OpenAIRE

Open Access Infrastructure for Research in Europe

Subject-specific pilots — linking data and publications across domains and infrastructures in the FP7 project OpenAIREplus

Mathias Lösch

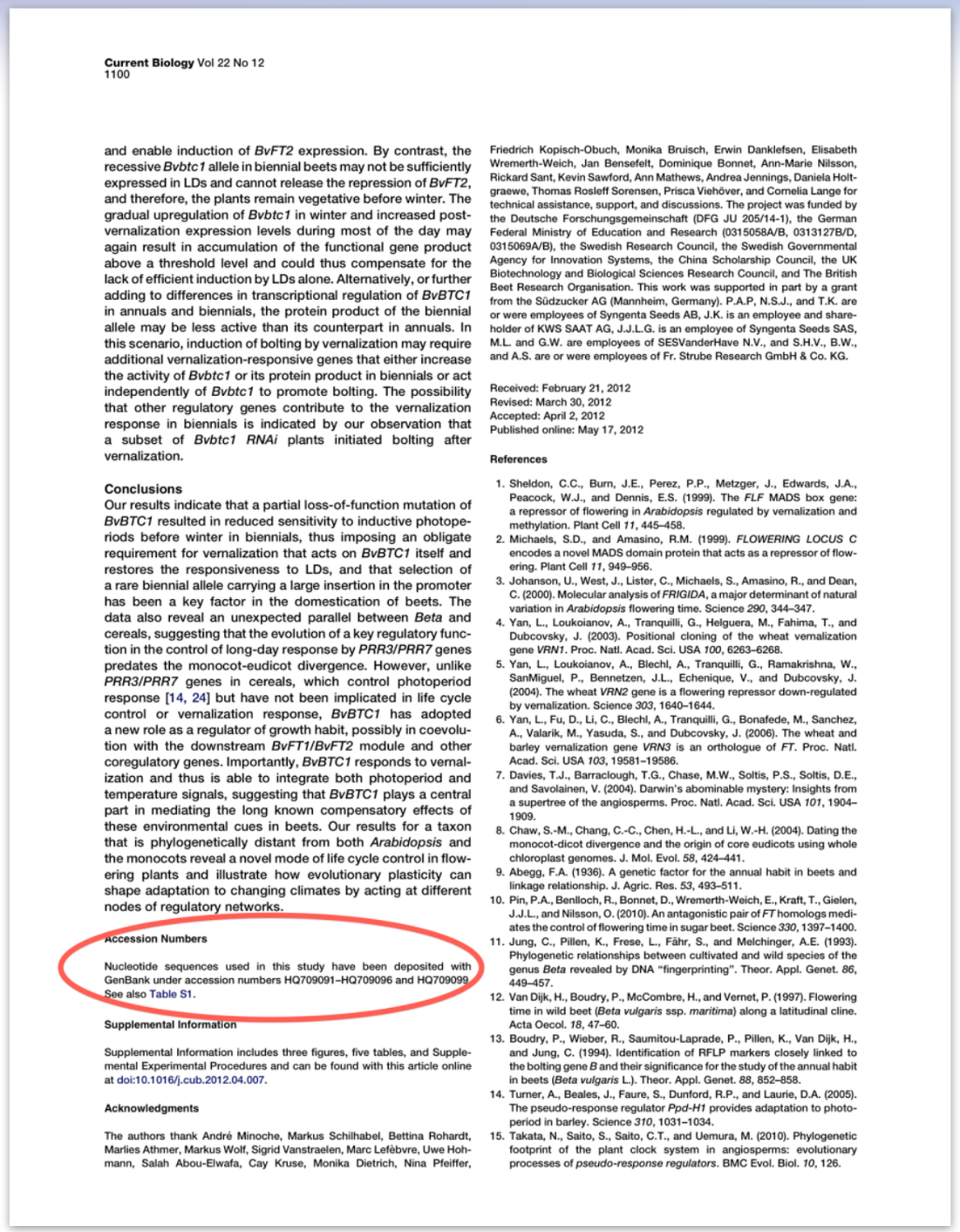
Kolloquium Wissensinfrastruktur WS 2012/13

16.11.2012

OpenAIRE and OpenAIREplus

- OpenAIRE (2009–2012): Open Access Pilot of the EU
  - Background: Open Access policy of the European Commission for FP7 (Special Clause 39 in 7 areas)
  - Goals:
    - Electronic infrastructure for EC-funded publications
    - European help desk for supporting the researchers in depositing their publications
- OpenAIREplus (2011–2014)
  - Goals:
    - Extend to other funders/programs
    - Extend the focus to research data
    - Foster linking of publications and research data

Linking Data and Publications in the Life Sciences



and enable induction of *BvFT2* expression. By contrast, the recessive *Bvbtc1* allele in biennial beets may not be sufficiently expressed in LDs and cannot release the repression of *BvFT2*, and therefore, the plants remain vegetative before winter. The gradual upregulation of *Bvbtc1* in winter and increased post-vernalization expression levels during most of the day may again result in accumulation of the functional gene product above a threshold level and could thus compensate for the lack of efficient induction by LDs alone. Alternatively, or further adding to differences in transcriptional regulation of *BvBTC1* in annuals and biennials, the protein product of the biennial allele may be less active than its counterpart in annuals. In this scenario, induction of bolting by vernalization may require additional vernalization-responsive genes that either increase the activity of *Bvbtc1* or its protein product in biennials or act independently of *Bvbtc1* to promote bolting. The possibility that other regulatory genes contribute to the vernalization response in biennials is indicated by our observation that a subset of *Bvbtc1 RNAi* plants initiated bolting after vernalization.

