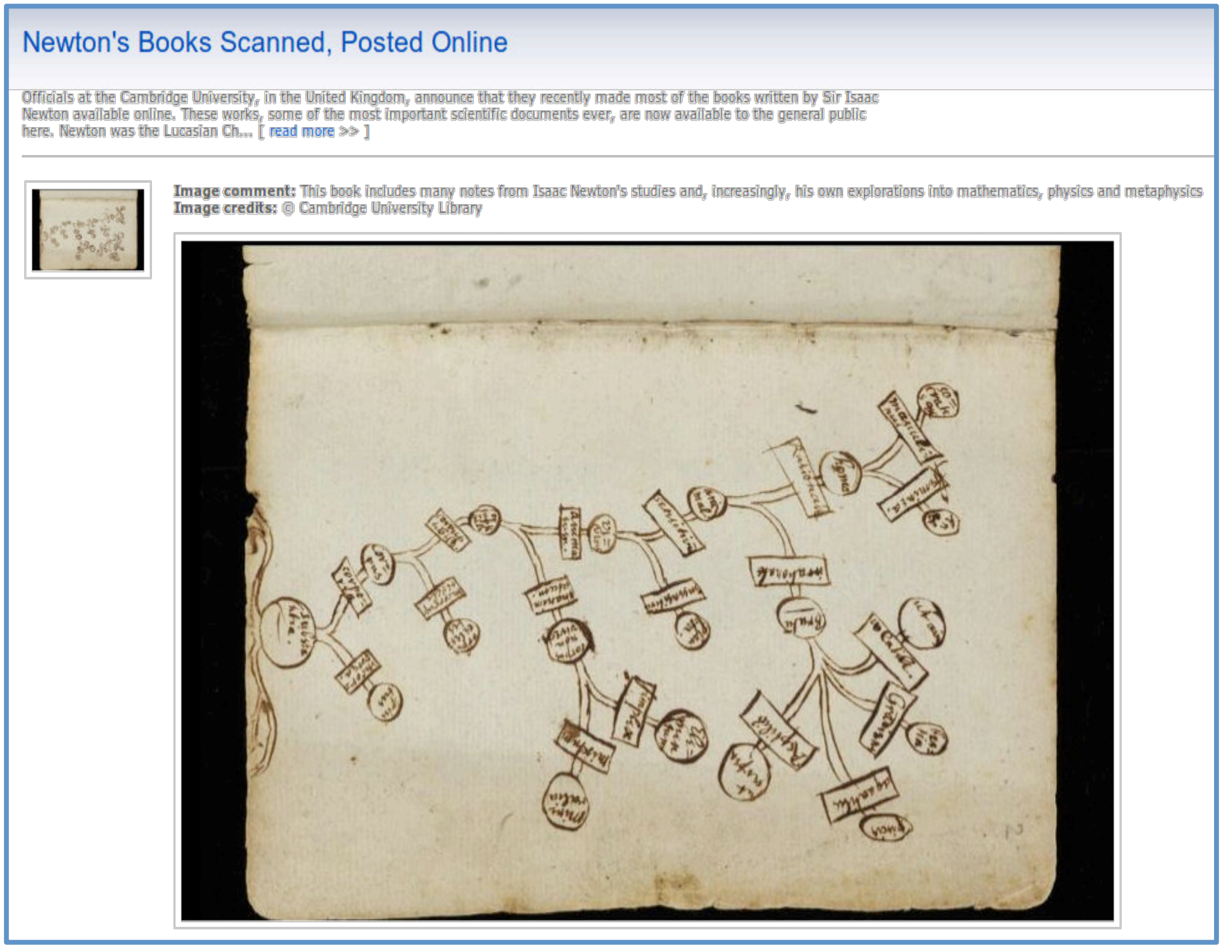
Newton's ideas and methods are preserved forever: how about yours?

Marco Roos, Kristina Hettne, Jun Zhao, Mark Thompson

Cloud and Workflows for Reproducible Bioinformatics

Shenzhen, December 19, 2012

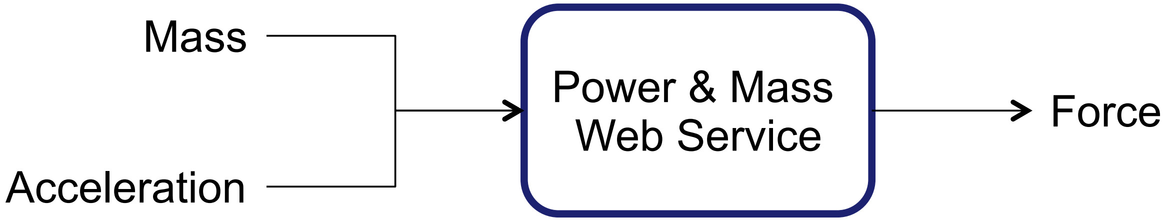


Newton's Books Scanned, Posted Online

Officials at the Cambridge University, in the United Kingdom, announce that they recently made most of the books written by Sir Isaac Newton available online. These works, some of the most important scientific documents ever, are now available to the general public here. Newton was the Lucasian Ch... [read more >>]

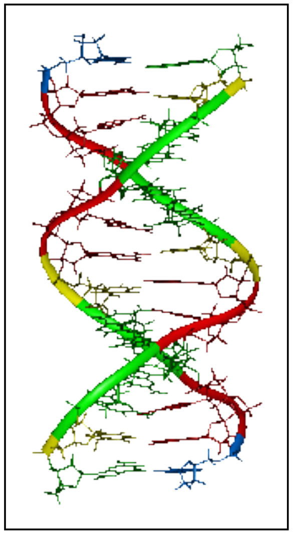
Image comment: This book includes many notes from Isaac Newton's studies and, increasingly, his own explorations into mathematics, physics and metaphysics
Image credits: © Cambridge University Library

Reproduced workflows



Case study

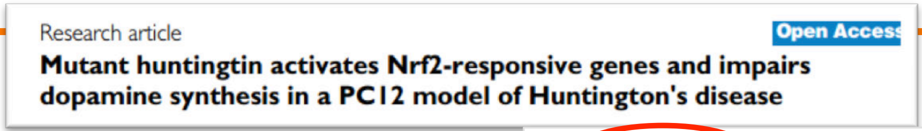
Bioinformatics analysis of Metabolic Syndrome
Kristina Hettne, Harish Dharuri



Genome Wide Association Studies

What is the genetic basis for the diseases associated with Metabolic Syndrome?

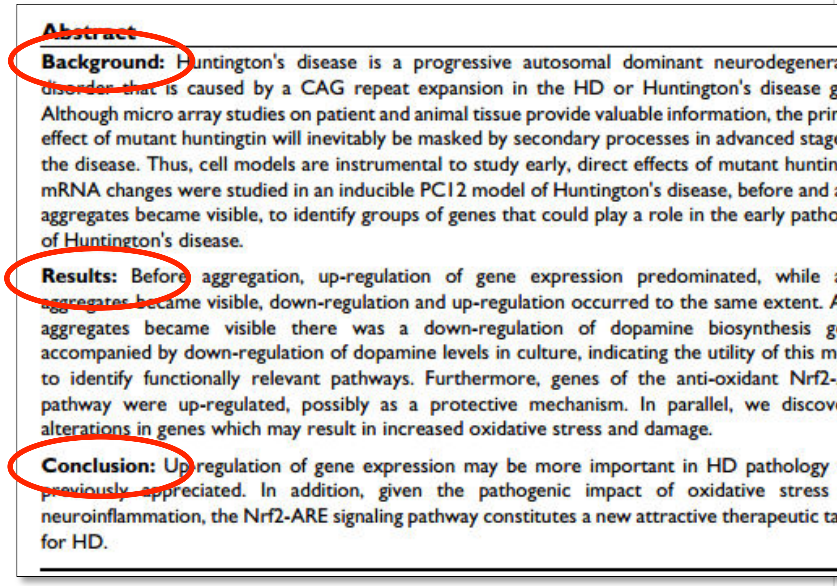
Reproducible Science



Research article

Open Access

Mutant huntingtin activates Nrf2-responsive genes and impairs dopamine synthesis in a PC12 model of Huntington's disease

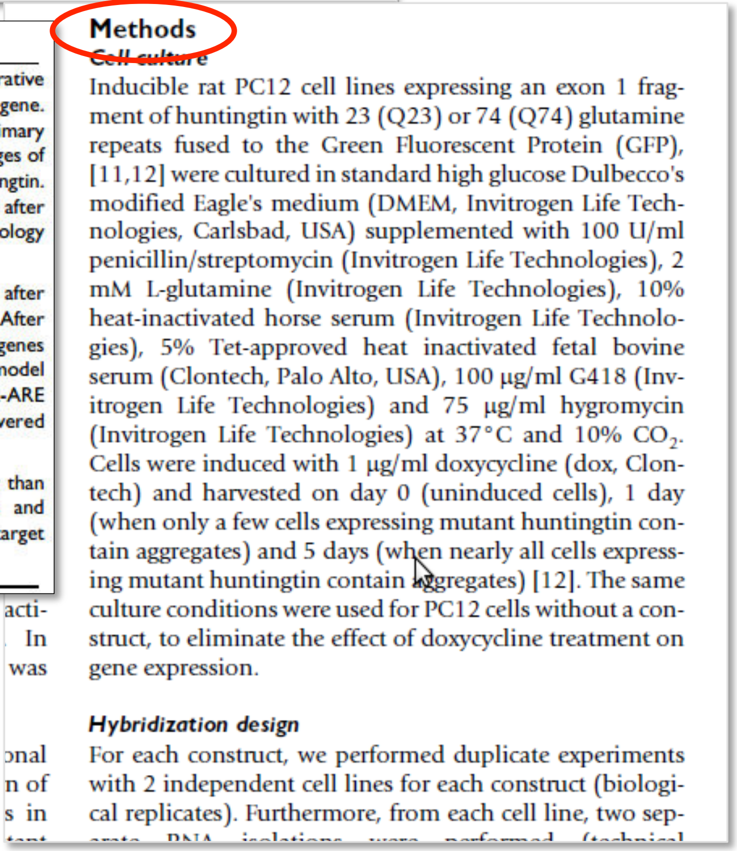


Abstract

Background: Huntington's disease is a progressive autosomal dominant neurodegenerative disorder that is caused by a CAG repeat expansion in the HD or Huntington's disease gene. Although micro array studies on patient and animal tissue provide valuable information, the primary effect of mutant huntingtin will inevitably be masked by secondary processes in advanced stages of the disease. Thus, cell models are instrumental to study early, direct effects of mutant huntingtin. mRNA changes were studied in an inducible PC12 model of Huntington's disease, before and after aggregates became visible, to identify groups of genes that could play a role in the early pathology of Huntington's disease.

Results: Before aggregation, up-regulation of gene expression predominated, while after aggregates became visible, down-regulation and up-regulation occurred to the same extent. After aggregates became visible there was a down-regulation of dopamine biosynthesis genes accompanied by down-regulation of dopamine levels in culture, indicating the utility of this model to identify functionally relevant pathways. Furthermore, genes of the anti-oxidant Nrf2-ARE pathway were up-regulated, possibly as a protective mechanism. In parallel, we discovered alterations in genes which may result in increased oxidative stress and damage.

Conclusion: Up-regulation of gene expression may be more important in HD pathology than previously appreciated. In addition, given the pathogenic impact of oxidative stress and neuroinflammation, the Nrf2-ARE signaling pathway constitutes a new attractive therapeutic target for HD.



acti-
In
was

onal
n of
s in

Methods

Cell culture

Inducible rat PC12 cell lines expressing an exon 1 fragment of huntingtin with 23 (Q23) or 74 (Q74) glutamine repeats fused to the Green Fluorescent Protein (GFP), [11,12] were cultured in standard high glucose Dulbecco's modified Eagle's medium (DMEM, Invitrogen Life Technologies, Carlsbad, USA) supplemented with 100 U/ml penicillin/streptomycin (Invitrogen Life Technologies), 2 mM L-glutamine (Invitrogen Life Technologies), 10% heat-inactivated horse serum (Invitrogen Life Technologies), 5% Tet-approved heat inactivated fetal bovine serum (Clontech, Palo Alto, USA), 100 μg/ml G418 (Invitrogen Life Technologies) and 75 μg/ml hygromycin (Invitrogen Life Technologies) at 37°C and 10% CO_2. Cells were induced with 1 μg/ml doxycycline (dox, Clontech) and harvested on day 0 (uninduced cells), 1 day (when only a few cells expressing mutant huntingtin contain aggregates) and 5 days (when nearly all cells expressing mutant huntingtin contain aggregates) [12]. The same culture conditions were used for PC12 cells without a construct, to eliminate the effect of doxycycline treatment on gene expression.

Hybridization design

For each construct, we performed duplicate experiments with 2 independent cell lines for each construct (biological replicates). Furthermore, from each cell line, two separate RNA isolations were performed (technical

Preservation for the
wet laboratory
scientist

From Van Roon-Mom *et al.*, BMC Molecular Biology 2008
doi: 10.1186/1471-2199-9-84.

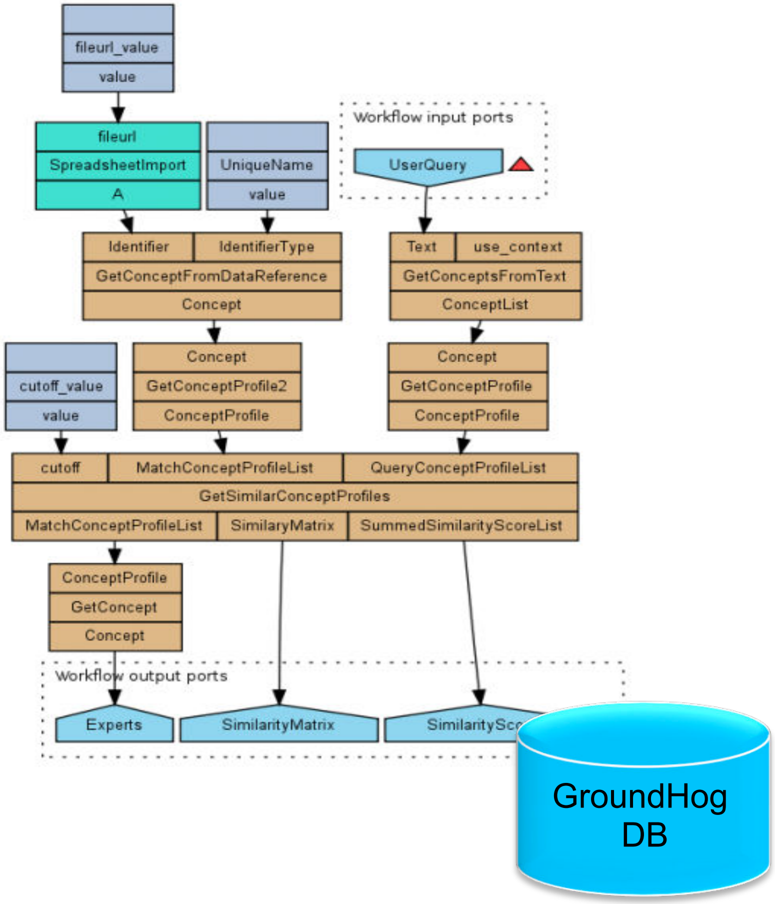
Reproducible Science?

What is the digital equivalent?

Is it equally good?

Can we do better?

- *or worse?*



Reproduced from Jelier *et al.*, Schuemie *et al.*, Hettne *et al.*, Haagen *et al.*, http://biosemantics.org , *myExperiment.org/workflows/2197*

Reproducible Science

What is our incentive?

Nobility

Good Reproducible Science

Greater Good

Serve the public

Reproducible Science

What is our incentive?

