BIG DATA EUROPE

Empowering Communities with Data Technologies

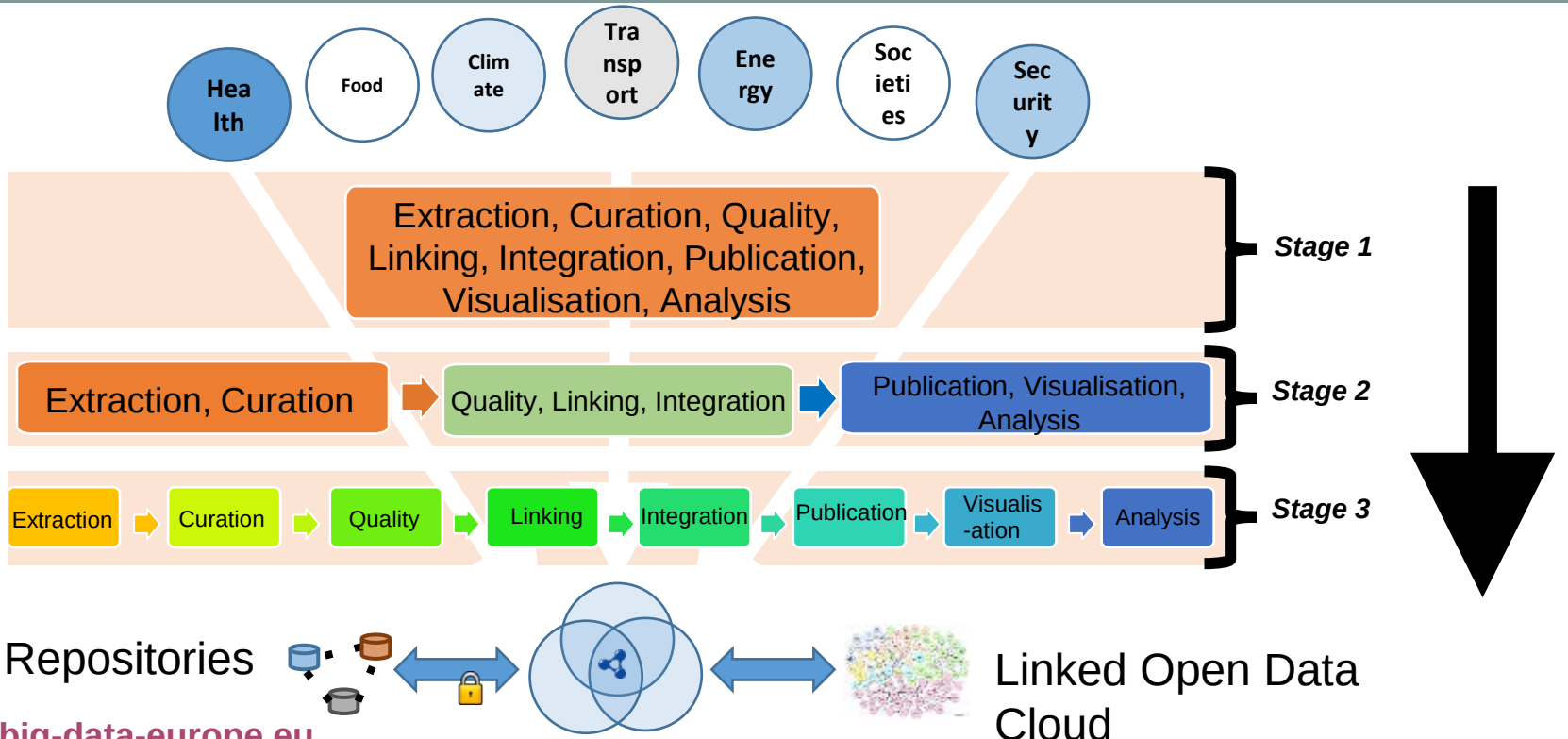
BIG DATA EUROPE

SC1 Hangout – Big Data Challenge in Health

Agenda for Today

- Welcome!
- Brief into and background (OPF)
- Introduction to the BDE Platform (Tenforce)
- Our Challenge and plan for pilot (VU)
- Q&A