Conclusions

Our results indicate that a partial loss-of-function mutation of *BvBTC1* resulted in reduced sensitivity to inductive photoperiods before winter in biennials, thus imposing an obligate requirement for vernalization that acts on *BvBTC1* itself and restores the responsiveness to LDs, and that selection of a rare biennial allele carrying a large insertion in the promoter has been a key factor in the domestication of beets. The data also reveal an unexpected parallel between *Beta* and cereals, suggesting that the evolution of a key regulatory function in the control of long-day response by *PRR3/PRR7* genes predates the monocot-eudicot divergence. However, unlike *PRR3/PRR7* genes in cereals, which control photoperiod response [14, 24] but have not been implicated in life cycle control or vernalization response, *BvBTC1* has adopted a new role as a regulator of growth habit, possibly in coevolution with the downstream *BvFT1/BvFT2* module and other coregulatory genes. Importantly, *BvBTC1* responds to vernalization and thus is able to integrate both photoperiod and temperature signals, suggesting that *BvBTC1* plays a central part in mediating the long known compensatory effects of these environmental cues in beets. Our results for a taxon that is phylogenetically distant from both *Arabidopsis* and the monocots reveal a novel mode of life cycle control in flowering plants and illustrate how evolutionary plasticity can shape adaptation to changing climates by acting at different nodes of regulatory networks.

Accession Numbers

Nucleotide sequences used in this study have been deposited with GenBank under accession numbers HQ709091–HQ709096 and HQ709099. See also Table S1.

Supplemental Information

Supplemental Information includes three figures, five tables, and Supplemental Experimental Procedures and can be found with this article online at doi:10.1016/j.cub.2012.04.007.

Acknowledgments

The authors thank André Minoche, Markus Schilhabel, Bettina Rohardt, Marlies Athmer, Markus Wolf, Sigrid Vanstraelen, Marc Lefèbvre, Uwe Hohmann, Salah Abou-Elwafa, Cay Kruse, Monika Dietrich, Nina Pfeiffer, Friedrich Kopisch-Obuch, Monika Bruisch, Erwin Danklefsen, Elisabeth Wremerth-Weich, Jan Bensefelt, Dominique Bonnet, Ann-Marie Nilsson, Rickard Sant, Kevin Sawford, Ann Mathews, Andrea Jennings, Daniela Holtgraewe, Thomas Rosleff Sorensen, Prisca Viehöver, and Cornelia Lange for technical assistance, support, and discussions. The project was funded by the Deutsche Forschungsgemeinschaft (DFG JU 205/14-1), the German Federal Ministry of Education and Research (0315058A/B, 0313127B/D, 0315069A/B), the Swedish Research Council, the Swedish Governmental Agency for Innovation Systems, the China Scholarship Council, the UK Biotechnology and Biological Sciences Research Council, and The British Beet Research Organisation. This work was supported in part by a grant from the Südzucker AG (Mannheim, Germany). P.A.P, N.S.J., and T.K. are or were employees of Syngenta Seeds AB, J.K. is an employee and shareholder of KWS SAAT AG, J.J.L.G. is an employee of Syngenta Seeds SAS, M.L. and G.W. are employees of SESVanderHave N.V., and S.H.V., B.W., and A.S. are or were employees of Fr. Strube Research GmbH & Co. KG.

Received: February 21, 2012
Revised: March 30, 2012
Accepted: April 2, 2012
Published online: May 17, 2012

References

1. Sheldon, C.C., Burn, J.E., Perez, P.P., Metzger, J., Edwards, J.A., Peacock, W.J., and Dennis, E.S. (1999). The *FLF* MADS box gene: a repressor of flowering in *Arabidopsis* regulated by vernalization and methylation. Plant Cell *11*, 445–458.
2. Michaels, S.D., and Amasino, R.M. (1999). *FLOWERING LOCUS C* encodes a novel MADS domain protein that acts as a repressor of flowering. Plant Cell *11*, 949–956.
3. Johanson, U., West, J., Lister, C., Michaels, S., Amasino, R., and Dean, C. (2000). Molecular analysis of *FRIGIDA*, a major determinant of natural variation in *Arabidopsis* flowering time. Science *290*, 344–347.
4. Yan, L., Loukoianov, A., Tranquilli, G., Helguera, M., Fahima, T., and Dubcovsky, J. (2003). Positional cloning of the wheat vernalization gene *VRN1*. Proc. Natl. Acad. Sci. USA *100*, 6263–6268.
5. Yan, L., Loukoianov, A., Blechl, A., Tranquilli, G., Ramakrishna, W., SanMiguel, P., Bennetzen, J.L., Echenique, V., and Dubcovsky, J. (2004). The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization. Science *303*, 1640–1644.
6. Yan, L., Fu, D., Li, C., Blechl, A., Tranquilli, G., Bonafede, M., Sanchez, A., Valarik, M., Yasuda, S., and Dubcovsky, J. (2006). The wheat and barley vernalization gene *VRN3* is an orthologue of *FT*. Proc. Natl. Acad. Sci. USA *103*, 19581–19586.
7. Davies, T.J., Barraclough, T.G., Chase, M.W., Soltis, P.S., Soltis, D.E., and Savolainen, V. (2004). Darwin's abominable mystery: Insights from a supertree of the angiosperms. Proc. Natl. Acad. Sci. USA *101*, 1904–1909.
8. Chaw, S.-M., Chang, C.-C., Chen, H.-L., and Li, W.-H. (2004). Dating the monocot-dicot divergence and the origin of core eudicots using whole chloroplast genomes. J. Mol. Evol. *58*, 424–441.
9. Abegg, F.A. (1936). A genetic factor for the annual habit in beets and linkage relationship. J. Agric. Res. *53*, 493–511.
10. Pin, P.A., Benlloch, R., Bonnet, D., Wremerth-Weich, E., Kraft, T., Gielen, J.J.L., and Nilsson, O. (2010). An antagonistic pair of *FT* homologs mediates the control of flowering time in sugar beet. Science *330*, 1397–1400.
11. Jung, C., Pillen, K., Frese, L., Fähr, S., and Melchinger, A.E. (1993). Phylogenetic relationships between cultivated and wild species of the genus *Beta* revealed by DNA "fingerprinting". Theor. Appl. Genet. *86*, 449–457.
12. Van Dijk, H., Boudry, P., McCombre, H., and Vernet, P. (1997). Flowering time in wild beet (*Beta vulgaris* ssp. *maritima*) along a latitudinal cline. Acta Oecol. *18*, 47–60.
13. Boudry, P., Wieber, R., Saumitou-Laprade, P., Pillen, K., Van Dijk, H., and Jung, C. (1994). Identification of RFLP markers closely linked to the bolting gene *B* and their significance for the study of the annual habit in beets (*Beta vulgaris* L.). Theor. Appl. Genet. *88*, 852–858.
14. Turner, A., Beales, J., Faure, S., Dunford, R.P., and Laurie, D.A. (2005). The pseudo-response regulator *Ppd-H1* provides adaptation to photoperiod in barley. Science *310*, 1031–1034.
15. Takata, N., Saito, S., Saito, C.T., and Uemura, M. (2010). Phylogenetic footprint of the plant clock system in angiosperms: evolutionary processes of *pseudo-response regulators*. BMC Evol. Biol. *10*, 126.