Fame and Glory

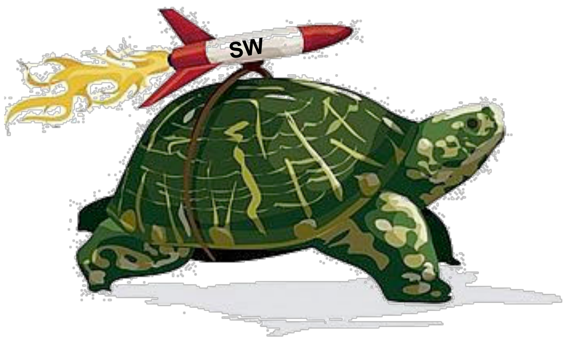
Getting on with it...

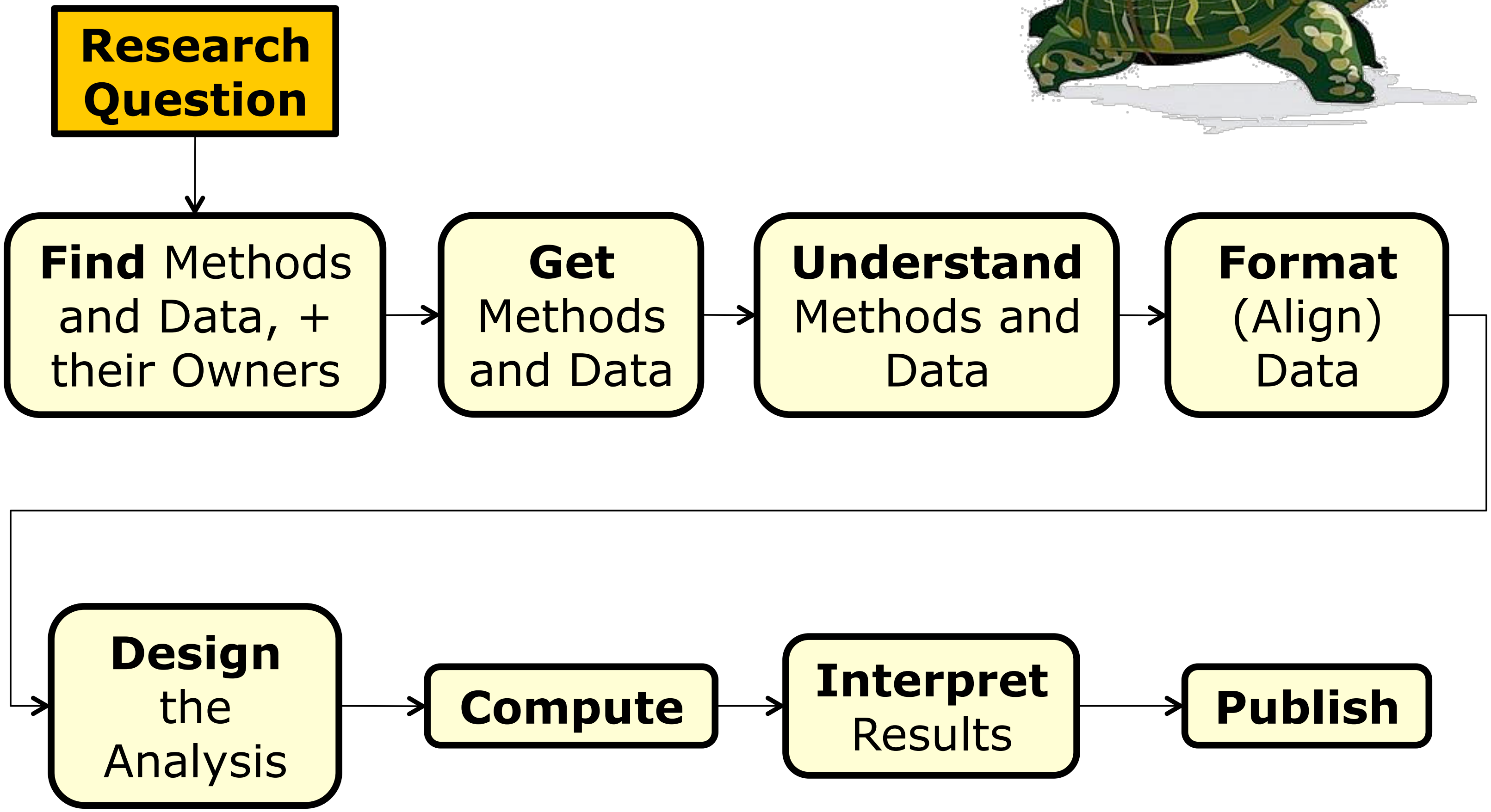
CHALLENGE

Stimulate preservation and reproducibility while speeding up the research process

Enhance the research cycle

What slows us down?





Research Question

Find Methods and Data, + their Owners

Get Methods and Data

Understand Methods and Data

Format (Align) Data

Design the Analysis

Compute

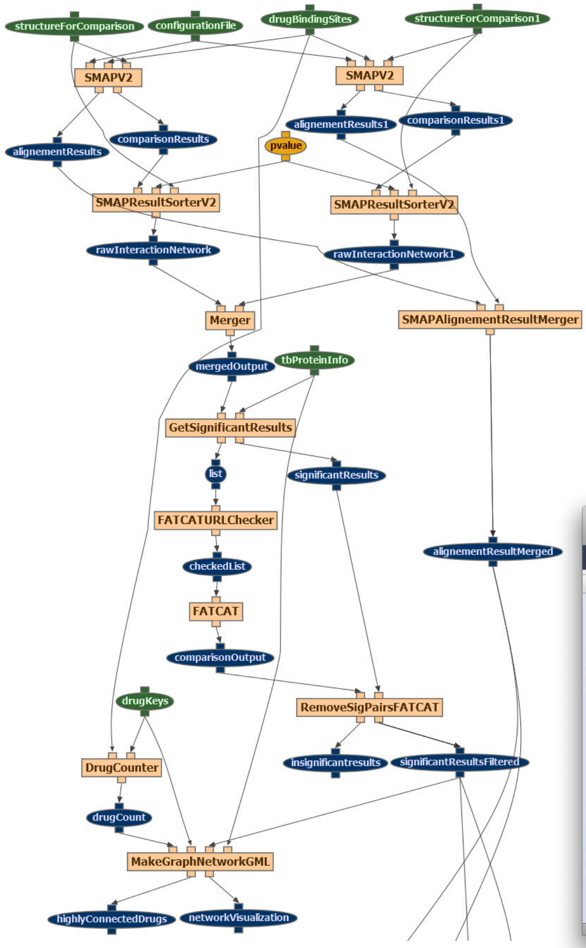
Interpret Results

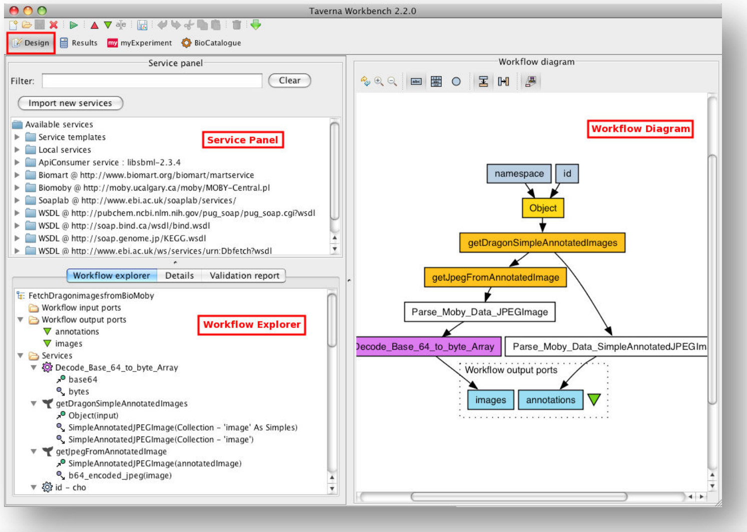
Publish

Bottlenecks

- Loosing track of what you did
- Messy storage
- Preparing material for a publication
- Understanding the computational procedure
- Communication with (non-technical) colleagues
- Keeping tools working
- Getting credit for digital results outside of traditional publications

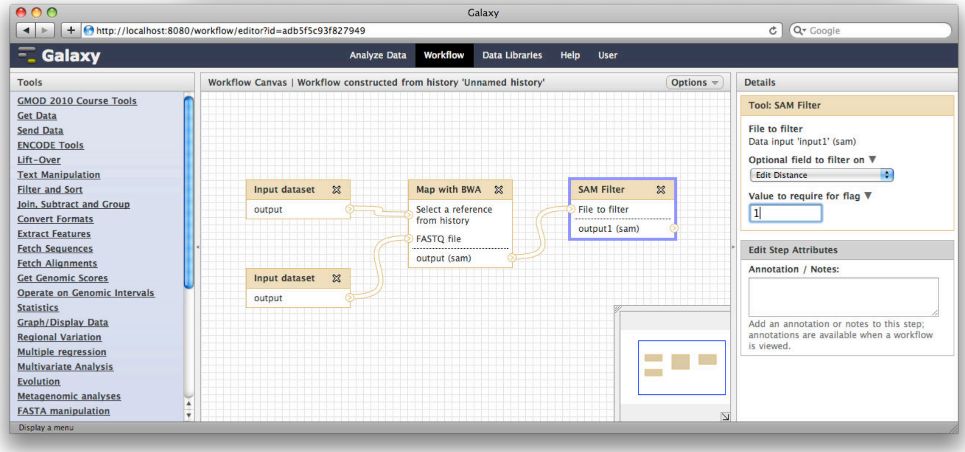
Getting on with workflows





- Available services
  - Service templates
  - Local services
  - ApiConsumer service : libsbml-2.3.4
  - Biomart @ http://www.biomart.org/biomart/martservice
  - Biomoby @ http://moby.ucalgary.ca/moby/MOBY-Central.pl
  - Soaplab @ http://www.ebi.ac.uk/soaplab/services/
  - WSDL @ http://pubchem.ncbi.nlm.nih.gov/pug_soap/pug_soap.cgi?wsdl
  - WSDL @ http://soap.bind.ca/wsdl/bind.wsdl
  - WSDL @ http://soap.genome.jp/KEGG.wsdl
  - WSDL @ http://www.ebi.ac.uk/ws/services/urn:Dbfetch?wsdl

- FetchDragonimagesfromBioMoby
  - Workflow input ports
  - Workflow output ports
    - annotations
    - images
  - Services
    - Decode_Base_64_to_byte_Array
      - base64
      - bytes
    - getDragonSimpleAnnotatedImages
      - Object(input)
      - SimpleAnnotatedJPEGImage(Collection - 'image' As Simples)
      - SimpleAnnotatedJPEGImage(Collection - 'image')
    - getJpegFromAnnotatedImage
      - SimpleAnnotatedJPEGImage(annotatedImage)
      - b64_encoded_jpeg(image)
    - id - cho



Tool: SAM Filter

File to filter: Data input 'input1' (sam)

Optional field to filter on: Edit Distance

Value to require for flag: 1

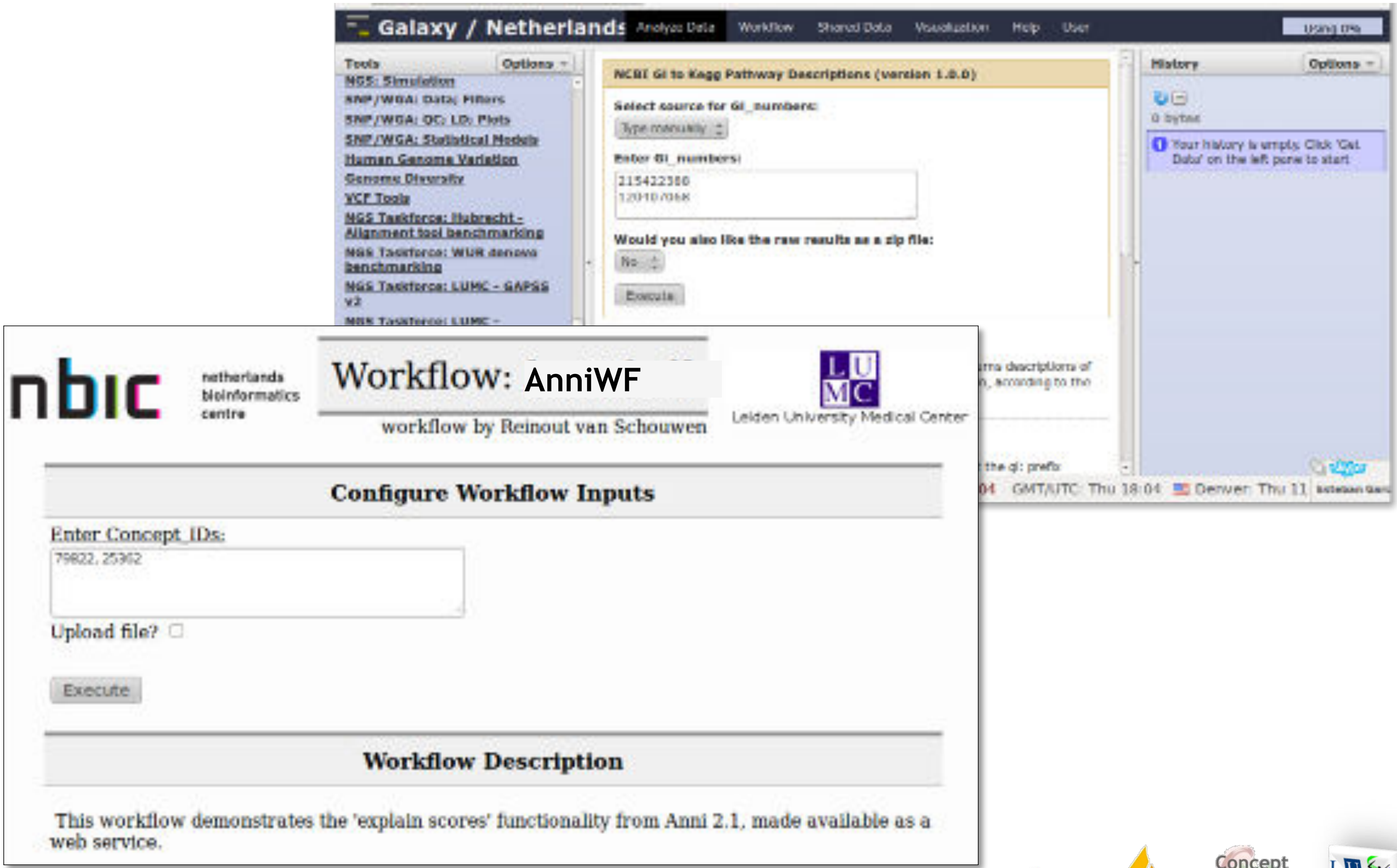
Annotation / Notes:

Add an annotation or notes to this step; annotations are available when a workflow is viewed.