Data Value Chain Evolution



www.big-data-europe.eu

Big Data in Europe: Obstacles

#1 Big Data “Variety“ problem

- Multiple Data Sources
- Required: Integration, Harmonisation

#2 Opening-up Data concerns

- Loss of control, lack of tracking
- Reservations about large corporations

#3 Limited Skills, Training, Technology

- Lack of Data Scientists
- Lack of Generic Architectures, components

Societal Domains, Focus Areas, Data Assets

Societal Domain	Preliminary Big Data Focus area	Selected Key Data assets
Life Sciences & Health	Heterogeneous data Linking & integration Biomedical Semantic Indexing & QA	ACD Labs / ChemSpider, ChEBI, ChEMBL, Con-ceptWiki, DrugBank, EN-ZYME, Gene Ontology, GO Annotation, Swis-sProt, UniProt, Wik-iPathways, PubMed, MeSH, Disease Ontology (DO), Joint Chemical Dic-tionary (Jochem), Bio-ASQ datasets
Food & Agriculture	Large-scale distributed data integration	INFOODS, AQUASTAT Green Learning Network (GLN), Agricultural Bibliography Network (ABN), AGRIS, AquaMaps, Fishbase
Energy	Real-time monitoring, stream processing, data analytics, and decision support	European Energy Exchange Data, smart meter measurement data, gas/fuels/energy market/price data, consumption statistics, equipment condition monitoring data)
Transport	Streaming sensor network & geo-spatial data integration	GTFS data, OSM/ LinkedGeoData, MobilityMaps, Transport sensor data, ROSATTE Road safety attributes, European Road Data Infrastructure - EuroRoadS
Climate	Real-time monitoring, stream processing, and data analytics.	European Grid Infrastructure (EGI), Databases hosting atmospheric data. Several software frameworks for simulation, calibration and reconstruction.
Social Sciences	Statistical and research data linking & integration	Federated social sciences data catalogs, statistical data from public data portals and statistical offices (e.g. EuroStats, UNESCO, WorldBank)
Security	Real-time monitoring, stream processing, and data analytics. Image data analysis	Earth Observation data (e.g. Very High Resolution Satellite Imagery acquired from commercial providers and governmental systems) and collateral data for supporting CFSP/CSDP missions and operations, Databases hosting atmospheric Data. Experimental and simulation data concerning dispersion of hazardous substances

Open PHACTS
Open Pharmacological Space

Why is it so hard to….

LETTERS NATURE CHEMICAL BIOLOGY VOLUME 3 NUMBER 5 MAY 2007 nature chemical biology

Chemical genetics reveals a complex functional ground state of neural stem cells

Phedias Diamandis[1–4], Jan Wildenhain[4], Ian D Clarke[1,2], Adrian G Sacher[1,2], Jeremy Graham[1,2], David S Bellows[3], Erick K M Ling[1,2,5], Ryan J Ward[1,2,5], Leanne G Jamieson[1,2,5], Mike Tyers[3,4] & Peter B Dirks[1,2,5,6]