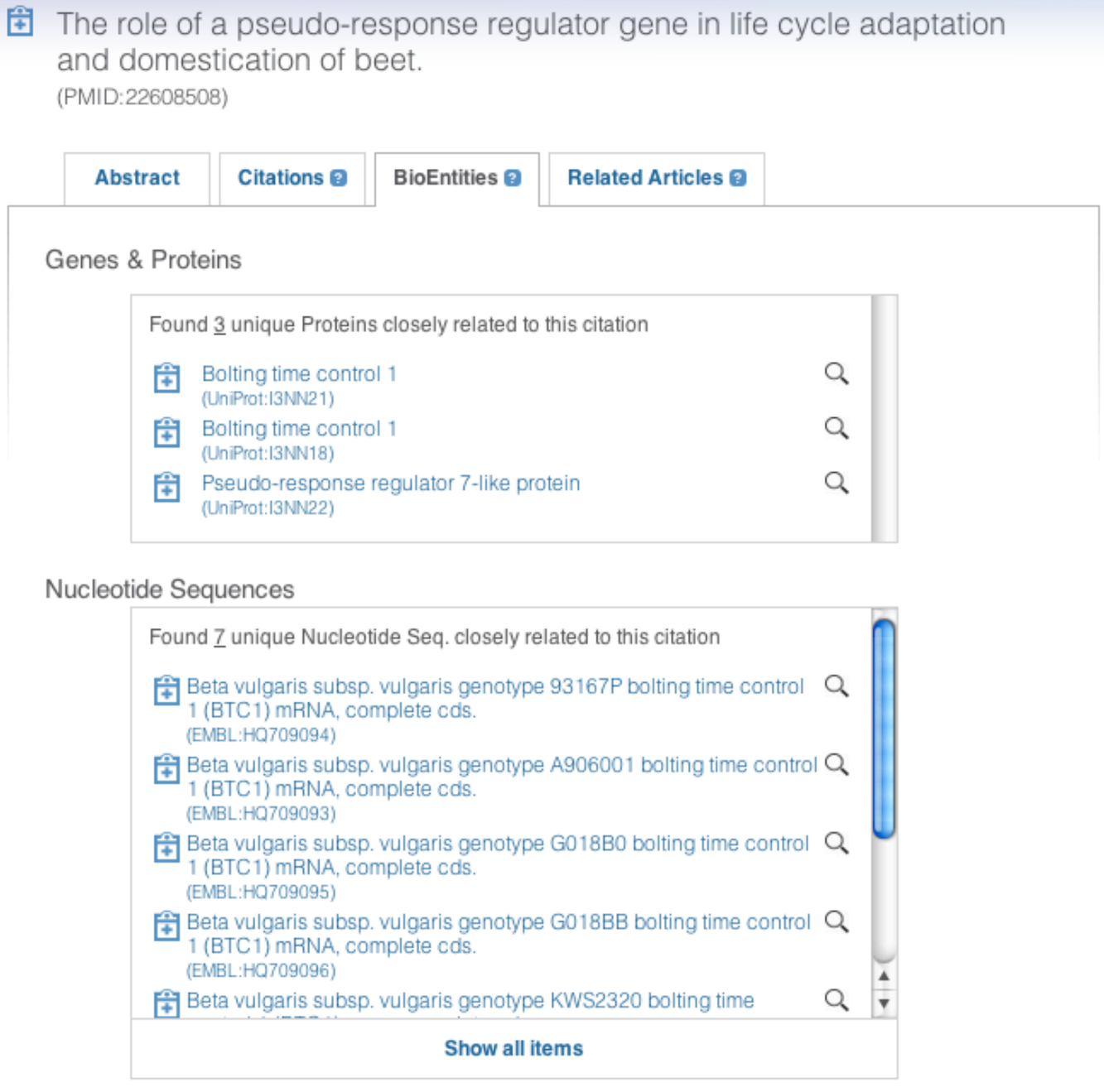
- Pin, P., W. Zhang, S. Vogt, N. Dally, B. Bu Büttner, G. Schulze-Buxloh, N. Jelly, T. Chia, et al. (2012). The role of a pseudo-response regulator gene in life cycle adaptation and domestication of beet. Current Biology.