Workflow 4Ever

~~Monolithic Tool~~ →
Web Services → Workflows → (Web) Tool

Example: Anni 2.0 → Anni workflows



NCBI GI to Kegg Pathway Descriptions (version 1.0.0)

Select source for GI_numbers:
Type manually

Enter GI_numbers:
215422380
120107068

Would you also like the raw results as a zip file:
No

Execute

0 bytes

Your history is empty. Click 'Get Data' on the left pane to start

nbic
netherlands bioinformatics centre
Leiden University Medical Center

Workflow: AnniWF

workflow by Reinout van Schouwen

Configure Workflow Inputs

Enter Concept IDs:
79822, 25362

Upload file? □

Execute

Workflow Description

This workflow demonstrates the 'explain scores' functionality from Anni 2.1, made available as a web service.

http://workflow.biosemantics.org/t2web/workflow/2725

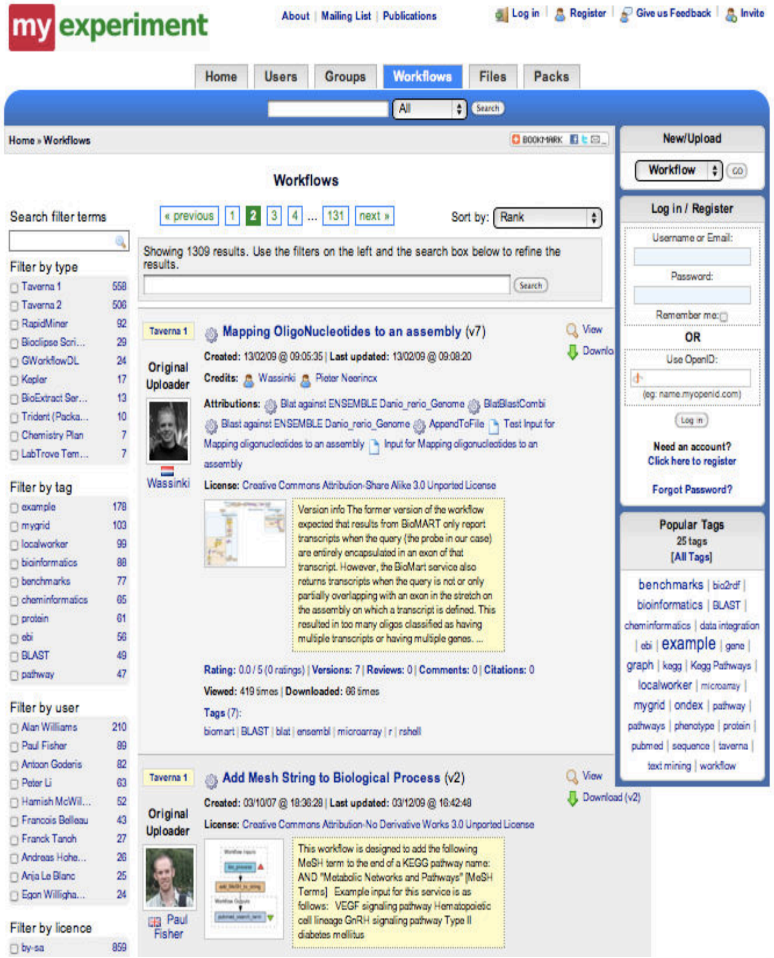
nbic Open PHACTS Concept Web Alliance LUMC

Digital Repository

myExperiment.org

The recipes store

- Find workflows
- Share workflows & files
- Find people
- Build communities
- Publish packages
- Tag workflows
- Score, rate, comment



Workflows

« previous | 1 | 2 | 3 | 4 | ... | 131 | next »

Sort by: Rank

Search filter terms

Filter by type

□ Taverna 1	558
□ Taverna 2	506
□ RapidMiner	92
□ Bioclipse Scri...	29
□ GWorkflowDL	24
□ Kepler	17
□ BioExtract Ser...	13
□ Trident (Packa...	10
□ Chemistry Plan	7
□ LabTrove Tem...	7

Filter by tag

□ example	178
□ mygrid	103
□ localworker	99
□ bioinformatics	88
□ benchmarks	77
□ cheminformatics	65
□ protein	61
□ ebi	56
□ BLAST	49
□ pathway	47

Filter by user

□ Alan Williams	210
□ Paul Fisher	89
□ Antoon Goderis	82
□ Peter Li	63
□ Hamish McWil...	52
□ Francois Belleau	43
□ Franck Tanoh	27
□ Andreas Hohe...	26
□ Anja Le Blanc	25
□ Egon Willigha...	24

Filter by licence

□ by-sa	859

Showing 1309 results. Use the filters on the left and the search box below to refine the results.

Search

Mapping OligoNucleotides to an assembly (v7)

Created: 13/02/09 @ 09:05:35 | **Last updated:** 13/02/09 @ 09:08:20

Credits: Wassinki Pieter Neerincx

Attributions: Blat against ENSEMBLE Danio_rerio_Genome BlatBlastCombi Blast against ENSEMBLE Danio_rerio_Genome AppendToFile Test Input for Mapping oligonucleotides to an assembly Input for Mapping oligonucleotides to an assembly

License: Creative Commons Attribution-Share Alike 3.0 Unported License

Version info The former version of the workflow expected that results from BioMART only report transcripts when the query (the probe in our case) are entirely encapsulated in an exon of that transcript. However, the BioMart service also returns transcripts when the query is not or only partially overlapping with an exon in the stretch on the assembly on which a transcript is defined. This resulted in too many oligos classified as having multiple transcripts or having multiple genes. ...

Rating: 0.0 / 5 (0 ratings) | **Versions:** 7 | **Reviews:** 0 | **Comments:** 0 | **Citations:** 0

Viewed: 419 times | **Downloaded:** 66 times

Tags (7):

biomart | BLAST | blat | ensembl | microarray | r | rshell

Add Mesh String to Biological Process (v2)

Created: 03/10/07 @ 18:36:28 | **Last updated:** 03/12/09 @ 16:42:48

License: Creative Commons Attribution-No Derivative Works 3.0 Unported License

This workflow is designed to add the following MeSH term to the end of a KEGG pathway name: AND "Metabolic Networks and Pathways" [MeSH Terms] Example input for this service is as follows: VEGF signaling pathway Hematopoietic cell lineage GnRH signaling pathway Type II diabetes mellitus

Log in / Register

Username or Email:

Password:

Remember me: □

OR

Use OpenID:

(eg: name.myopenid.com)

Log in

Need an account?
Click here to register

Forgot Password?

Popular Tags
25 tags
[All Tags]

benchmarks | bio2rdf | bioinformatics | BLAST | cheminformatics | data integration | ebi | example | gene | graph | kegg | Kegg Pathways | localworker | microarray | mygrid | ondex | pathway | pathways | phenotype | protein | pubmed | sequence | taverna | text mining | workflow

Instructions for workflow authors

10 Best Practices for creating workflows

1. Make a sketch workflow
2. Use modules
3. Think about the output
4. Provide example inputs and outputs
5. Annotate
6. Test execution from outside local environment
7. Choose services carefully
8. Reuse existing workflows
9. Advertise
10. Maintain

10/10

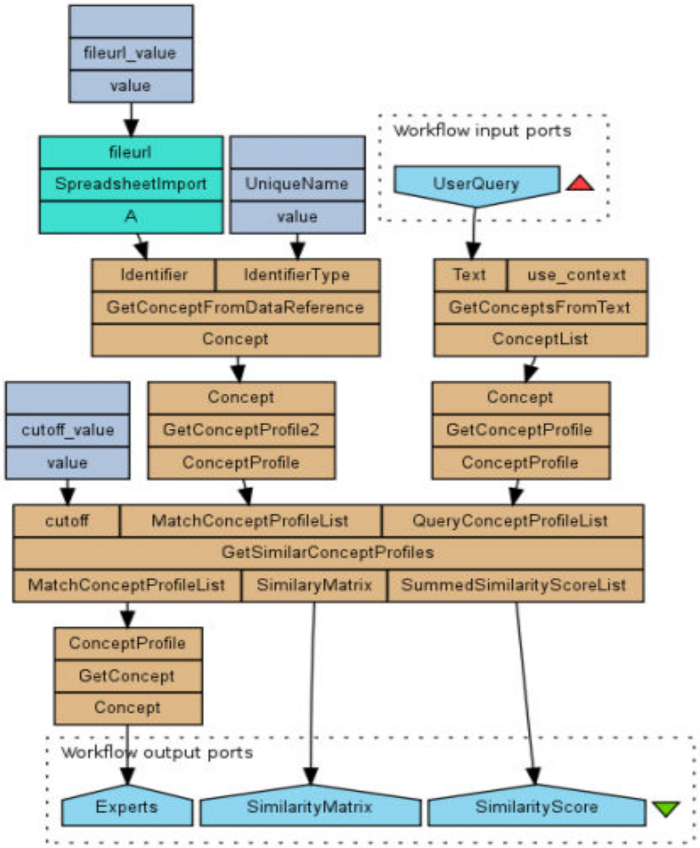
Reproducible Science

Is a workflow sufficient?

Useful Preservation = Understandable Objects

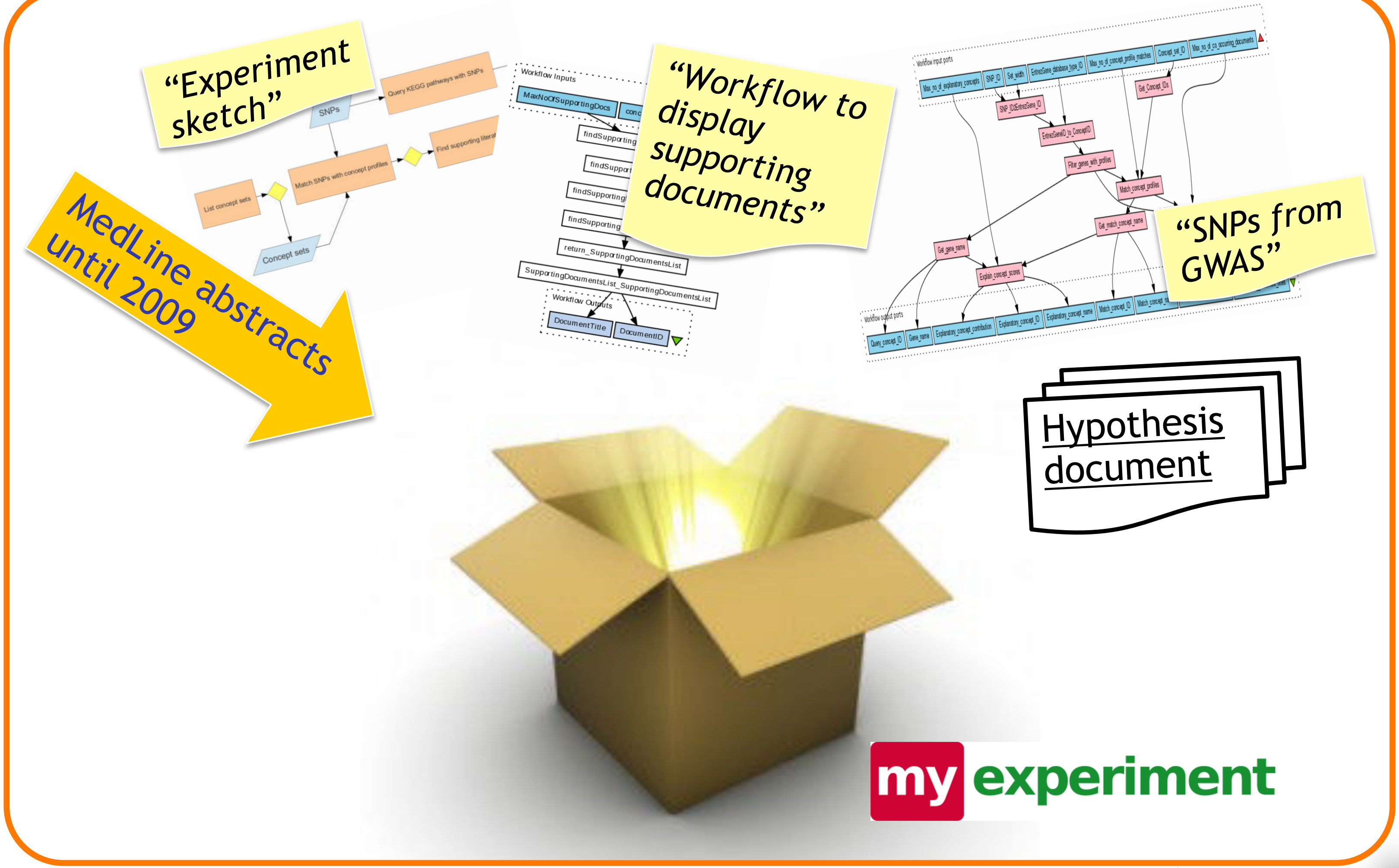
Reproduce, Reuse, Repurpose, Repair, ...

What is this doing?



Reproduced from Jelier *et al.*, Schuemie *et al.*, Hettne *et al.*, Haagen *et al.*, http://biosemantics.org , *myExperiment.org/workflows/2197*

Useful preservation 1
myExperiment Packs



“Experiment sketch”

MedLine abstracts until 2009

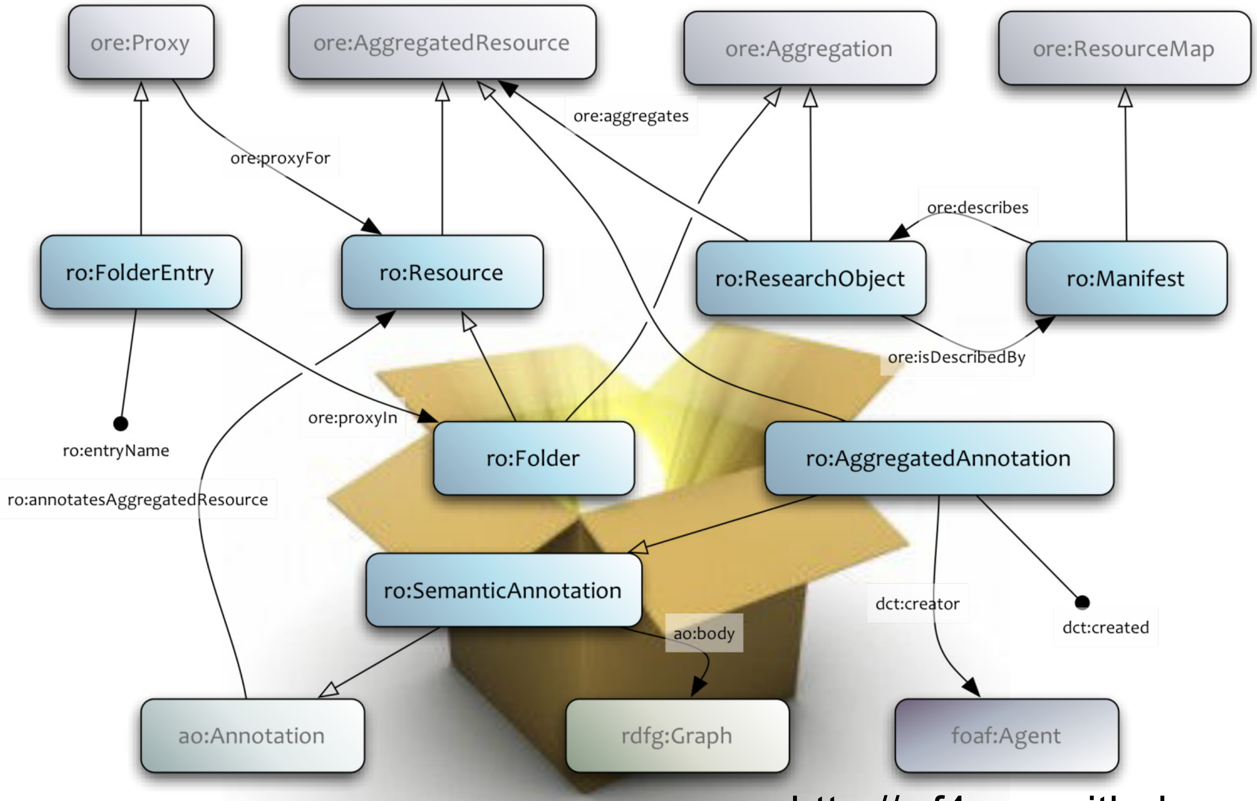
“Workflow to display supporting documents”

“SNPs from GWAS”

Hypothesis document

Useful preservation
Research Object Model

Research Object Model
Aggregation and Annotation Model for Digital Methods



http://wf4ever.github.com/ro/

Research Object (RO) Model

RO = ORE + AO + vocabularies

Object Re-use and Exchange (OAI-ORE)

- Describes aggregations of resources:
- data, metadata, papers, *etc*.

