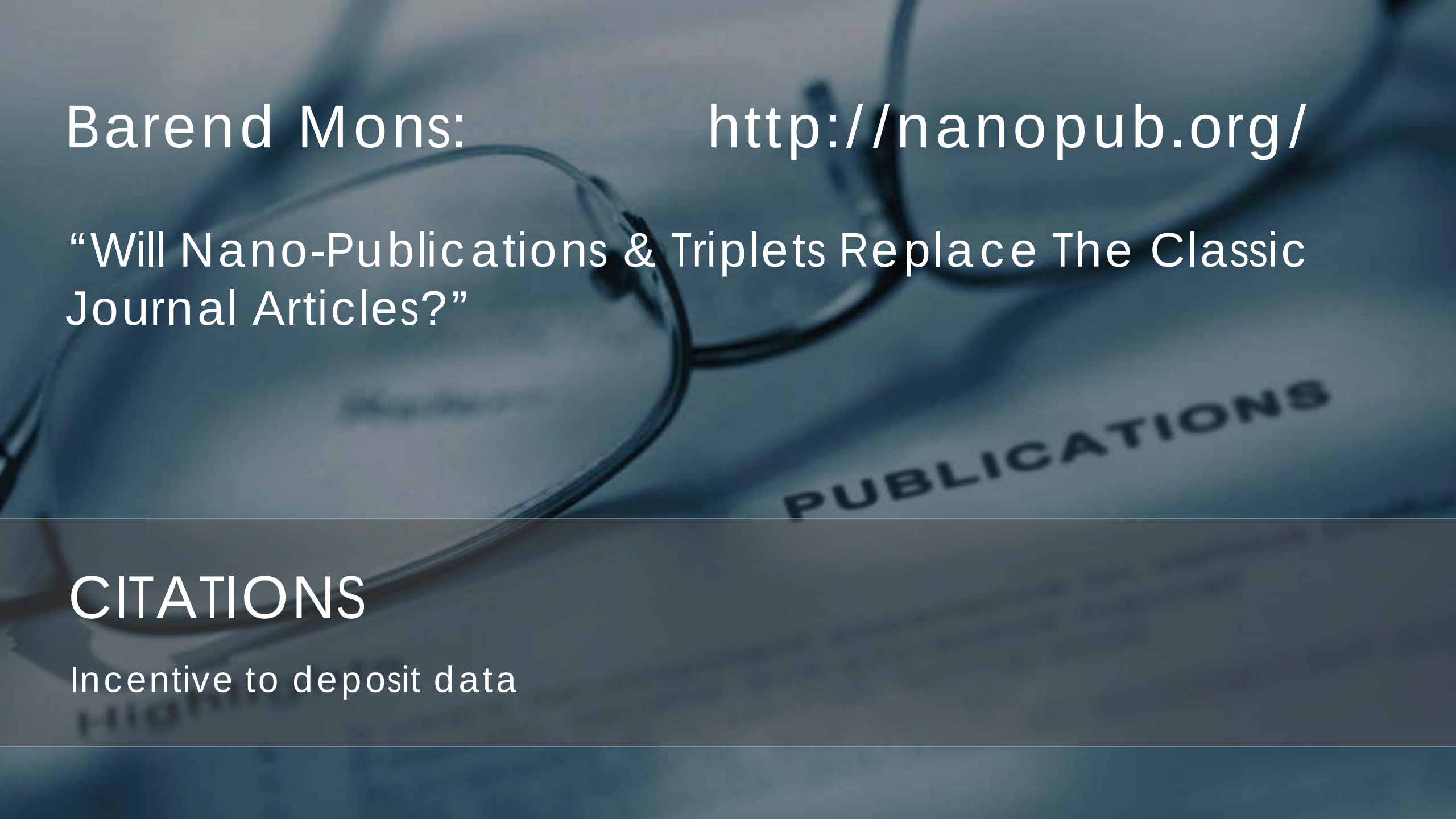
SEMANTIC WEB AND TRIPLET

Subject

Predicate

Object



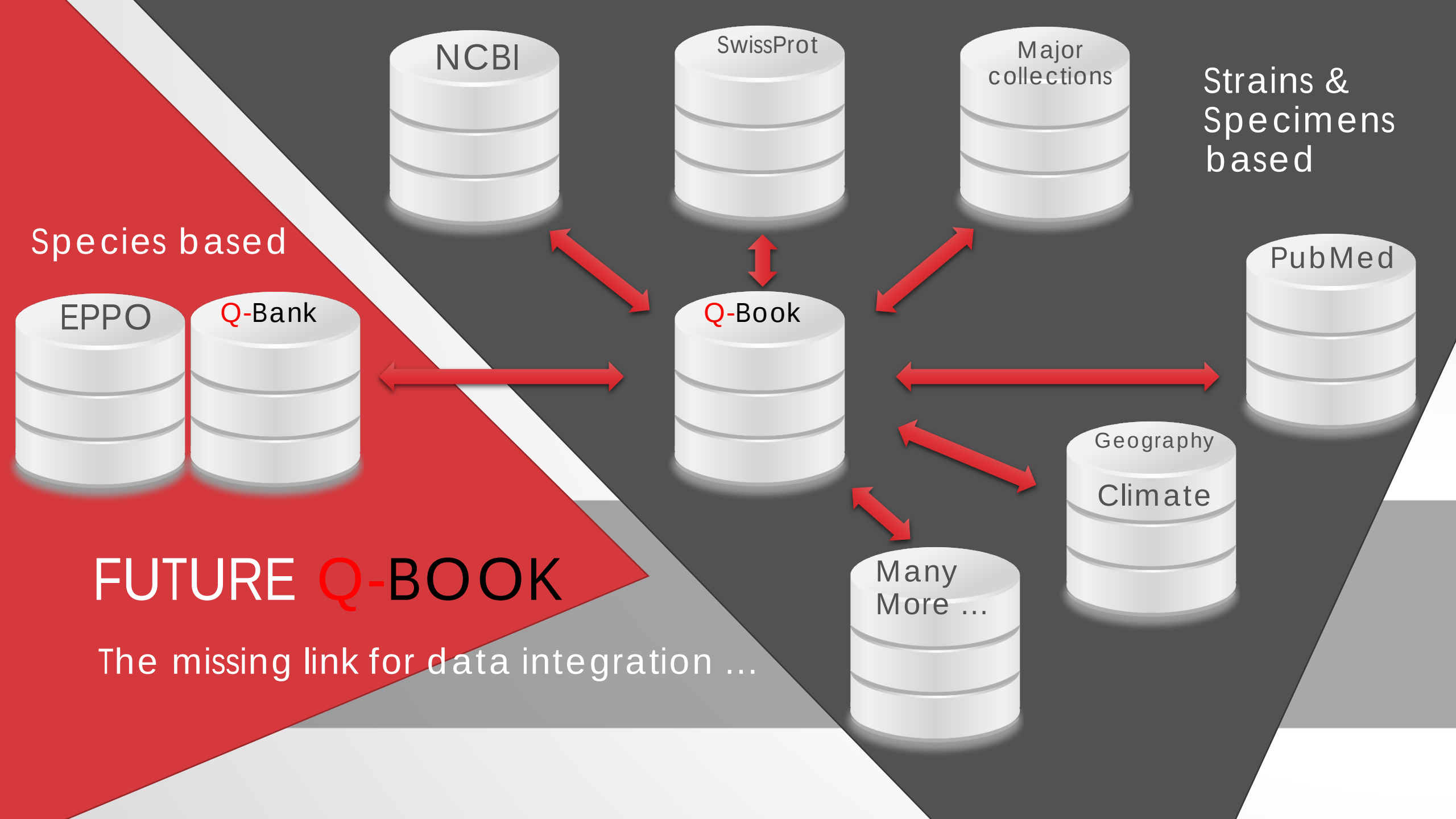


Barend Mons: http://nanopub.org/

“Will Nano-Publications & Triplets Replace The Classic Journal Articles?”

CITATIONS

Incentive to deposit data


NCBI
SwissProt
Major collections
Strains & Specimens based
Species based
EPPO
Q-Bank
Q-Book
PubMed
Geography
Climate
Many More ...
FUTURE Q-BOOK
The missing link for data integration ...