- Lists GenBank accession numbers for the used nucleotide sequences

Accession Numbers

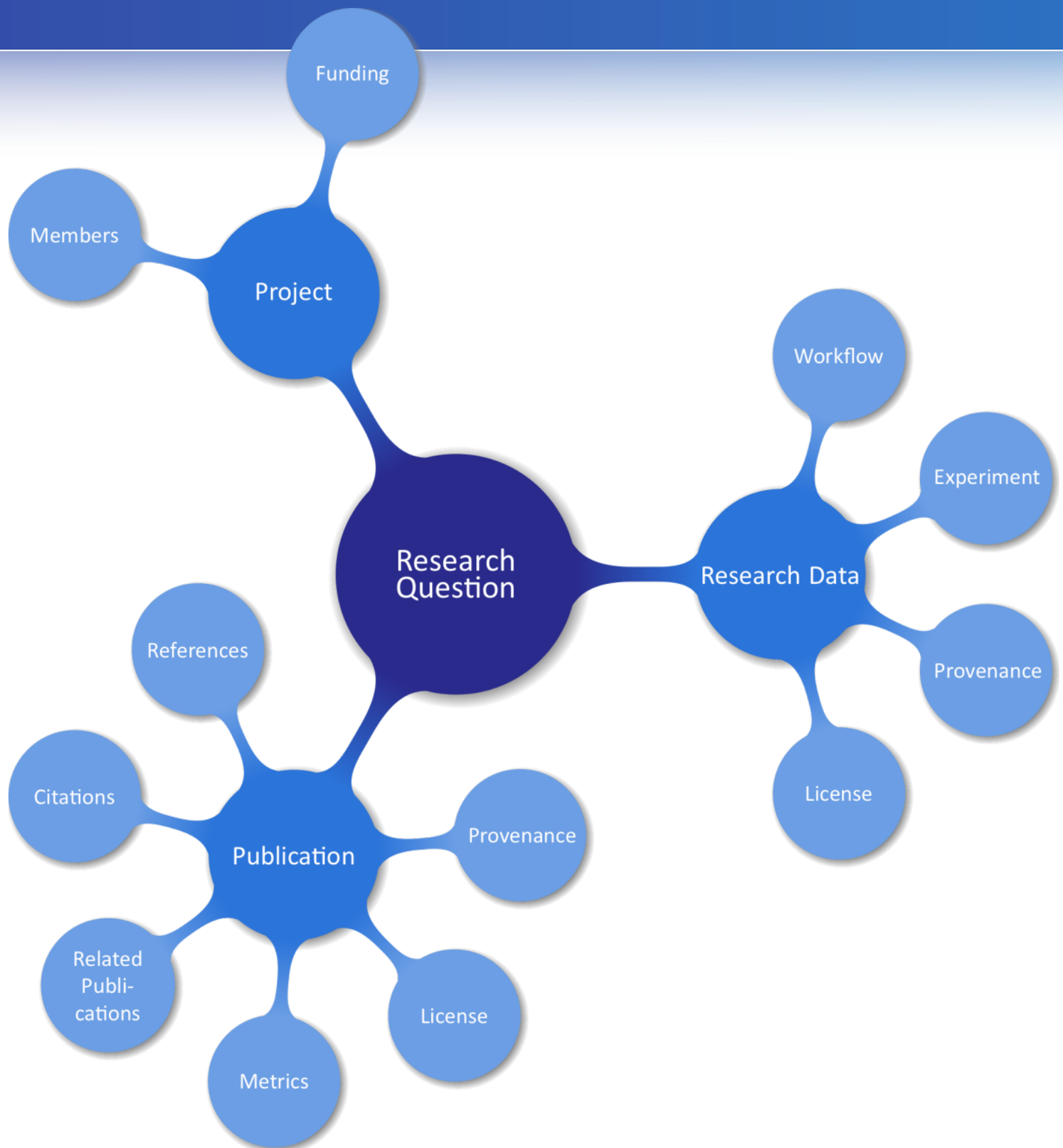
Nucleotide sequences used in this study have been deposited with GenBank under accession numbers HQ709091–HQ709096 and HQ709099. See also Table S1.

EuropePMC
http://europepmc.org/



- Displays related sequences as direct database links
- Displays more information from many sources:
  - Related publications (PubMed, PubMed Central, UK PubMed Central)
  - Expert-curated links to Life Science databases
  - Automatically mined biological terms as links to Life Science Databases
  - Citations
  - References

“Information in Context”



Advantages of Context Information

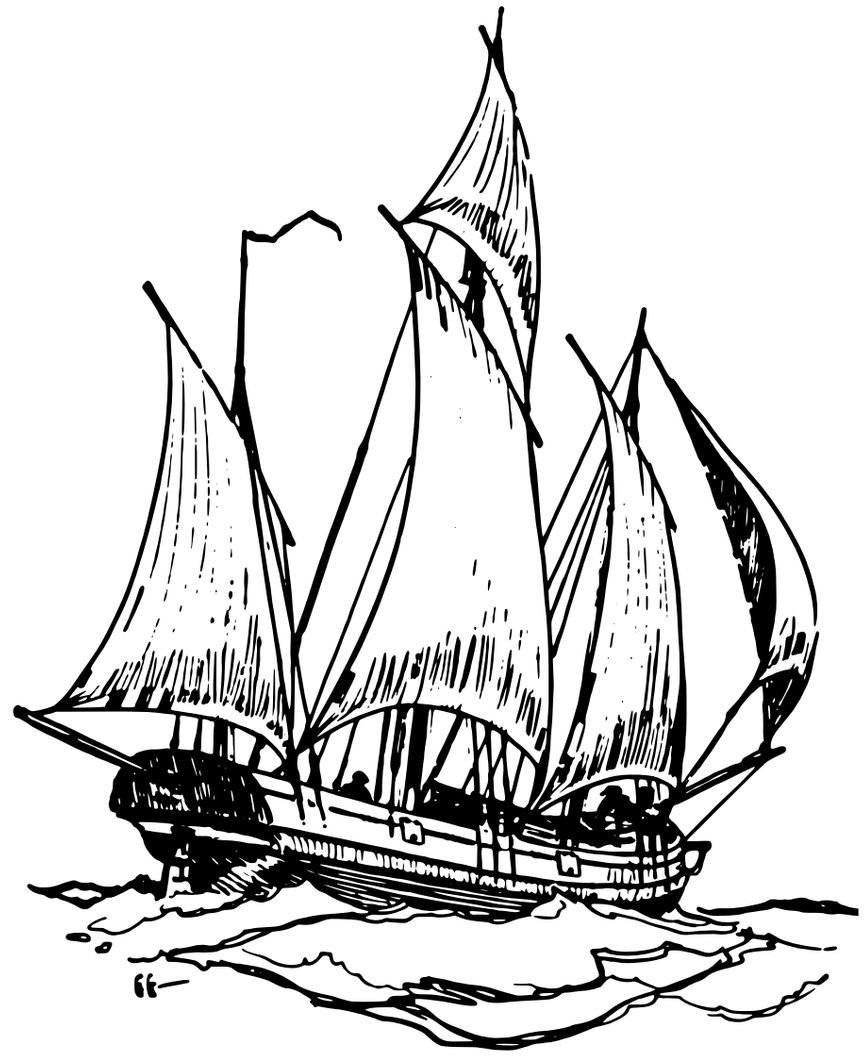
- Better discoverability and re-use of research data (datasets)
- Easier verification of research results (datasets)
- Better discoverability of related research (references, citations, related publications)
- Easier assessability of a publication's importance in a research area (citations, usage statistics)