Annotation Ontology (AO)

- Associates RDF metadata descriptions with resources

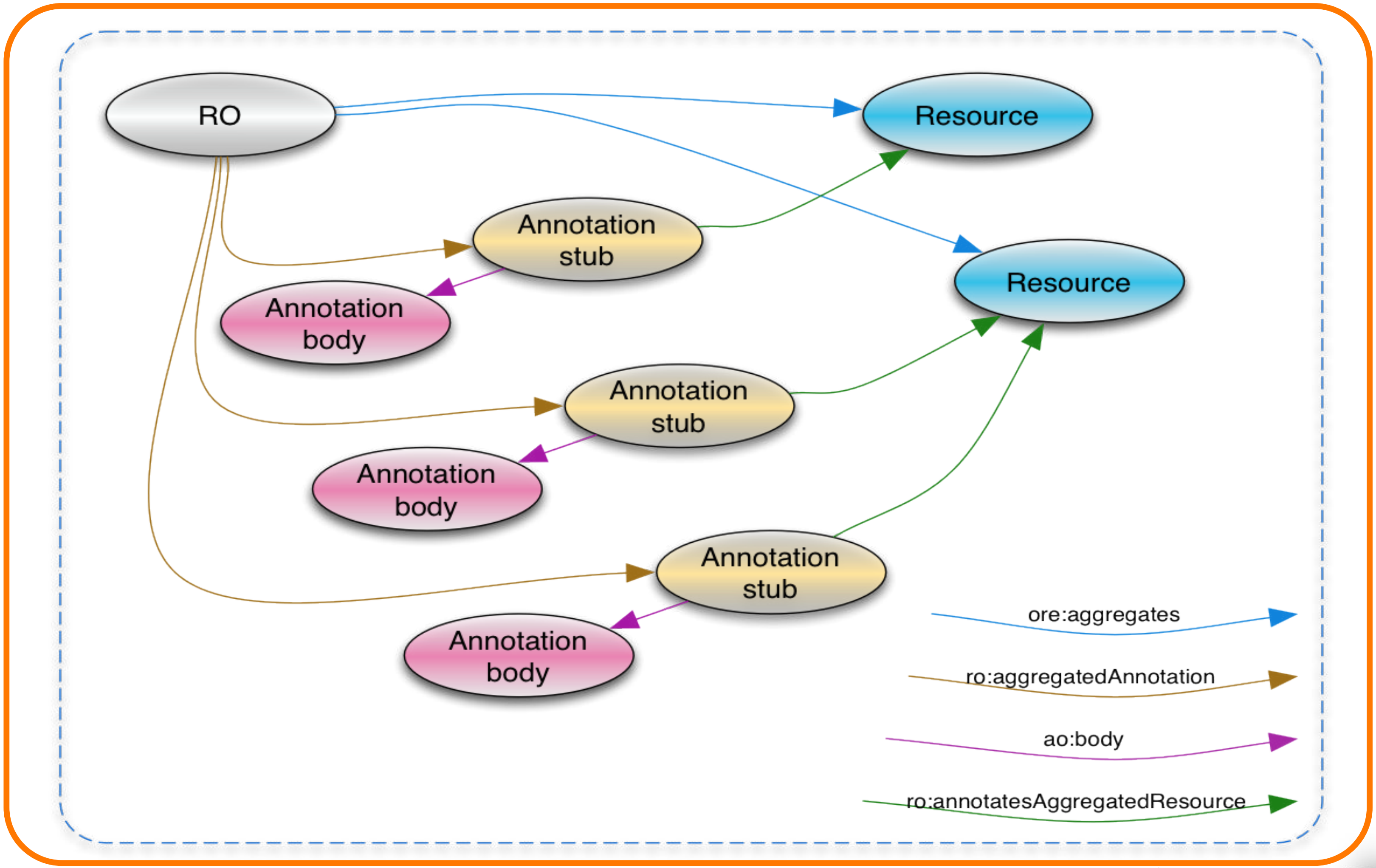
Generic and domain-specific vocabularies

- Used in annotation bodies to provide information about resources (types, dependencies, descriptions, *etc*.)

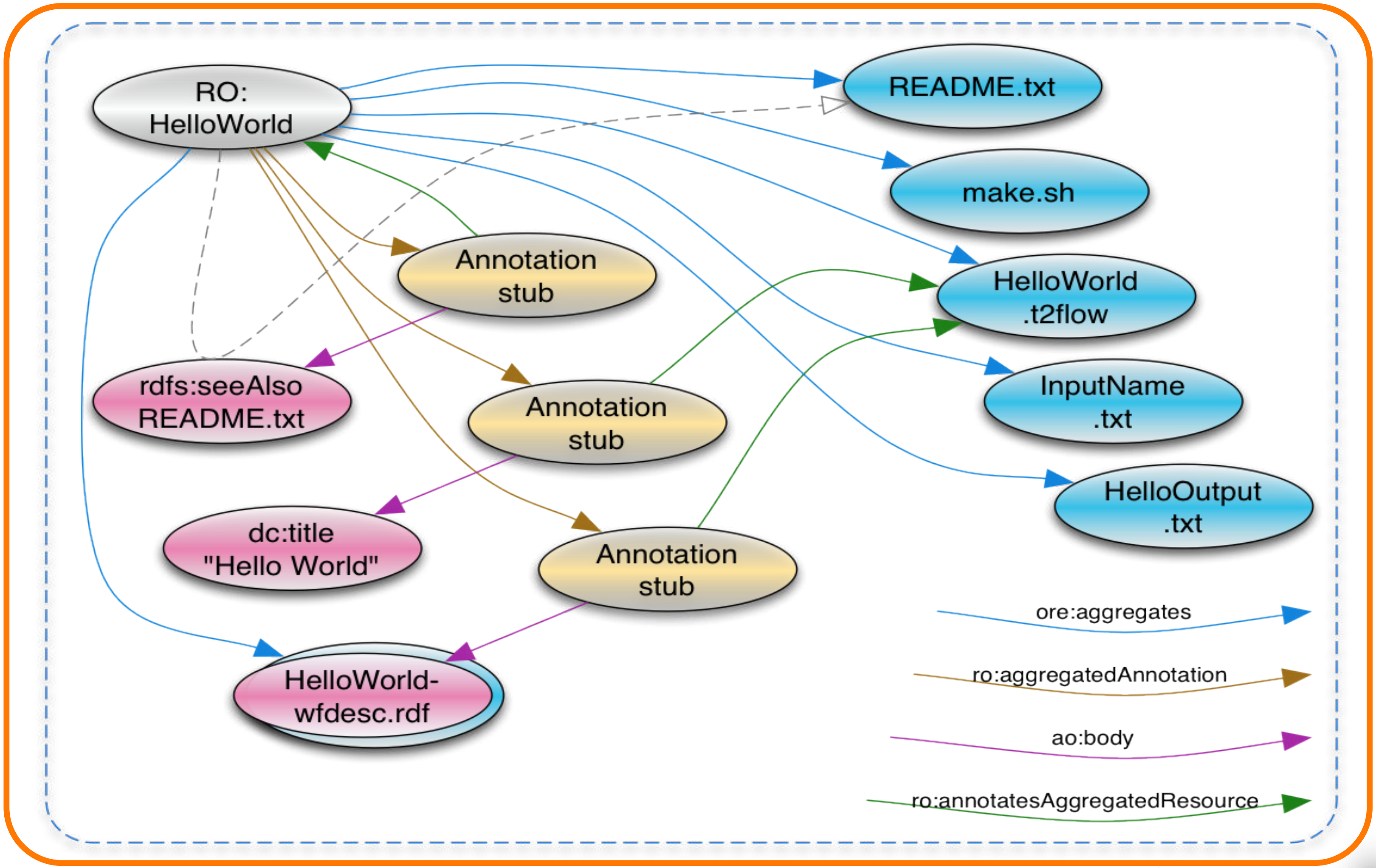
Builds on RDF, leading to RDF as a natural implementation choice

Model specification: http://wf4ever.github.com/ro/

Research Object Model



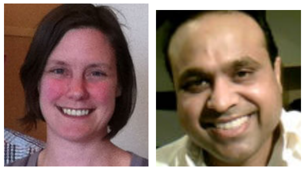
Research Object: “Hello World”



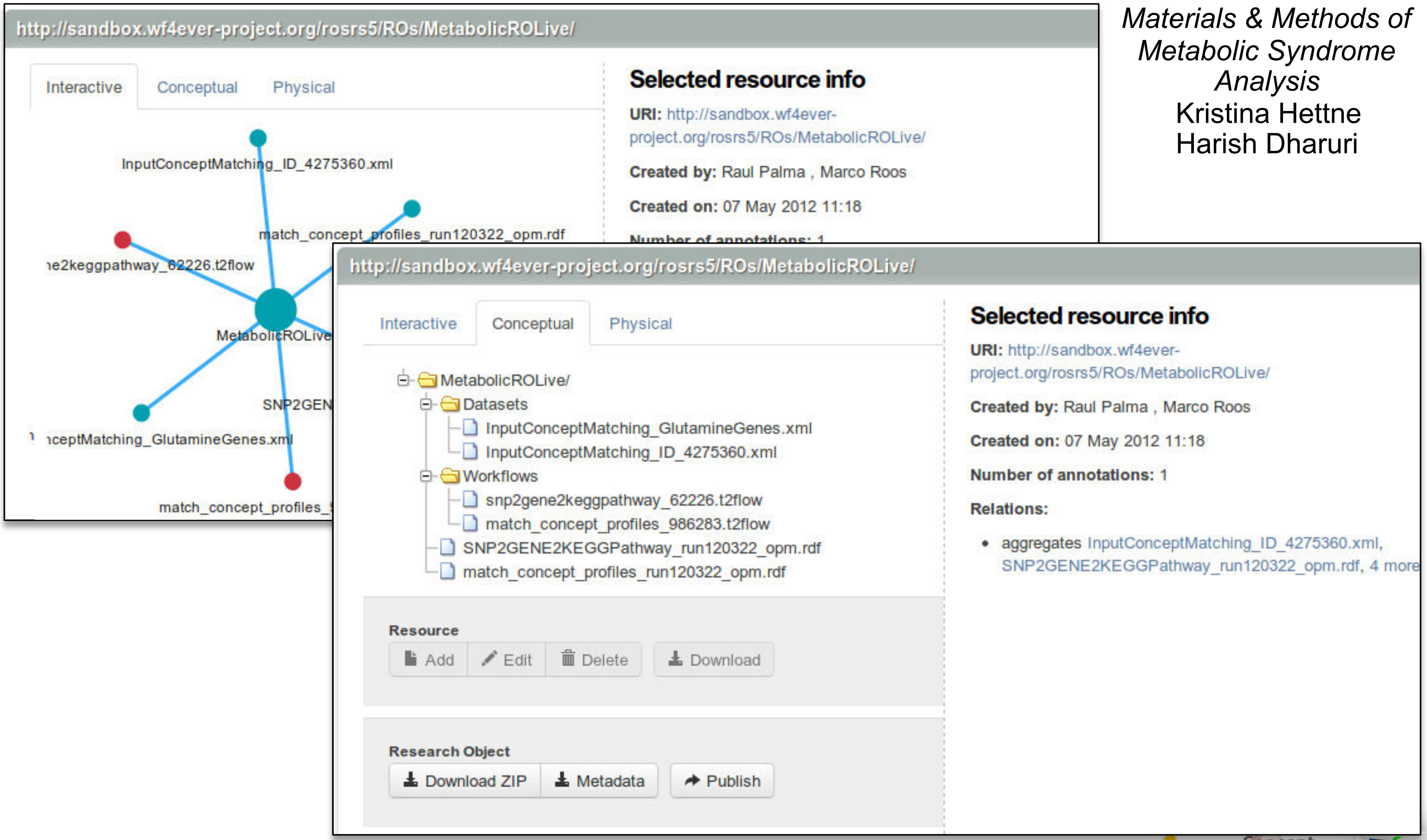
https://github.com/wf4ever/ro-catalogue/tree/master/v0.1/HelloWorld

Help organize the materials and methods of computational analysis

Research Object Portal



Materials & Methods of Metabolic Syndrome Analysis
Kristina Hettne
Harish Dharuri



http://sandbox.wf4ever-project.org/rosrs5/ROs/MetabolicROLive/

Interactive
Conceptual
Physical
InputConceptMatching_ID_4275360.xml
match_concept_profiles_run120322_opm.rdf
ne2keggpathway_62226.t2flow
MetabolicROLive
SNP2GEN
nceptMatching_GlutamineGenes.xml
match_concept_profiles_

Selected resource info

URI: http://sandbox.wf4ever-project.org/rosrs5/ROs/MetabolicROLive/

Created by: Raul Palma , Marco Roos

Created on: 07 May 2012 11:18

Number of annotations: 1

http://sandbox.wf4ever-project.org/rosrs5/ROs/MetabolicROLive/

Interactive
Conceptual
Physical

```
MetabolicROLive/
  Datasets
    InputConceptMatching_GlutamineGenes.xml
    InputConceptMatching_ID_4275360.xml
  Workflows
    snp2gene2keggpathway_62226.t2flow
    match_concept_profiles_986283.t2flow
  SNP2GENE2KEGGPathway_run120322_opm.rdf
  match_concept_profiles_run120322_opm.rdf
```

Selected resource info

URI: http://sandbox.wf4ever-project.org/rosrs5/ROs/MetabolicROLive/

Created by: Raul Palma , Marco Roos

Created on: 07 May 2012 11:18

Number of annotations: 1

Relations:

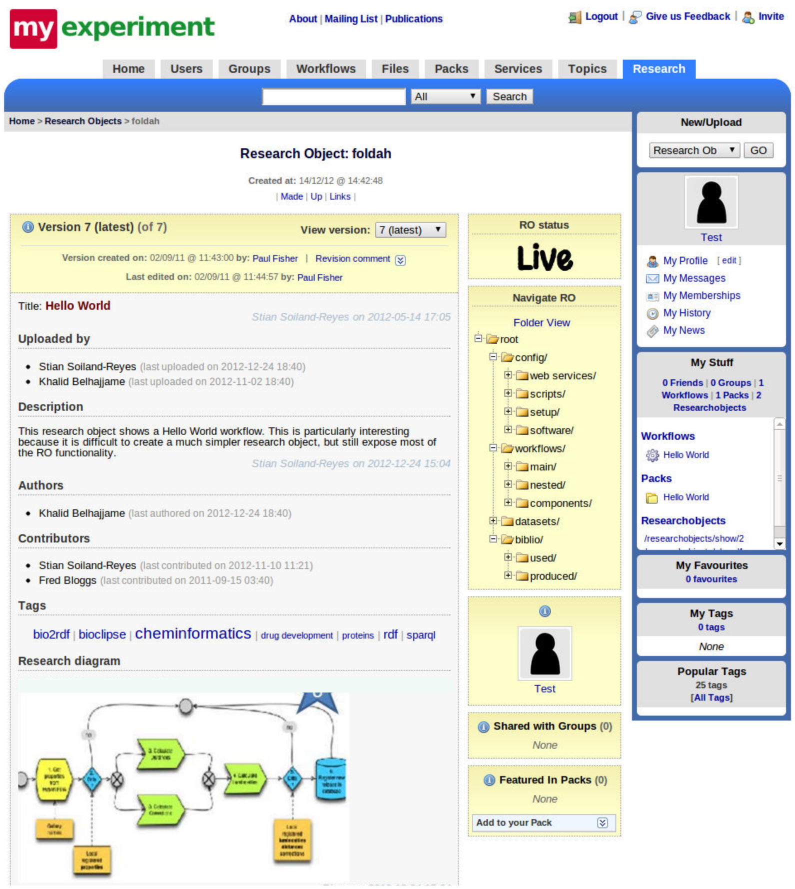
- aggregates InputConceptMatching_ID_4275360.xml, SNP2GENE2KEGGPathway_run120322_opm.rdf, 4 more

http://sandbox.wf4ever-project.org/portal

Expected on myExperiment

Research Objects inside!