Q-Statistics
analytical tools

Q-Semantic
navigation tools

Q-bank

Linked
external
resources
&
Collections

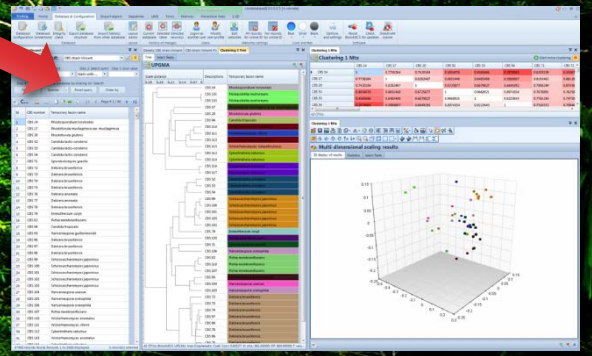
Q-Book
System
&
DB
Cloud based

EPPO

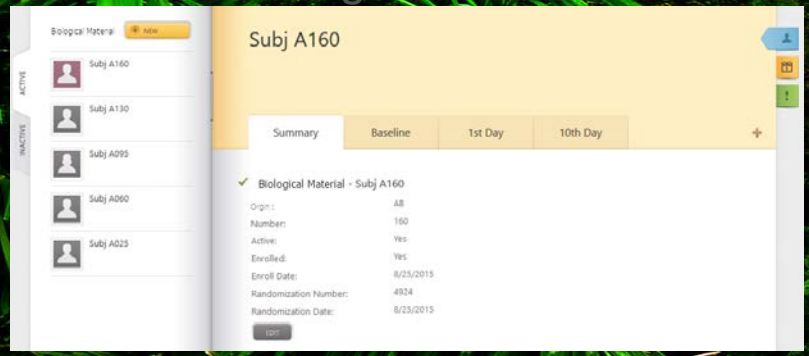
Basic
(meta)data
retrieval
tools

Local
PC or
server
DB

Local strains
management soft

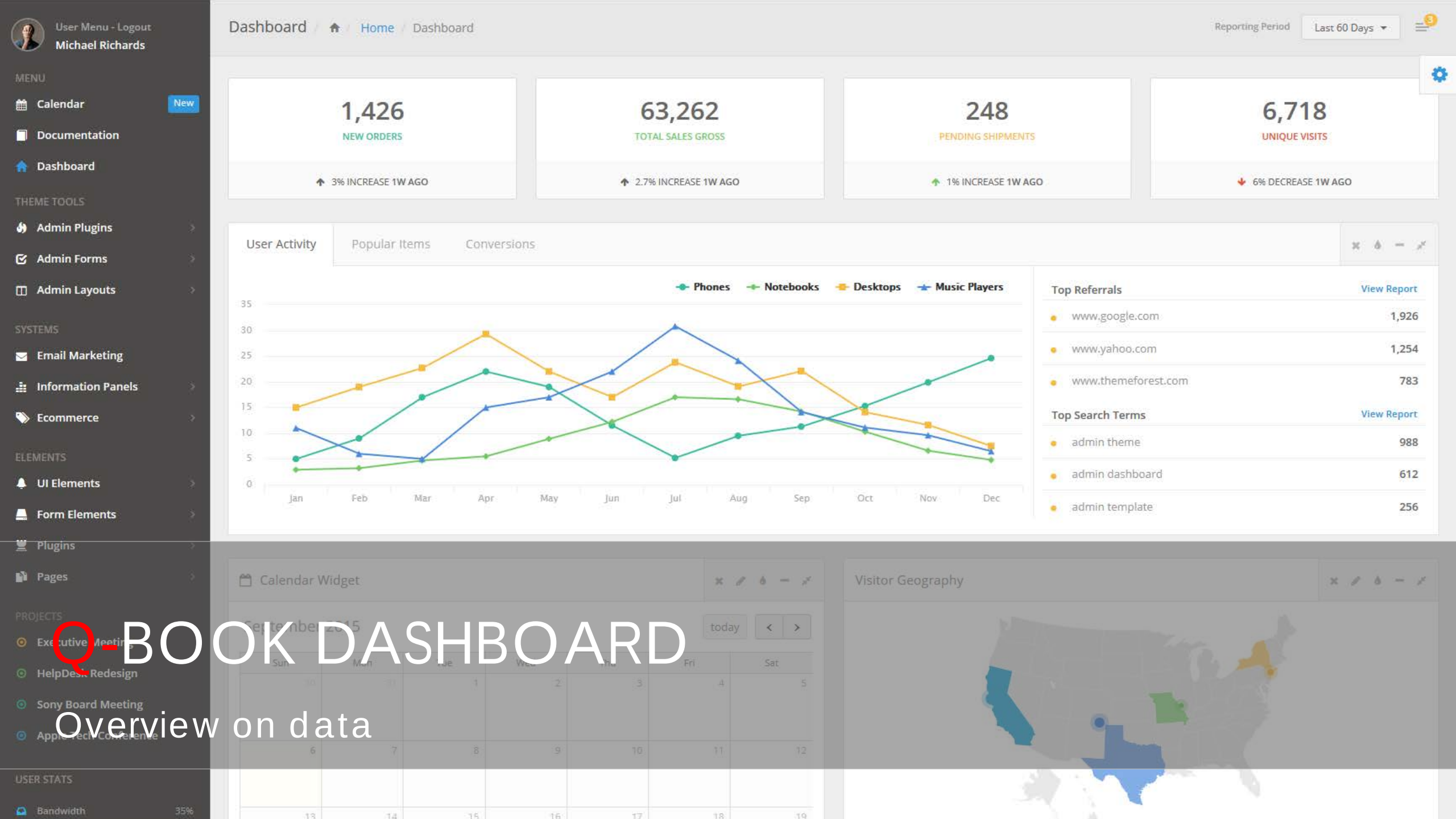


Web based strains
management soft