OpenAIREplus WP3 – “Studies on practices and principles of OA”

- Follow-up package on OpenAIRE WP7 – “Subject-specific requirements and data challenges”
  - Meier zu Verl, C. and W. Horstmann (eds.): *Studies on subject-specific requirements for Open Access infrastructure. Bielefeld, 2011. DOI:10.2390/PUB-2011-1*
- OpenAIREplus T3.1:
  - Create subject-specific demonstrators for “Enhanced Publications”
  - Showcase how publications + context information can be exchanged between infrastructures

Goals of T3.1

- Research
  - Status quo of how context information is managed in different subjects
  - Types of context information
- Prototype
  - Development of two prototypes showcasing the exchange and display of context information between infrastructures
- Evaluate & Report
  - Get feedback from researchers on the prototypes
  - Formulate recommendations on how to represent and exchange context information in OpenAIRE



Partners of OpenAIREplus Subject-Specific Pilots

- Scientific Partners
  - European Bioinformatics Institute (EMBL-EBI, Life Sciences)
  - Data Archiving and Networked Services (DANS, Social Sciences & Humanities)
  - Science & Technology Facilities Council (STFC, Climate Science)
- Technical Partners
  - Bielefeld University Library (UNIBI, task lead)
  - Consiglio Nazionale delle Ricerche (CNR)
  - National and Kapodestrian University of Athens (NKUA)
  - Interdisciplinary Centre for Mathematical and Computational Modelling (ICM)

Typology of Context Information

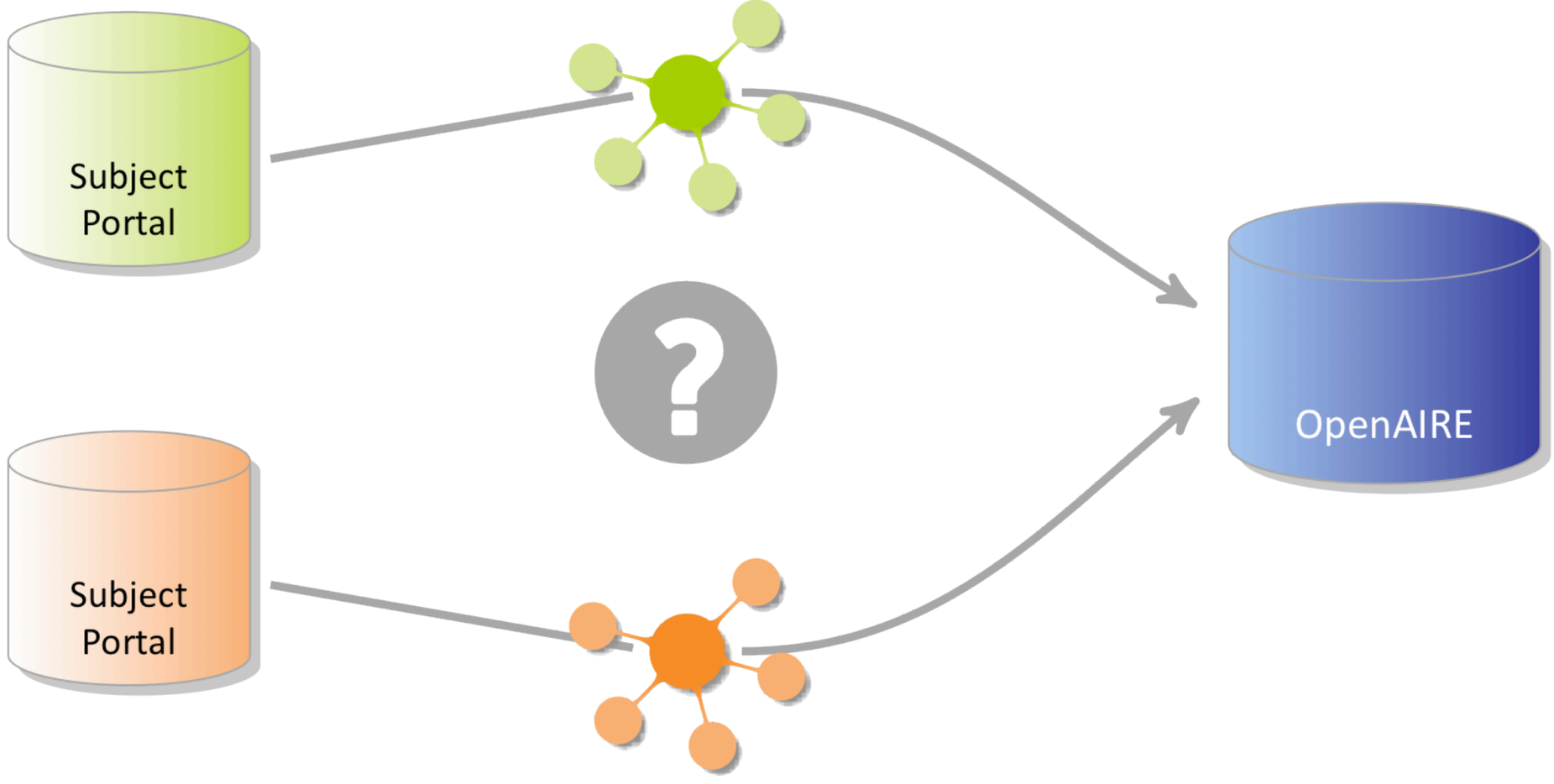
- Research Datasets
  - Research data (e.g., statistical results, questionnaires, etc.)
  - “Database links” (referencing canonical data entities)
- “Supplementary Material”: additional tables, figures etc.
- References/Citations
- Metrics/Usage Statistics
- Project/Funding information
- “Related publications”: automatically recommended through usage analysis or Text Mining