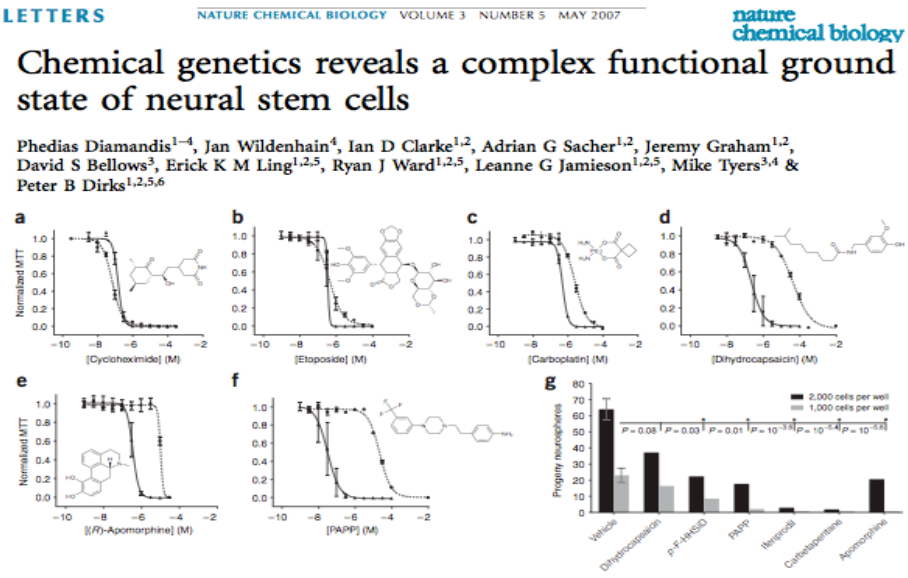
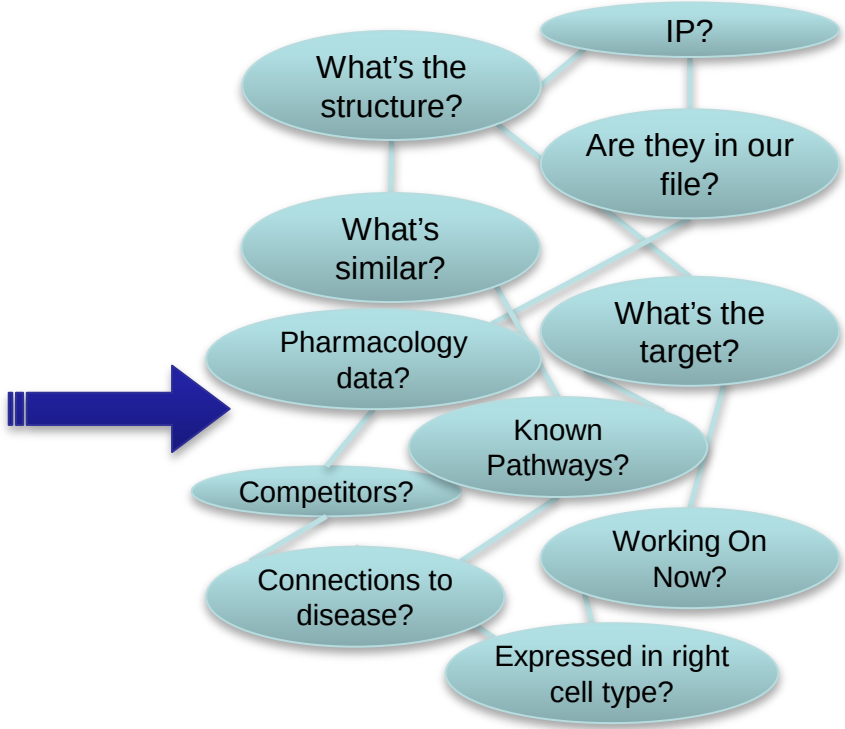


Figure 2 Identification of potent NPC-specific compounds. (a–f) Dose-response curves and chemical structures of controls: cycloheximide (a), etoposide (b) and carboplatin (c), and of selected newly identified compounds: dihydrocapsaicin (d), apomorphine (e) and PAPP (f). Each plot shows the fitted sigmoidal logistic curve to MTT proliferation assay readings of both astrocytes (- -●- -) and neurosphere cultures (–▲–). Values represent the mean and