- Packs more prominent
- Start a pack when you upload a workflow
- Upload wizards, pack management, export
- Checklists, automated star ratings
- Add workflow runs and example data
- Sticky annotations



Home > Research Objects > foldah

Research Object: foldah

Created at: 14/12/12 @ 14:42:48

| Made | Up | Links |

Version 7 (latest) (of 7) View version: 7 (latest)

Version created on: 02/09/11 @ 11:43:00 by: Paul Fisher | Revision comment

Last edited on: 02/09/11 @ 11:44:57 by: Paul Fisher

Title: Hello World

Stian Soiland-Reyes on 2012-05-14 17:05

Uploaded by

- Stian Soiland-Reyes (last uploaded on 2012-12-24 18:40)
- Khalid Belhajjame (last uploaded on 2012-11-02 18:40)

Description

This research object shows a Hello World workflow. This is particularly interesting because it is difficult to create a much simpler research object, but still expose most of the RO functionality.

Stian Soiland-Reyes on 2012-12-24 15:04

Authors

- Khalid Belhajjame (last authored on 2012-12-24 18:40)

Contributors

- Stian Soiland-Reyes (last contributed on 2012-11-10 11:21)
- Fred Bloggs (last contributed on 2011-09-15 03:40)

Tags

bio2rdf | bioclipse | cheminformatics | drug development | proteins | rdf | sparql

Research diagram

RO status

Live

Navigate RO

Folder View

- root
  - config/
    - web services/
    - scripts/
    - setup/
    - software/
  - workflows/
    - main/
    - nested/
    - components/
  - datasets/
  - biblio/
    - used/
    - produced/

Test

Shared with Groups (0)

None

Featured In Packs (0)

None

Add to your Pack

My Stuff

0 Friends | 0 Groups | 1 Workflows | 1 Packs | 2 Researchobjects

Workflows

Hello World

Packs

Hello World

Researchobjects

/researchobjects/show/2

My Favourites

0 favourites

My Tags

0 tags

None

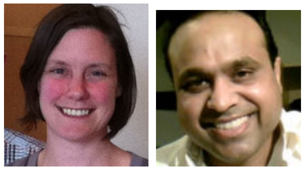
Popular Tags

25 tags

[All Tags]

RO-enabled myExperiment mockup

Fame and Glory



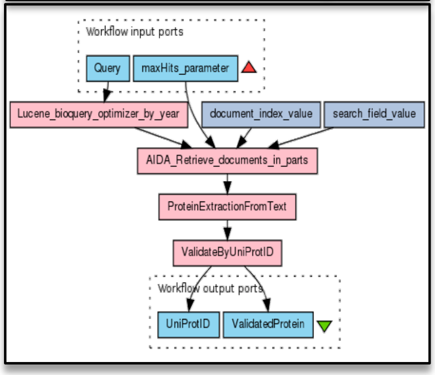
It was me, me, me!

What I found

HDAC1 *interacts with* Parvb

Discovered by: me
Published by: me

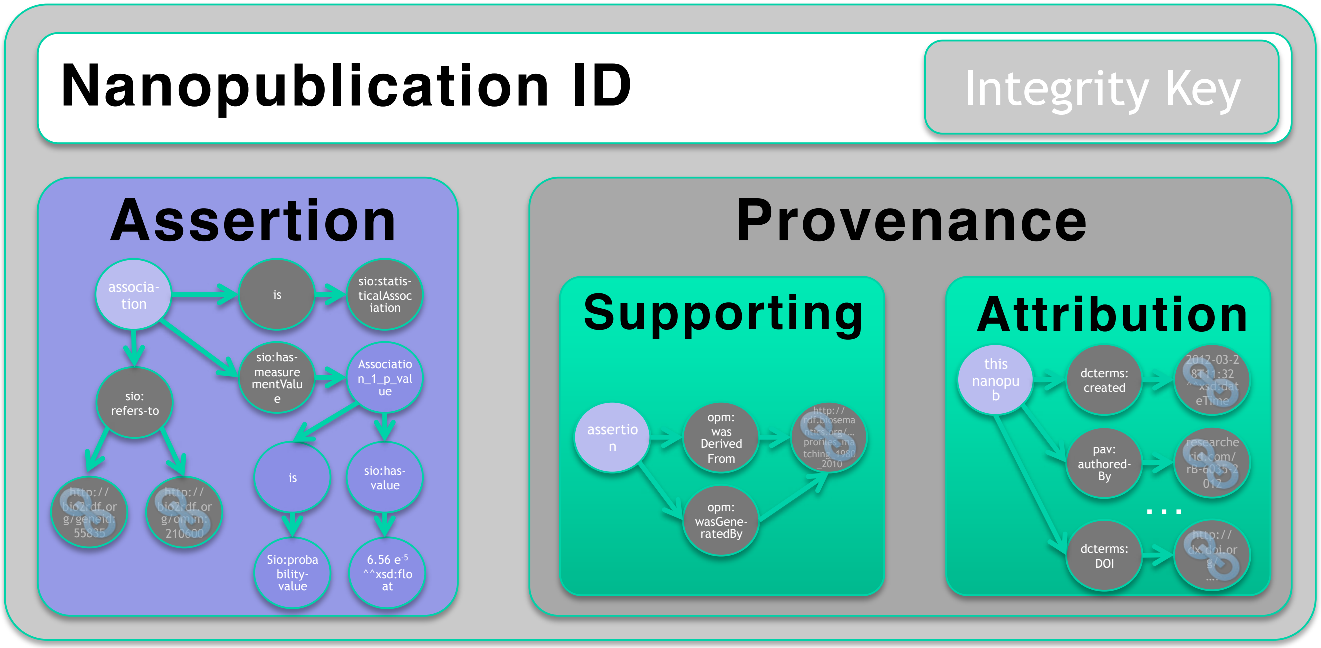
Research Object



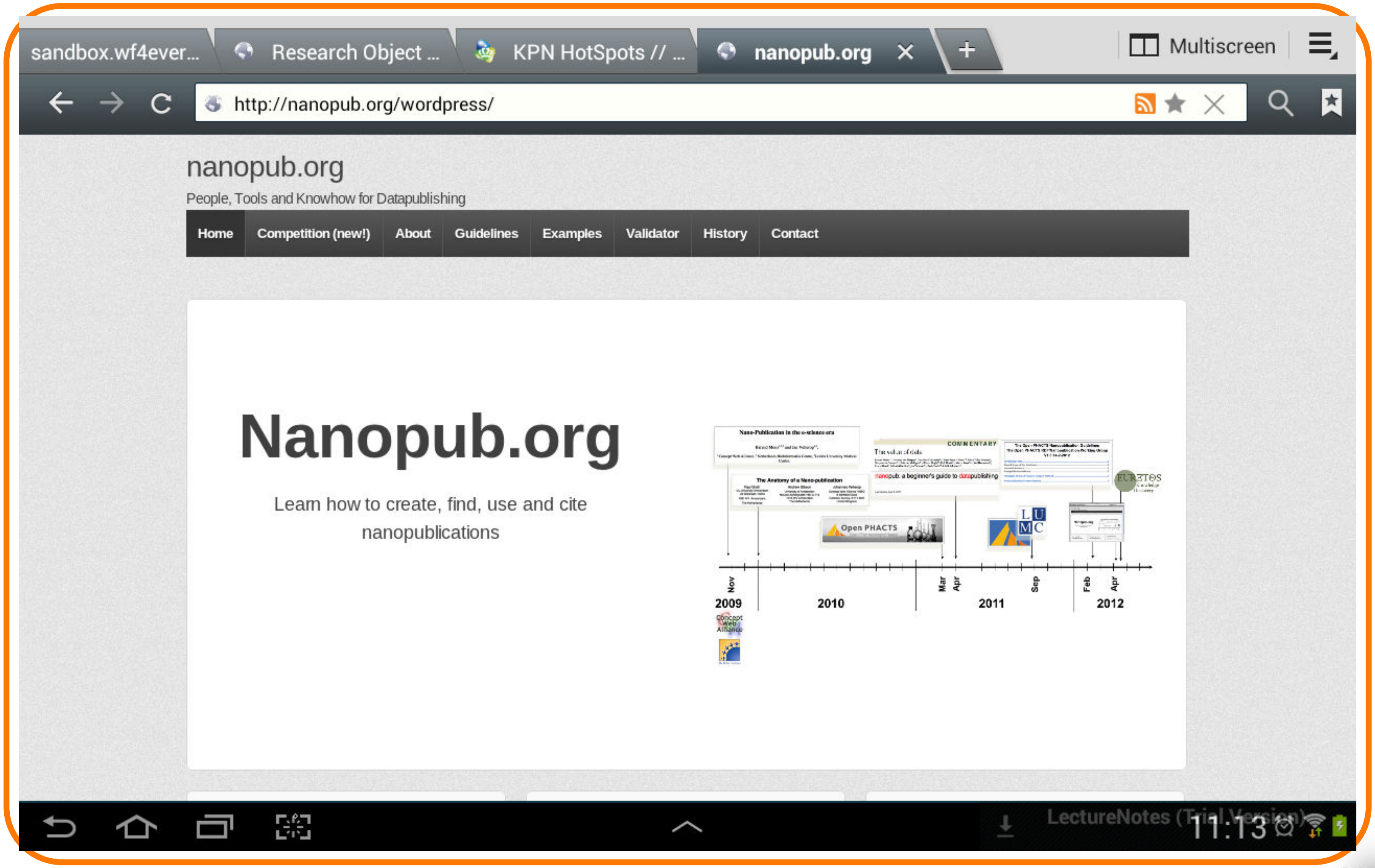
How I found it

Nanopublication Model

Getting credit for digital results



Nanopub.org



nanopub.org

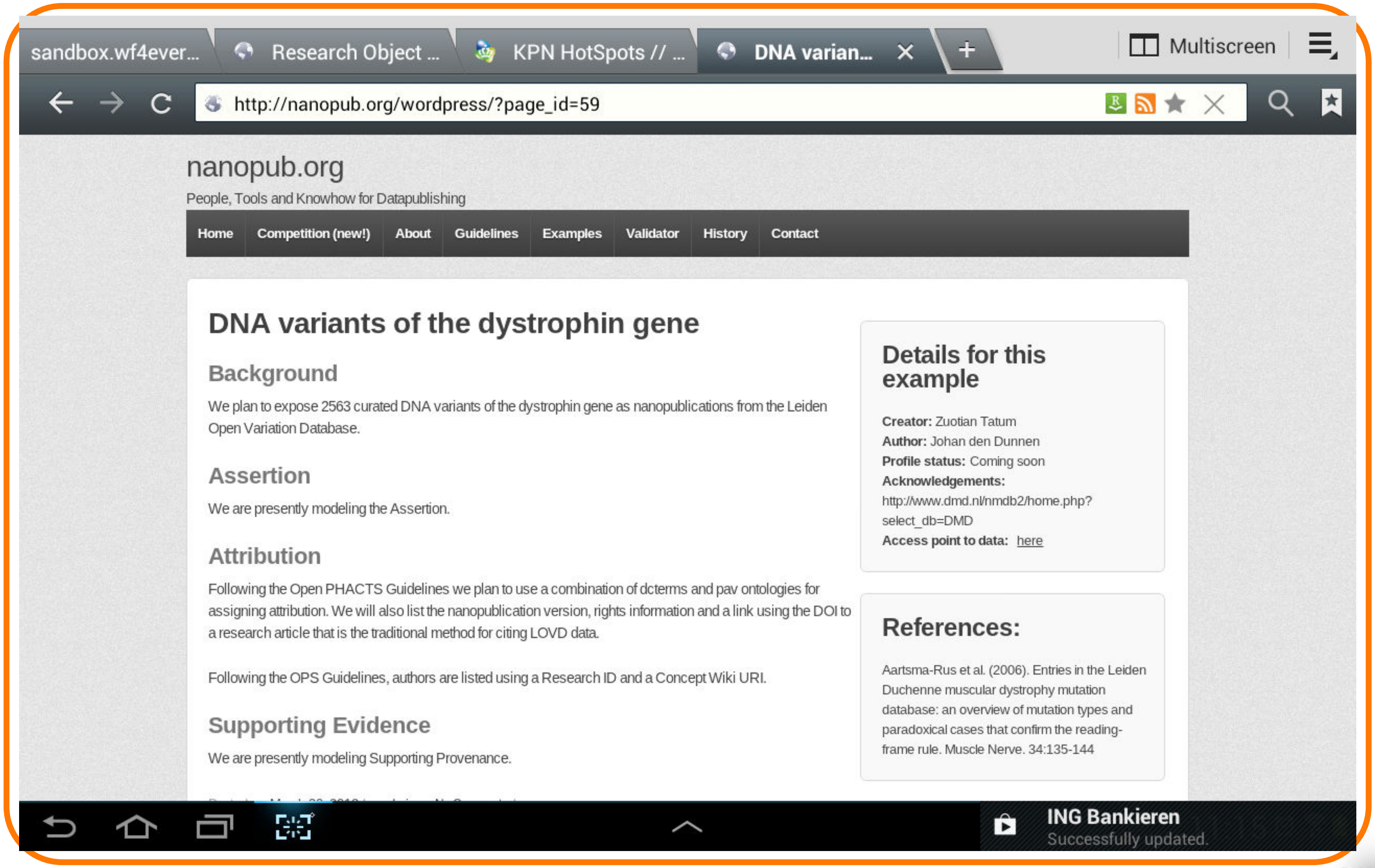
People, Tools and Knowhow for Datapublishing

Nanopub.org

Learn how to create, find, use and cite nanopublications

Nano-Publication in the e-science era
The Anatomy of a Nano-publication
COMMENTARY
The value of data
nanopub: a beginner's guide to datapublishing
Open PHACTS
EURETOS
Nov
2009
2010
Mar
Apr
2011
Sep
Feb
Apr
2012
LectureNotes (Trial Version)
11:13

Examples



nanopub.org

People, Tools and Knowhow for Datapublishing

DNA variants of the dystrophin gene

Background

We plan to expose 2563 curated DNA variants of the dystrophin gene as nanopublications from the Leiden Open Variation Database.

Assertion

We are presently modeling the Assertion.

Attribution

Following the Open PHACTS Guidelines we plan to use a combination of dcterms and pav ontologies for assigning attribution. We will also list the nanopublication version, rights information and a link using the DOI to a research article that is the traditional method for citing LOVD data.

Following the OPS Guidelines, authors are listed using a Research ID and a Concept Wiki URI.

Supporting Evidence

We are presently modeling Supporting Provenance.