Q-BOOK DATA IMPORTATION & EDITION

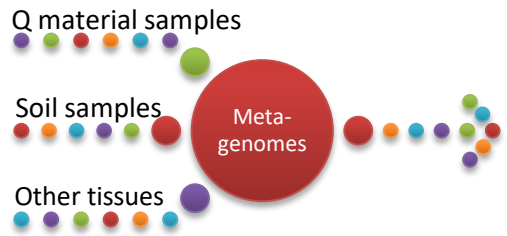
Data deposit using different methods

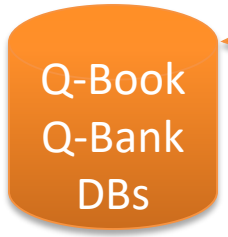


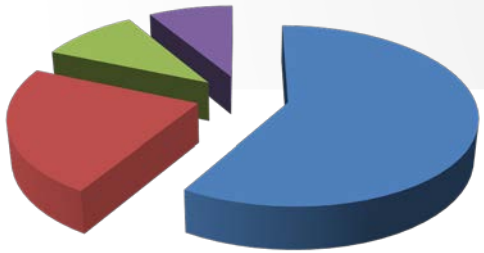
Q-BOOK DASHBOARD

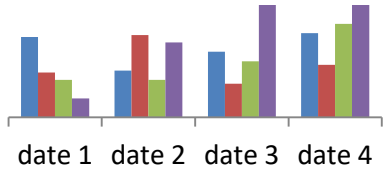
Overview on data

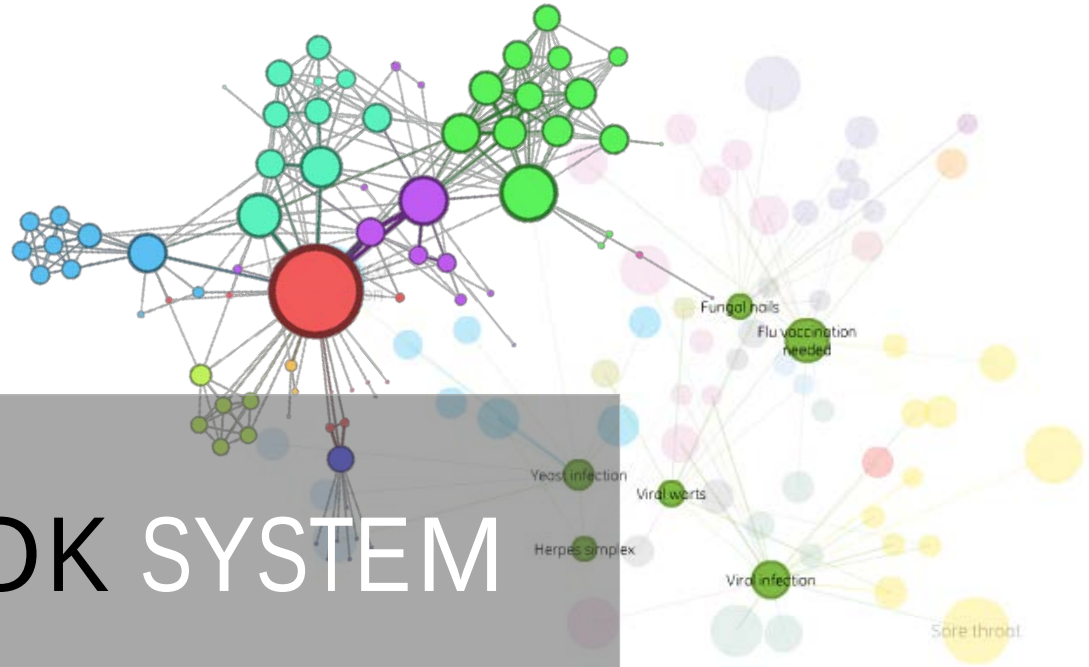
Fast identification pipeline for NGS – Sample to Diagnostic = +-24H