Typology II

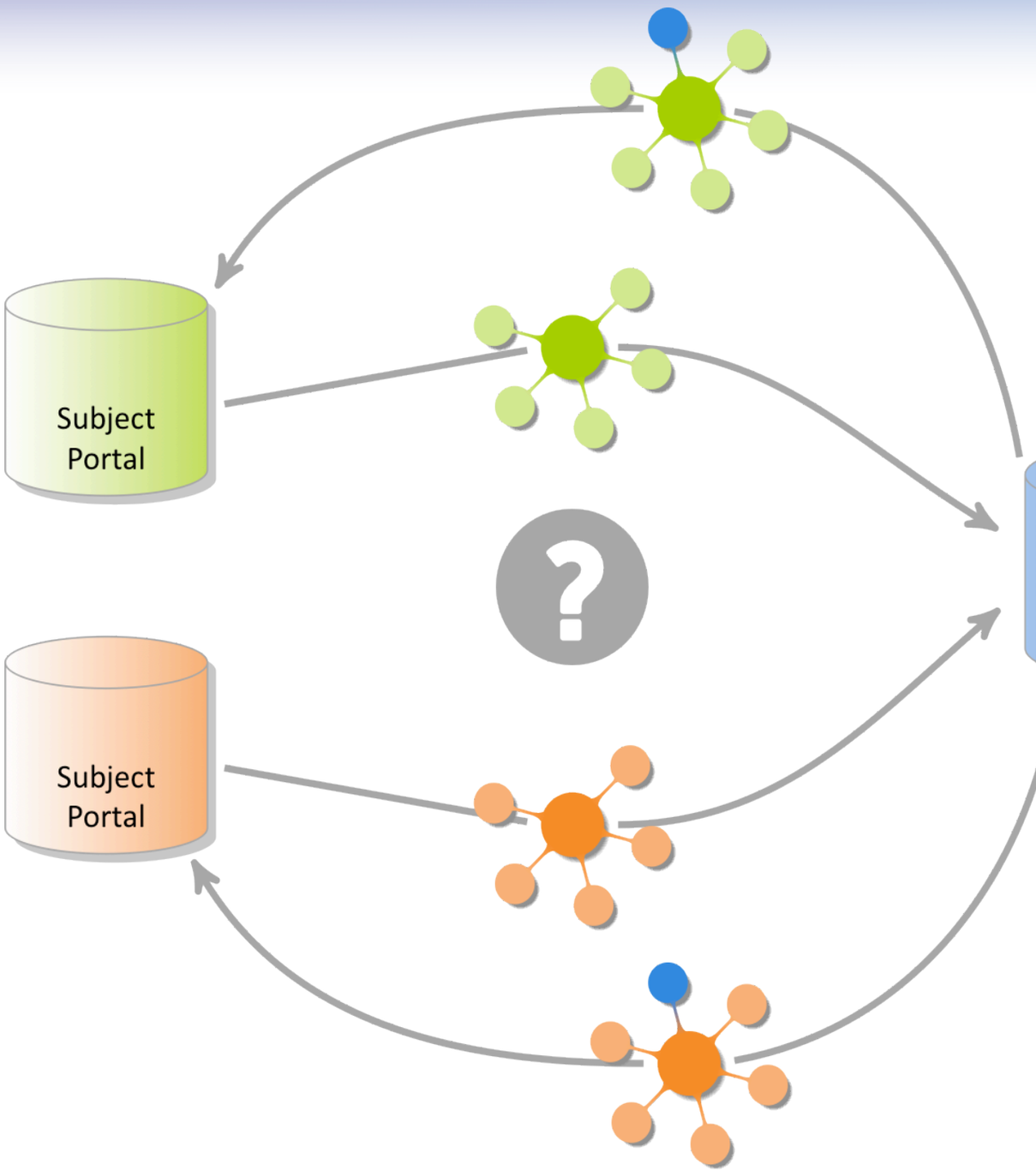
- When is the context information produced?
  - At publication time
  - Post publication
- Who produces it?
  - Author
  - Data curator (domain expert)
  - Automatic inference (Machine)
  - [Anyone (“Crowdsourcing”, “Citizen science”)]

Challenge: Exchange

- How to package publications and context information into “Compound Objects” to exchange them between different infrastructures?
- How to generalize this task for different subjects?