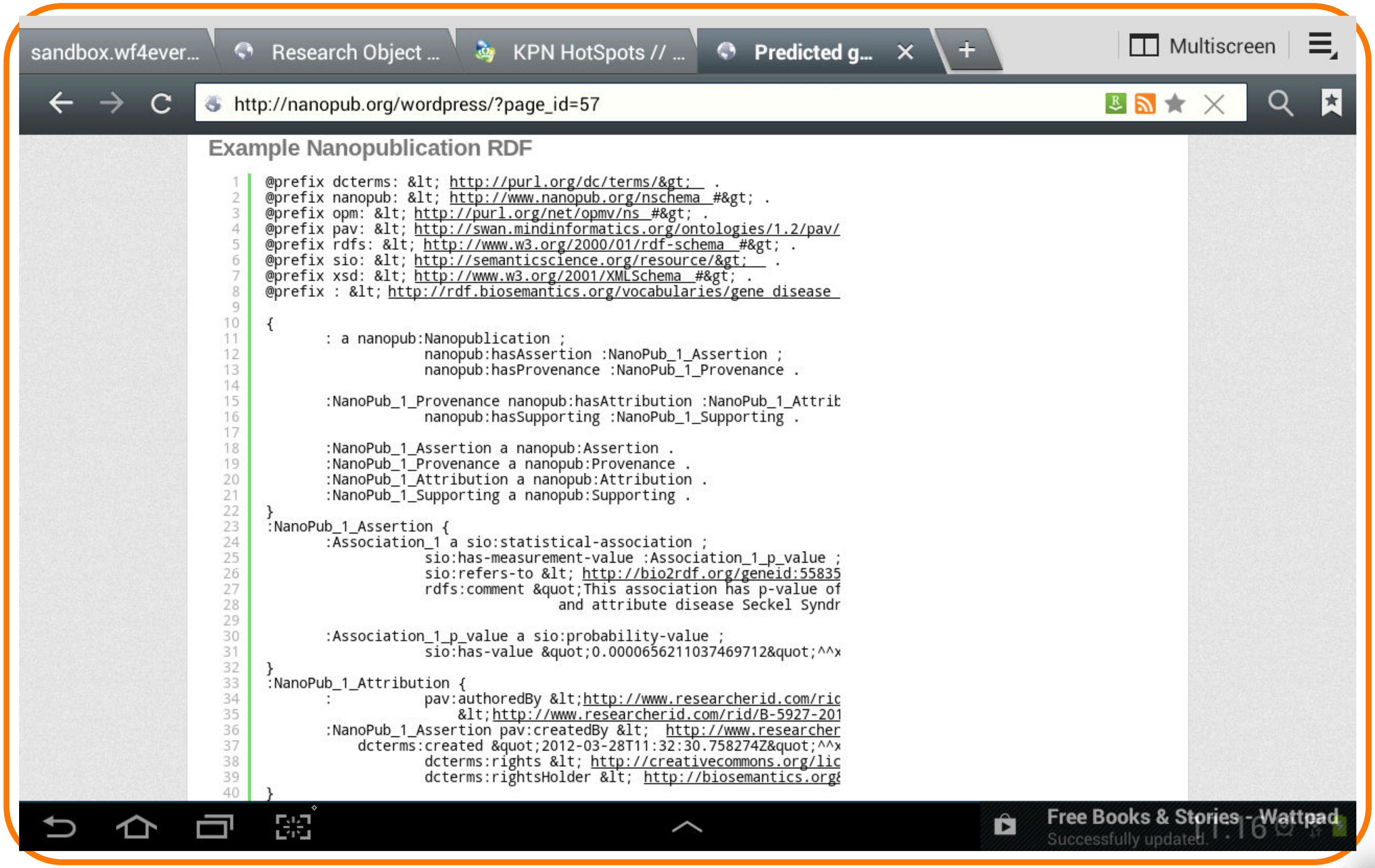
Details for this example

Creator: Zuotian Tatum
Author: Johan den Dunnen
Profile status: Coming soon
Acknowledgements: http://www.dmd.nl/nmdb2/home.php?select_db=DMD
Access point to data: here

References:

Aartsma-Rus et al. (2006). Entries in the Leiden Duchenne muscular dystrophy mutation database: an overview of mutation types and paradoxical cases that confirm the reading-frame rule. Muscle Nerve. 34:135-144

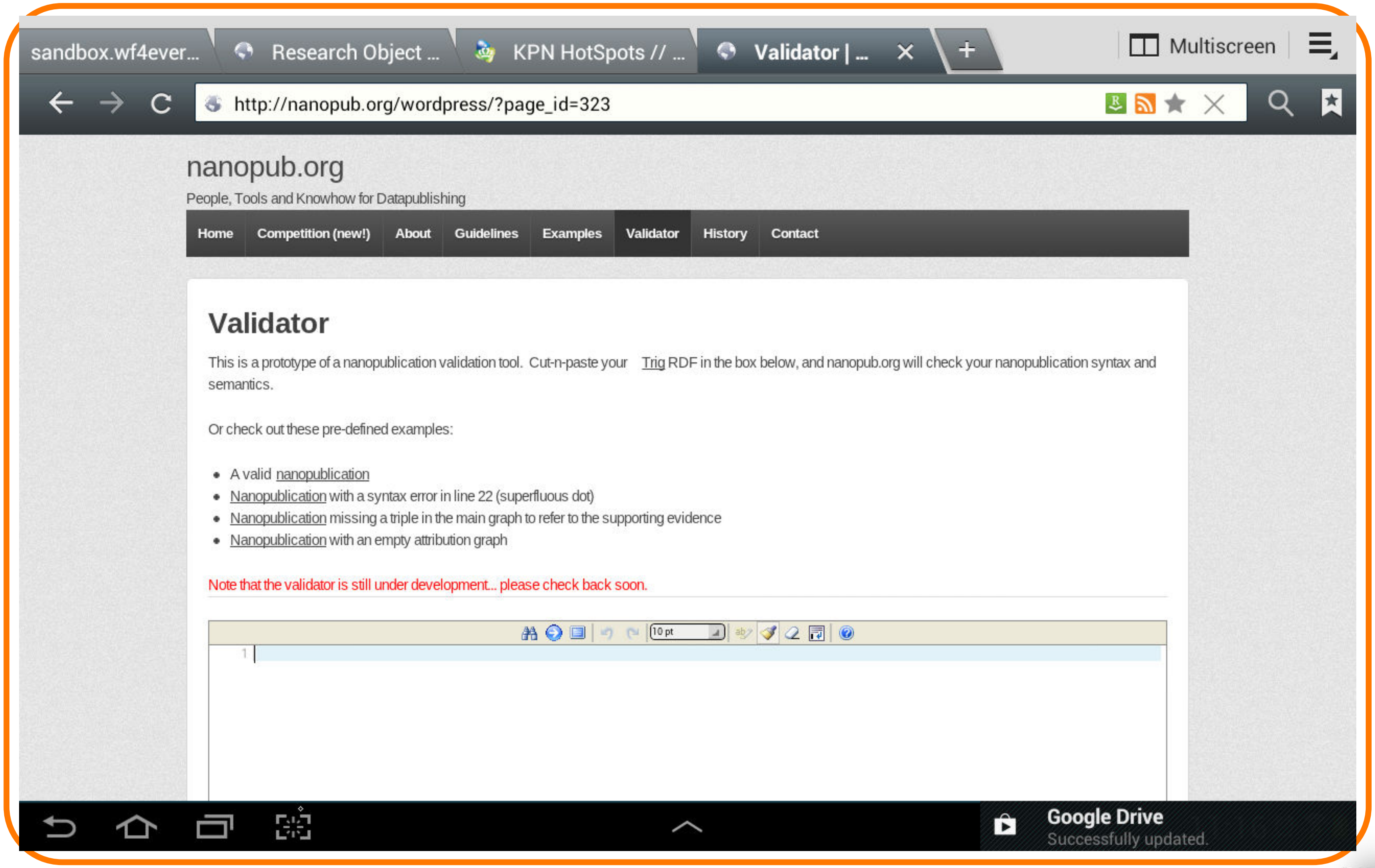
Examples in RDF format



Example Nanopublication RDF

```
@prefix dcterms: &lt; http://purl.org/dc/terms/&gt;   .
@prefix nanopub: &lt; http://www.nanopub.org/nschema  #&gt; .
@prefix opm: &lt; http://purl.org/net/opmv/ns  #&gt; .
@prefix pav: &lt; http://swan.mindinformatics.org/ontologies/1.2/pav/
@prefix rdfs: &lt; http://www.w3.org/2000/01/rdf-schema  #&gt; .
@prefix sio: &lt; http://semanticscience.org/resource/&gt;   .
@prefix xsd: &lt; http://www.w3.org/2001/XMLSchema  #&gt; .
@prefix : &lt; http://rdf.biosemantics.org/vocabularies/gene_disease_

{
       : a nanopub:Nanopublication ;
                   nanopub:hasAssertion :NanoPub_1_Assertion ;
                   nanopub:hasProvenance :NanoPub_1_Provenance .

       :NanoPub_1_Provenance nanopub:hasAttribution :NanoPub_1_Attrib
                   nanopub:hasSupporting :NanoPub_1_Supporting .

       :NanoPub_1_Assertion a nanopub:Assertion .
       :NanoPub_1_Provenance a nanopub:Provenance .
       :NanoPub_1_Attribution a nanopub:Attribution .
       :NanoPub_1_Supporting a nanopub:Supporting .
}
:NanoPub_1_Assertion {
       :Association_1 a sio:statistical-association ;
                   sio:has-measurement-value :Association_1_p_value ;
                   sio:refers-to &lt; http://bio2rdf.org/geneid:55835
                   rdfs:comment &quot;This association has p-value of
                                   and attribute disease Seckel Syndr

       :Association_1_p_value a sio:probability-value ;
                   sio:has-value &quot;0.000065621103746971 2&quot;^^x
}
:NanoPub_1_Attribution {
       :           pav:authoredBy &lt;http://www.researcherid.com/ric
                       &lt;http://www.researcherid.com/rid/B-5927-201
       :NanoPub_1_Assertion pav:createdBy &lt;  http://www.researcher
           dcterms:created &quot;2012-03-28T11:32:30.758274Z&quot;^^x
                   dcterms:rights &lt; http://creativecommons.org/lic
                   dcterms:rightsHolder &lt; http://biosemantics.org&
}
```

Validator



nanopub.org

People, Tools and Knowhow for Datapublishing

Validator

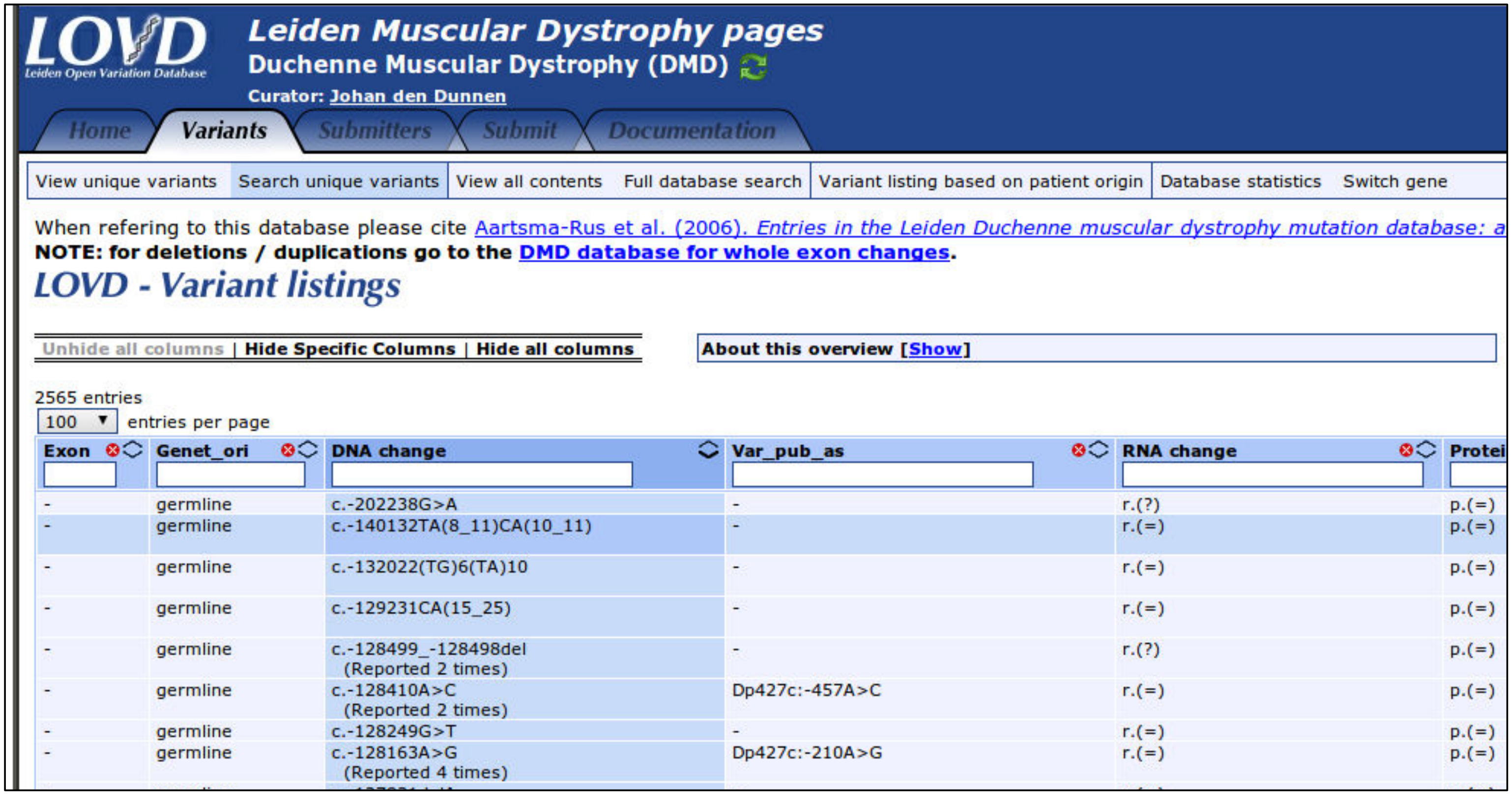
This is a prototype of a nanopublication validation tool. Cut-n-paste your Trig RDF in the box below, and nanopub.org will check your nanopublication syntax and semantics.

Or check out these pre-defined examples:

- A valid nanopublication
- Nanopublication with a syntax error in line 22 (superfluous dot)
- Nanopublication missing a triple in the main graph to refer to the supporting evidence
- Nanopublication with an empty attribution graph

Note that the validator is still under development... please check back soon.

Example: LOVD



Leiden Muscular Dystrophy pages

Duchenne Muscular Dystrophy (DMD)

Curator: Johan den Dunnen

When refering to this database please cite Aartsma-Rus et al. (2006). *Entries in the Leiden Duchenne muscular dystrophy mutation database: a*

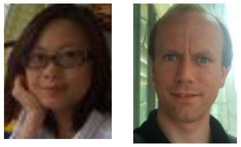
NOTE: for deletions / duplications go to the DMD database for whole exon changes.

LOVD - Variant listings

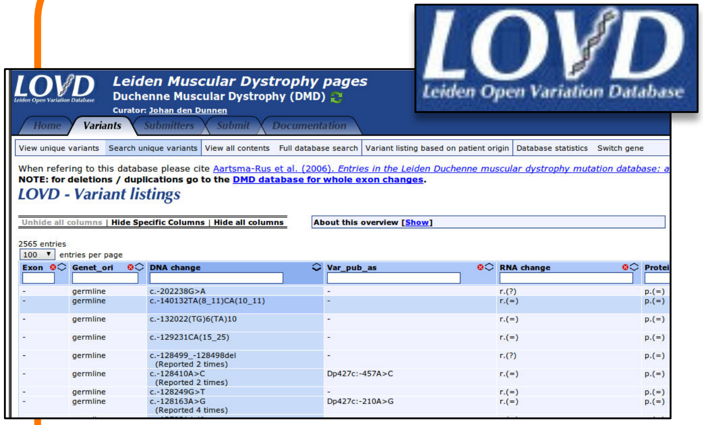
2565 entries

Exon	Genet_ori	DNA change	Var_pub_as	RNA change	Protei
-	germline	c.-202238G>A	-	r.(?)	p.(=)
-	germline	c.-140132TA(8_11)CA(10_11)	-	r.(=)	p.(=)
-	germline	c.-132022(TG)6(TA)10	-	r.(=)	p.(=)
-	germline	c.-129231CA(15_25)	-	r.(=)	p.(=)
-	germline	c.-128499_-128498del (Reported 2 times)	-	r.(?)	p.(=)
-	germline	c.-128410A>C (Reported 2 times)	Dp427c:-457A>C	r.(=)	p.(=)
-	germline	c.-128249G>T	-	r.(=)	p.(=)
-	germline	c.-128163A>G (Reported 4 times)	Dp427c:-210A>G	r.(=)	p.(=)

Nanopublications of Genetic Variations visualized on the genome



Zuotian Tatum, Jesse van Dam



Leiden Muscular Dystrophy pages

Duchenne Muscular Dystrophy (DMD)

Curator: Johan den Dunnen

View unique variants | Search unique variants | View all contents | Full database search | Variant listing based on patient origin | Database statistics | Switch gene

When refering to this database please cite Aartsma-Rus et al. (2006). *Entries in the Leiden Duchenne muscular dystrophy mutation database: a*

NOTE: for deletions / duplications go to the DMD database for whole exon changes.