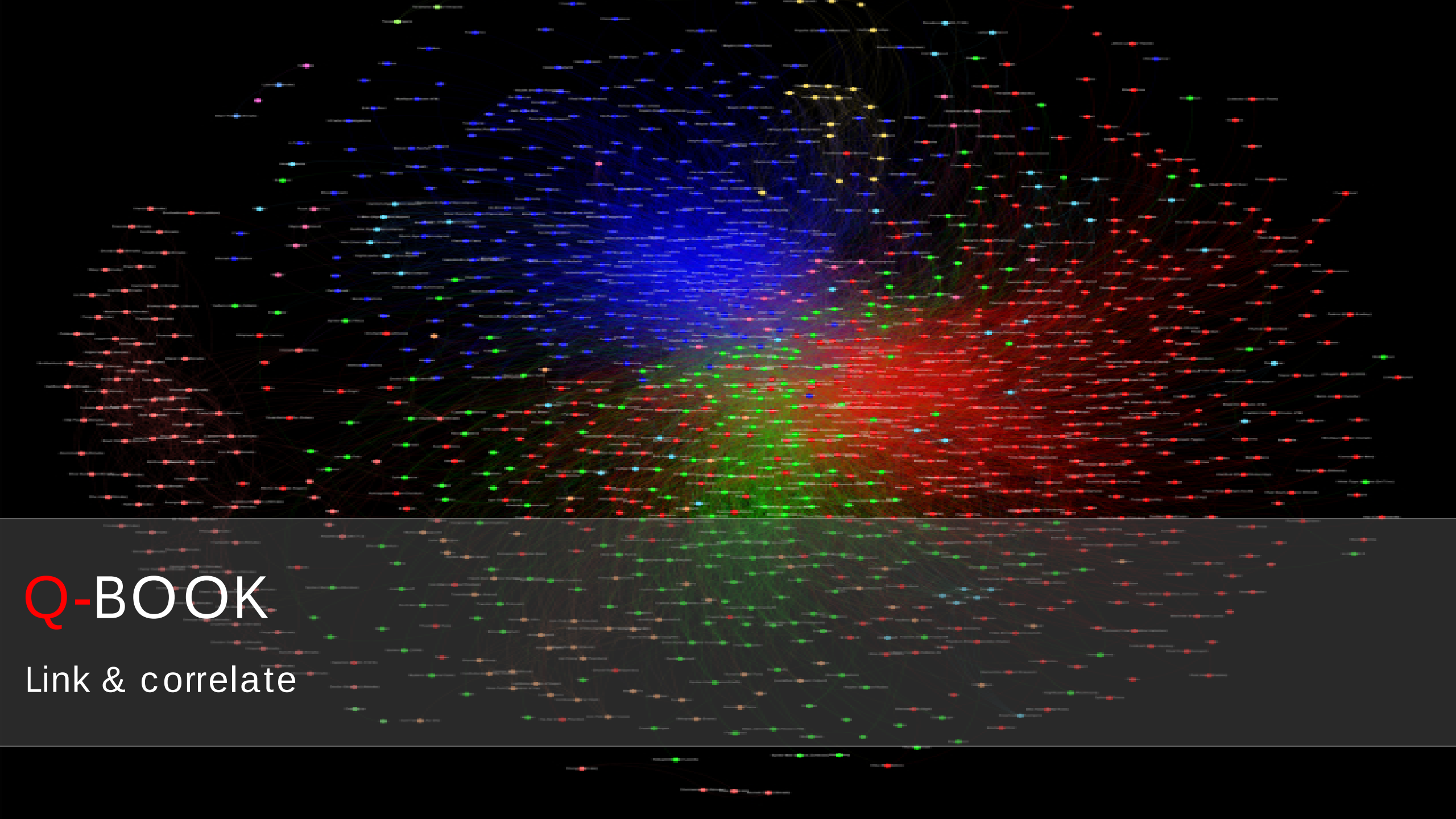




Q-BOOK SYSTEM

Advantages:

- Precise identification
- Very fast answer
- Action proposal
- Single method for all samples



Q-BOOK

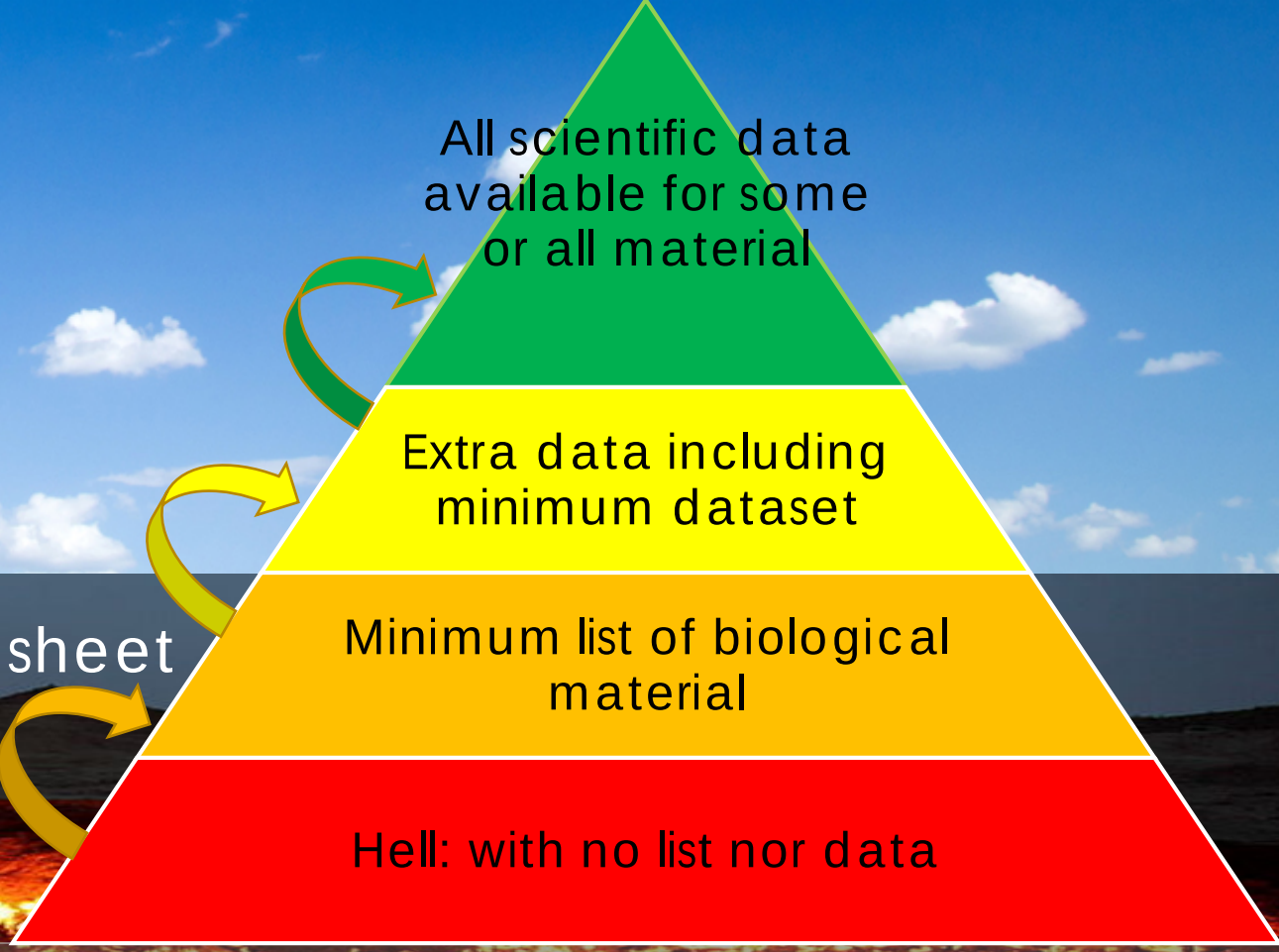
Link & correlate

Steps to move up:

3. Include all possible scientific data & importation/link in/to Q-Book

2. List with min dataset in Excel sheet & Simple importation in Q-Book

1. Simple list biological material in Excel sheet & Simple importation in Q-Book



Q-BOOK FOR ALL

Collections at all digitalization level are welcome

THANKS TO ALL OF YOU