Challenge: Mutual Exchange

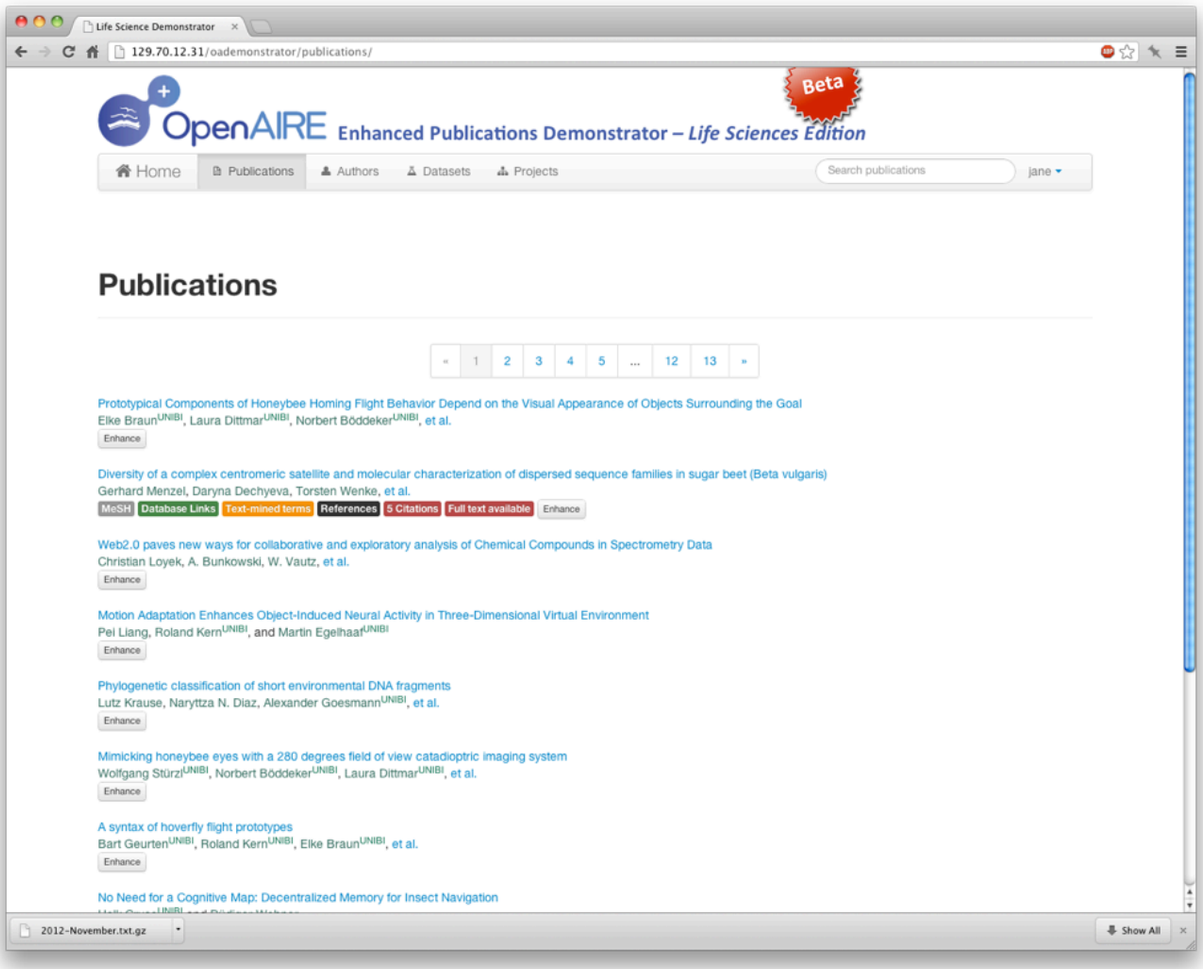


- Project information
  - Description
  - Funding program
  - Other publications of this project
- Usage statistics
- Text mining/usage analysis
  - Related publications
  - References
  - Citations
- How to pass this information back to the original sources?

Two Demonstrators for Managing Information in Context

- Social Sciences & Humanities (Development: DANS)
  - http://openaire.dans.knaw.nl/about/Language/EN
  - Data Sources:
    - Repositories (Publication Metadata)
    - DANS EASY Archive (Datasets)
    - DANS Narcis Portal (Dataset links)
    - OpenAIRE (Publication Metadata, project Information)
- Life Sciences (Development: UNIBI)
  - http://129.70.12.31/oademonstrator/
  - Data Sources:
    - Repositories (Publication Metadata)
    - Europe PMC (References, Citations)
    - EBI Life Science Databases, EBI Web Service (Database links)
    - OpenAIRE (Publication Metadata, project Information)

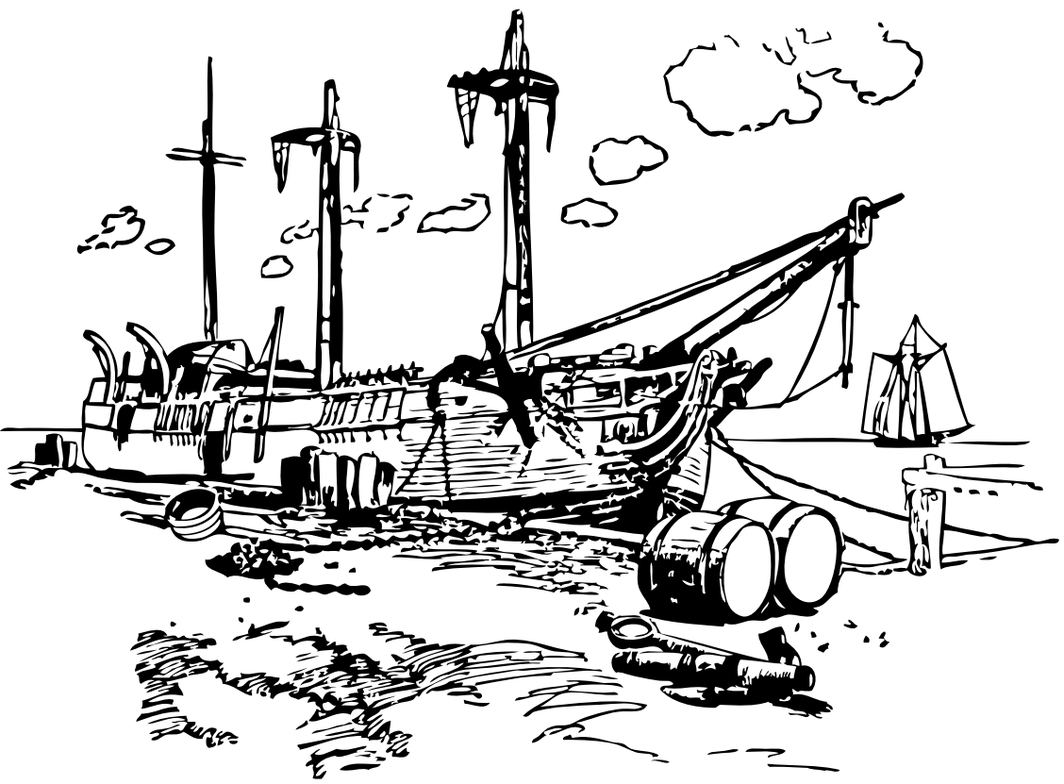
Life Science Prototype (developed by UNIBI)