LOVD - Variant listings

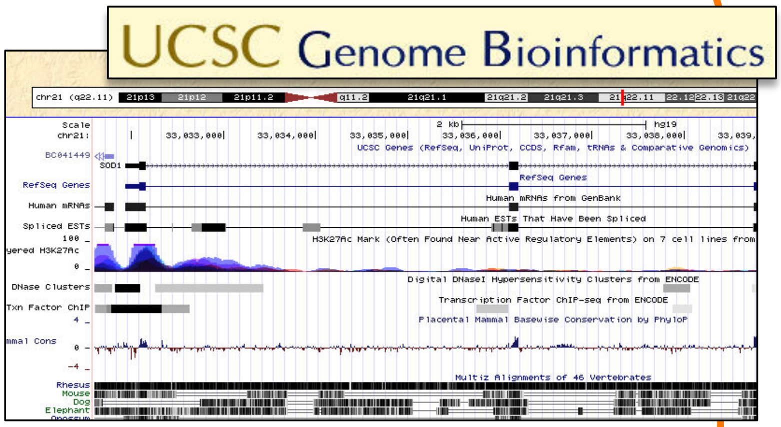
Unhide all columns | **Hide Specific Columns** | **Hide all columns**

About this overview [Show]

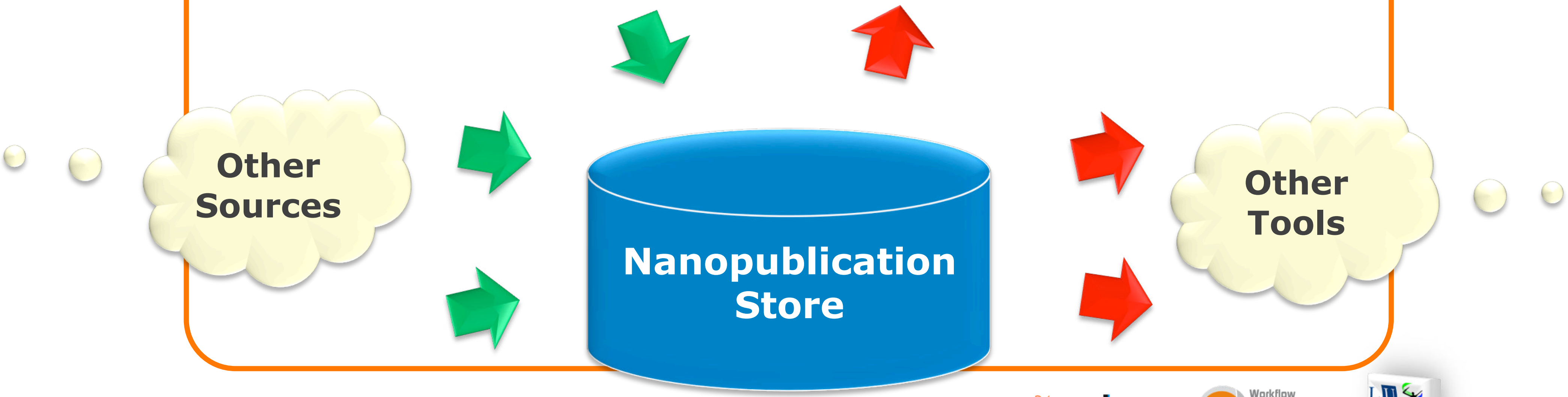
2565 entries

100 ▼ entries per page

Exon	Genet_ori	DNA change	Var_pub_as	RNA change	Protei
-	germline	c.-202238G>A	-	r.(?)	p.(=)
-	germline	c.-140132TA(8_11)CA(10_11)	-	r.(=)	p.(=)
-	germline	c.-132022(TG)6(TA)10	-	r.(=)	p.(=)
-	germline	c.-129231CA(15_25)	-	r.(=)	p.(=)
-	germline	c.-128499_-128498del (Reported 2 times)	-	r.(?)	p.(=)
-	germline	c.-128410A>C (Reported 2 times)	Dp427c:-457A>C	r.(=)	p.(=)
-	germline	c.-128249G>T	-	r.(=)	p.(=)
-	germline	c.-128163A>G (Reported 4 times)	Dp427c:-210A>G	r.(=)	p.(=)



UCSC Genome Bioinformatics
chr21 (q22.11)
21p13
21p12
21p11.2
q11.2
21q21.1
21q21.2
21q21.3
21q22.11
22.12
22.13
21q22
Scale
chr21:
2 kb
hg19
33,033,000
33,034,000
33,035,000
33,036,000
33,037,000
33,038,000
33,039,
UCSC Genes (RefSeq, UniProt, CCDS, Rfam, tRNAs & Comparative Genomics)
BC041449
SOD1
RefSeq Genes
RefSeq Genes
Human mRNAs
Human mRNAs from GenBank
Spliced ESTs
Human ESTs That Have Been Spliced
100
yered H3K27Ac
0
H3K27Ac Mark (Often Found Near Active Regulatory Elements) on 7 cell lines from
DNase Clusters
Digital DNaseI Hypersensitivity Clusters from ENCODE
Txn Factor ChIP
Transcription Factor ChIP-seq from ENCODE
4
Placental Mammal Basewise Conservation by PhyloP
mmal Cons
0
-4
Multiz Alignments of 46 Vertebrates
Rhesus
Mouse
Dog
Elephant
Opossum



Fame and Glory

It was me, me, me!

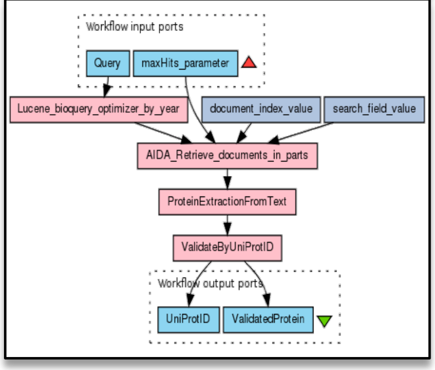
What I found

Nanopublication

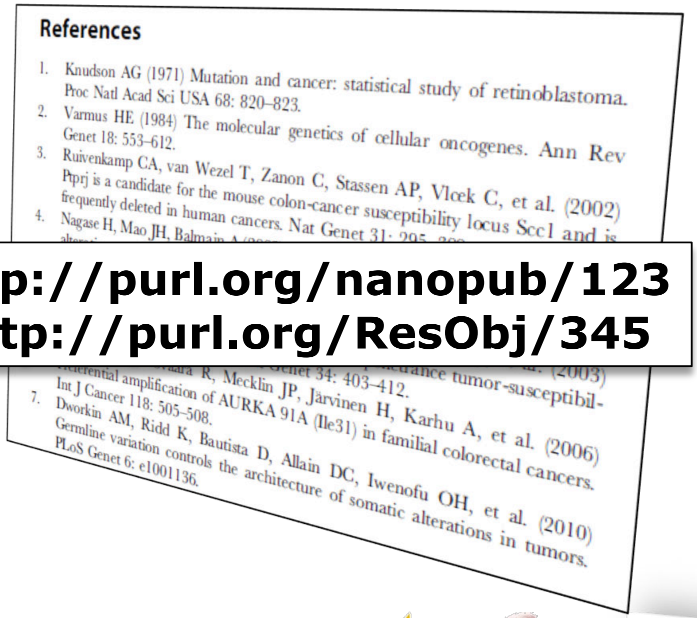
<CS7183> *<associatedWith>* <MetS>

Discovered by: me
Published by: me

Research Object



How I found it



References

1. Knudson AG (1971) Mutation and cancer: statistical study of retinoblastoma. Proc Natl Acad Sci USA 68: 820–823.
2. Varmus HE (1984) The molecular genetics of cellular oncogenes. Ann Rev Genet 18: 553–612.
3. Ruivenkamp CA, van Wezel T, Zanon C, Stassen AP, Vlcek C, et al. (2002) Ptprj is a candidate for the mouse colon-cancer susceptibility locus Scc1 and is frequently deleted in human cancers. Nat Genet 31: 295–300.
4. Nagase H, Mao JH, Balmain A (200

ance tumor-susceptibil- (2003) Genet 34: 403–412.

aara R, Mecklin JP, Järvinen H, Karhu A, et al. (2006) ferential amplification of AURKA 91A (Ile31) in familial colorectal cancers. Int J Cancer 118: 505–508.

7. Dworkin AM, Ridd K, Bautista D, Allain DC, Iwenofu OH, et al. (2010) Germline variation controls the architecture of somatic alterations in tumors. PLoS Genet 6: e1001136.

http://purl.org/nanopub/123
http://purl.org/ResObj/345

Summary (1/2)

- Preservation under the hood of digital research tools
- Research Object Model: annotated aggregates
- Nanopublication: fine-grained digital credit
  Check Nanopub.org to stay updated

Summary (2/2)

- Semantic Web for exchange and interoperability
- In progress: RO-enabling myExperiment
  Watch myExperiment.org in 2013!
- Plans to RO-enable Taverna, Galaxy, GenomeSpace

Workflow 4Ever

Acknowledgements

EU Wf4Ever project (270129)
funded under EU FP7 (ICT- 2009.4.1).
(http://www.wf4ever-project.org)

(GIGA)^{n} SCIENCE

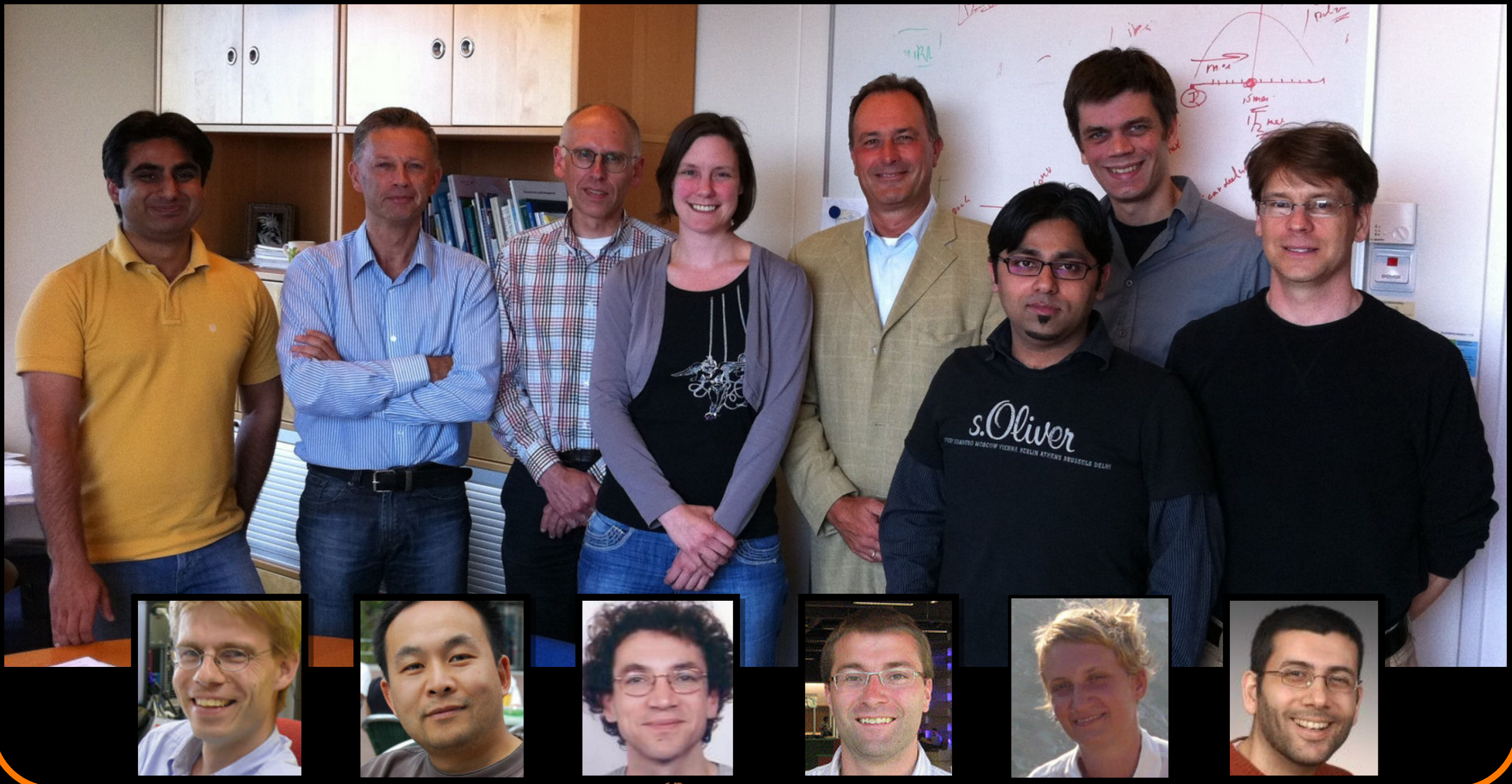
nbic

Open PHACTS

Concept Web Alliance

LUMC

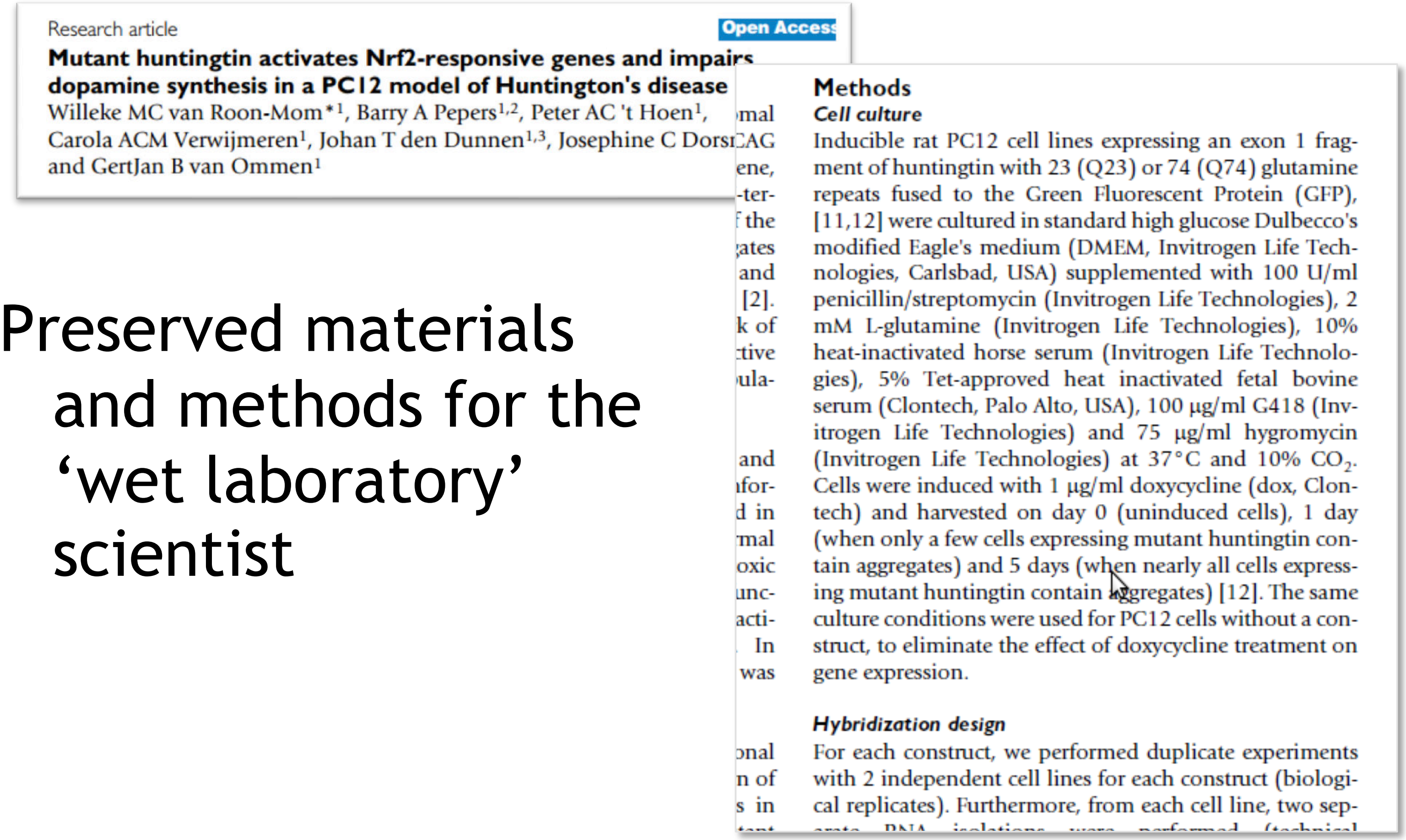
Thank you for your attention



http://biosemantics.org

Reproducible Science

Preserved materials and methods for the ‘wet laboratory’ scientist



Research article

Open Access

Mutant huntingtin activates Nrf2-responsive genes and impairs dopamine synthesis in a PC12 model of Huntington's disease