© 2007 Nature Publishing Group http://www.nature.com/naturechemicalbiology



Open PHACTS
Open Pharmacological Space

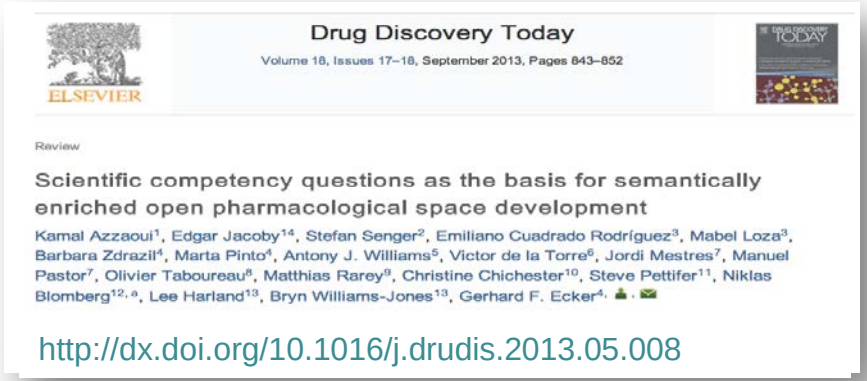
Open PHACTS Mission:

Integrate Multiple Research Biomedical Data Resources Into A Single Open, Sustainable & Free Access Point

Open PHACTS
Open Pharmacological Space

The Open PHACTS Discovery Platform

- **Cloud-Based "Production" Level System. Secure & Private**
- **Guided By Business Questions**
- **Uses Semantic Web Technology <u>But</u> provides a simple REST-ful API for everyone else**



ELSEVIER

Drug Discovery Today

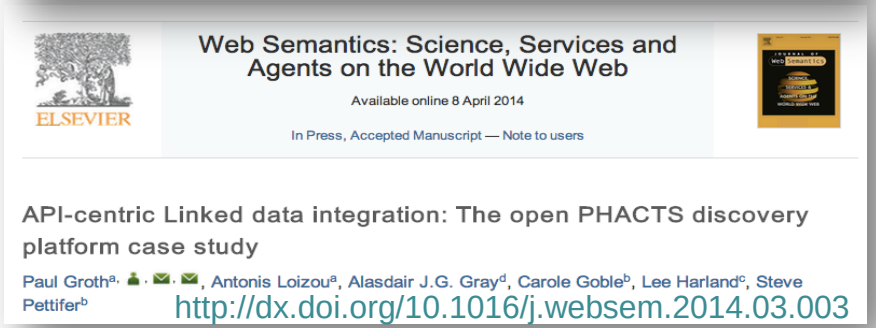
Volume 18, Issues 17–18, September 2013, Pages 843–852

Review

Scientific competency questions as the basis for semantically enriched open pharmacological space development

Kamal Azzaoui[1], Edgar Jacoby[14], Stefan Senger[2], Emiliano Cuadrado Rodríguez[3], Mabel Loza[3], Barbara Zdrazil[4], Marta Pinto[4], Antony J. Williams[5], Victor de la Torre[6], Jordi Mestres[7], Manuel Pastor[7], Olivier Taboureau[8], Matthias Rarey[9], Christine Chichester[10], Steve Pettifer[11], Niklas Blomberg[12,a], Lee Harland[13], Bryn Williams-Jones[13], Gerhard F. Ecker[4]

http://dx.doi.org/10.1016/j.drudis.2013.05.008



ELSEVIER

Web Semantics: Science, Services and Agents on the World Wide Web

Available online 8 April 2014

In Press, Accepted Manuscript — Note to users

API-centric Linked data integration: The open PHACTS discovery platform case study

Paul Groth[a], Antonis Loizou[a], Alasdair J.G. Gray[d], Carole Goble[b], Lee Harland[c], Steve Pettifer[b]

http://dx.doi.org/10.1016/j.websem.2014.03.003

Come and Talk to Us

- BDE Societal Workshops
  - Schedule on Website
- W3C Interest Groups set-up
  - Please Join and Engage!
  - https://www.w3.org/community/bde-health/join
- Stakeholders more than welcome!
- Get in touch, spread the word, ask us questions