- Example publications from UNIBI repository “PUB”
- Context information from the EBI Web service
  - Database links
  - Text-mined terms
  - References & Citations
  - MeSH subject headings
- Context information from OpenAIRE
  - Project Information

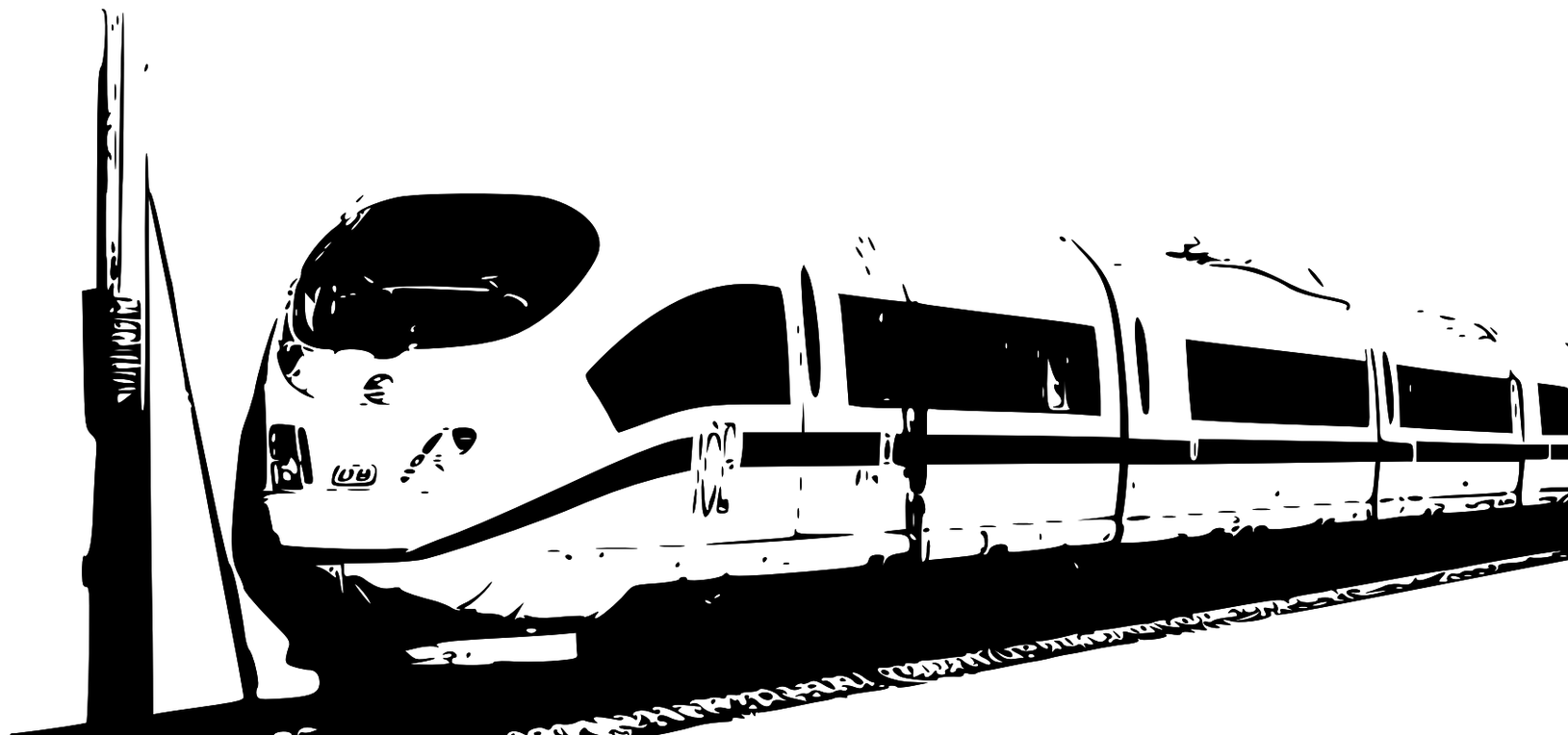
Initial Feedback from Researchers

- Generally positive!
- Suggestions:
  - Consider linking to further data sources
  - Reasonably limit amount of displayed information or link to next-higher entity (e.g., organism)



Life Sciences Demonstrator – Development

- Rapid prototyping
  - ~3 months of development
  - ~1000 lines of Python code
  - ~1000 lines of HTML code



Tools Used

- Logos link to the web pages of the projects

Bootstrap

WHOOSH
Python search library

Flask
web development,
one drop at a time

Jinja

SQLAlchemy

python™

Outlook: Local Perspective

- Adaption to local repository “PUB – Publications at Bielefeld University”:
  - Import database links for publications with PubMed IDs via the EBI Web service
  - Display the additional information on the PUB splash pages

Thank you!

Contact

Mathias Lösch
Mathias.Loesch@uni-bielefeld.de

www.openaire.eu

Thanks to: Alessia Bardi, Sarah Callaghan, Arjan Hogenaar, Maarten Hoogerwerf, Katerina Iatropoulou, Jakub Jurkiewicz, Dimitra Keramida, Oliver Kilian, Mateusz Kobos, Natalia Manola, Paolo Manghi, Johanna McEntyre, Najla Rettberg, Jochen Schirrwagen, Stephen Spencer