Willeke MC van Roon-Mom*[1], Barry A Pepers[1,2], Peter AC 't Hoen[1], Carola ACM Verwijmeren[1], Johan T den Dunnen[1,3], Josephine C Dors
and GertJan B van Ommen[1]

mal
CAG
ene,
-ter-
f the
gates
and
[2].
k of
ctive
ula-

and
nfor-
d in
mal
oxic
unc-
acti-
. In
was

onal
n of
s in

Methods

Cell culture

Inducible rat PC12 cell lines expressing an exon 1 fragment of huntingtin with 23 (Q23) or 74 (Q74) glutamine repeats fused to the Green Fluorescent Protein (GFP), [11,12] were cultured in standard high glucose Dulbecco's modified Eagle's medium (DMEM, Invitrogen Life Technologies, Carlsbad, USA) supplemented with 100 U/ml penicillin/streptomycin (Invitrogen Life Technologies), 2 mM L-glutamine (Invitrogen Life Technologies), 10% heat-inactivated horse serum (Invitrogen Life Technologies), 5% Tet-approved heat inactivated fetal bovine serum (Clontech, Palo Alto, USA), 100 µg/ml G418 (Invitrogen Life Technologies) and 75 µg/ml hygromycin (Invitrogen Life Technologies) at 37°C and 10% CO_2. Cells were induced with 1 µg/ml doxycycline (dox, Clontech) and harvested on day 0 (uninduced cells), 1 day (when only a few cells expressing mutant huntingtin contain aggregates) and 5 days (when nearly all cells expressing mutant huntingtin contain aggregates) [12]. The same culture conditions were used for PC12 cells without a construct, to eliminate the effect of doxycycline treatment on gene expression.

Hybridization design

For each construct, we performed duplicate experiments with 2 independent cell lines for each construct (biological replicates). Furthermore, from each cell line, two sep-

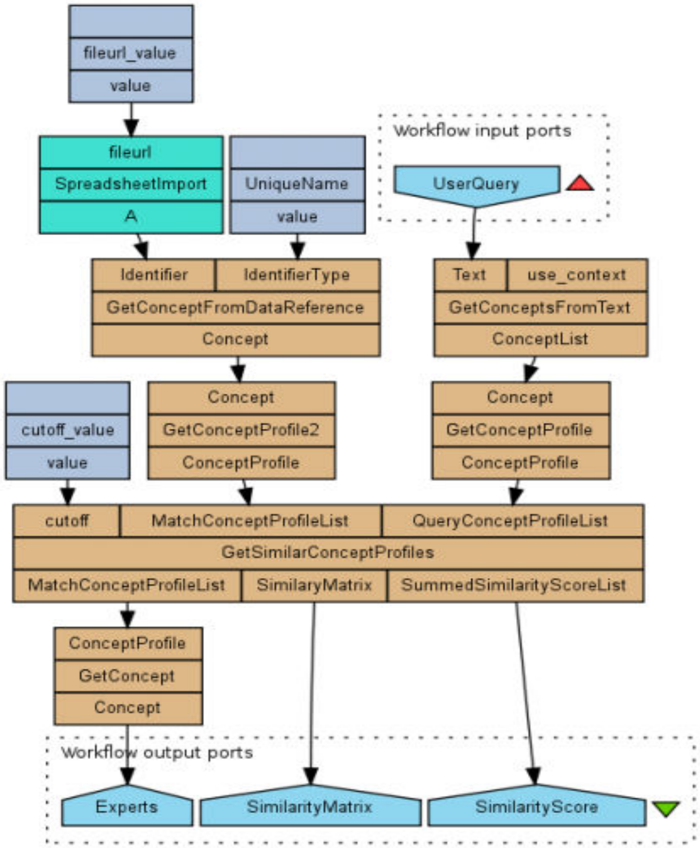
From Van Roon-Mom *et al.*, BMC Molecular Biology 2008
doi: 10.1186/1471-2199-9-84.

Reproducible Science?

What is the digital equivalent?

Is it equally good?

Can we do better?
 - *or worse?*



Reproduced from Jelier *et al.*, Schuemie *et al.*, Hettne *et al.*, Haagen *et al.*, http://biosemantics.org , *myExperiment.org/workflows/2197*

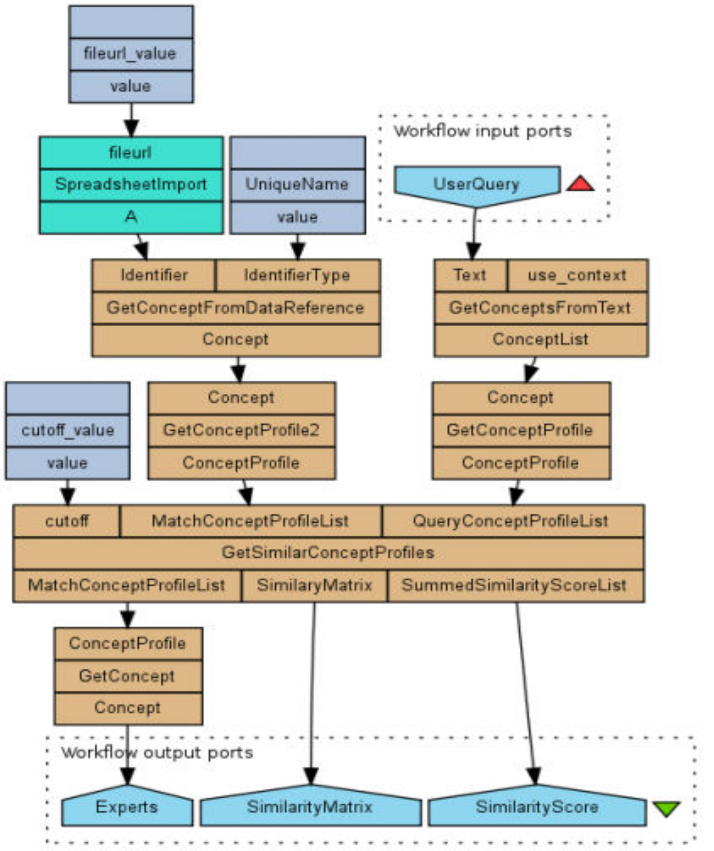
Reproducible Science

What is the digital equivalent?

Is it equally good?

Can we do better?
- *or worse?*

Can you tell what this is doing?



Reproduced from Jelier *et al.*, Schuemie *et al.*, Hettne *et al.*, Haagen *et al.*, http://biosemantics.org , *myExperiment.org/workflows/2197*

Reproducible Science

What is our incentive?

Nobility

Good Reproducible Science

Greater Good

Serve the public

Reproducible Science

What is our incentive?

Fame and Glory

Getting on with it...

Our aim

‘Useful’ preservation

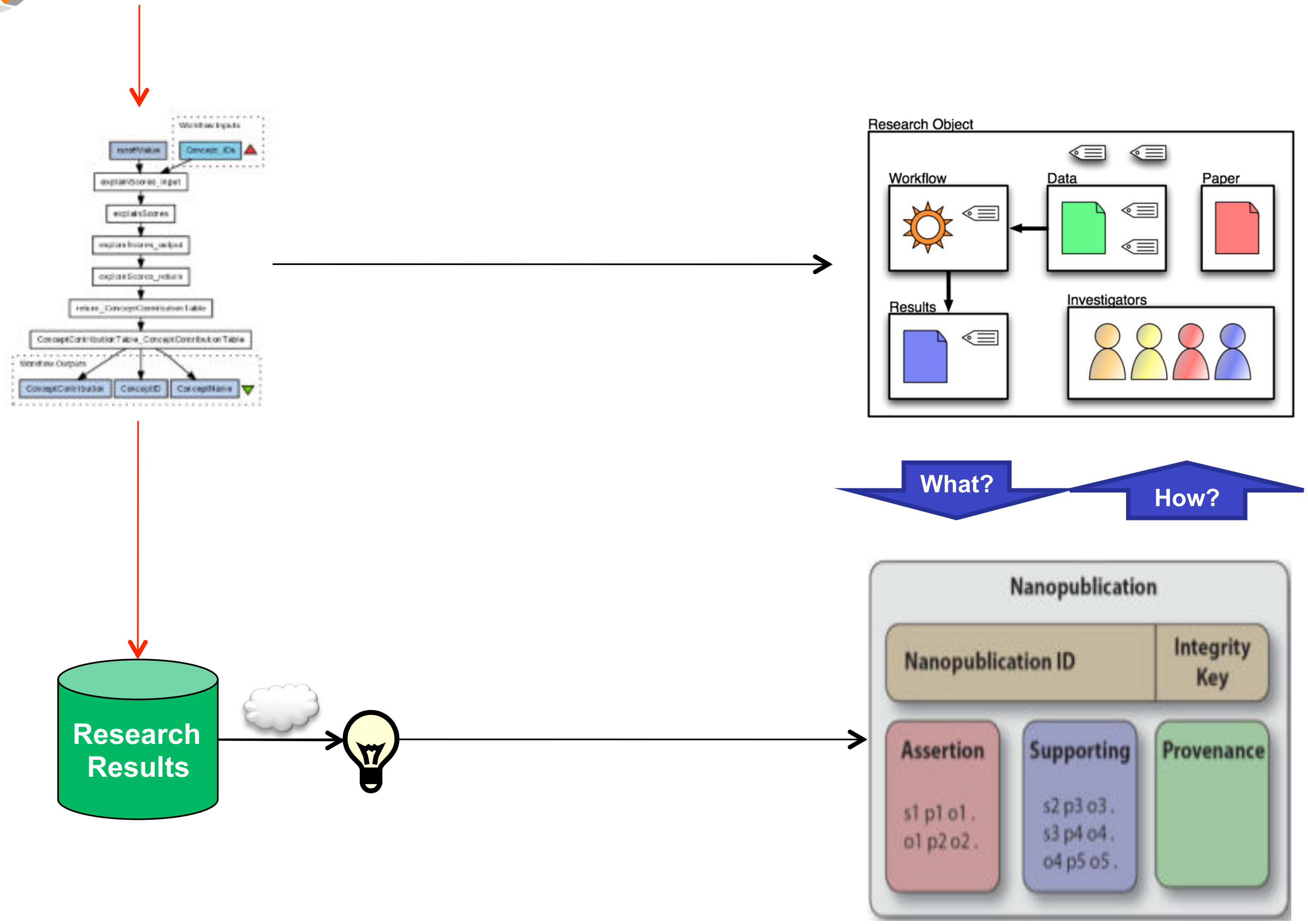
Support reproducibility
in tools and by guidelines that

speed up your research

get you acknowledgement

Workflow 4Ever

Preservation



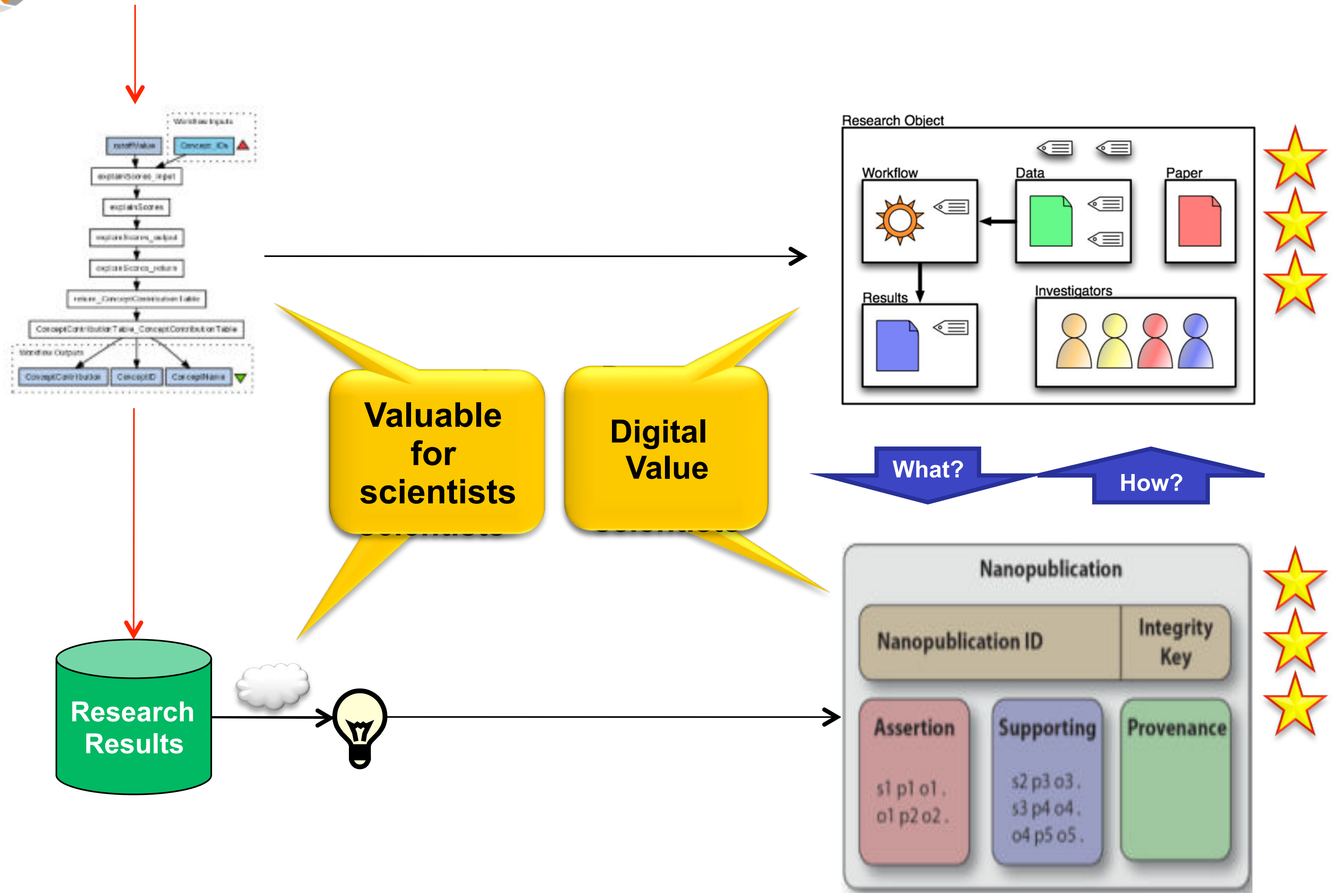
Research Results

What?

How?

nbic Open PHACTS Concept Web Alliance LUMC

Preservation



Acknowledgements

http://biosemantics.org/

- Erik Schultes
- Andrew Gibson
- Reinout van Schouwen
- Kostas Karasavvas
- Kristina Hettne
- Harish Dharuri
- Eleni Mina
- Jesse van Dam
- Herman van Haagen
- Zuotian Tatum
- Johan den Dunnen
- Peter-Bram ‘t Hoen
- Barend Mons
- Gert-Jan van Ommen

- Paul Groth
- Frank van Harmelen

- Erik van Mulligen
- Bharat Singh
- Jan Kors

- Christine Chichester
- Kees Burger - NBIC
- Spyros Kotoulas - VU
- Antonis Loizou - VU
- Valery Tkachenko - RSC
- Andra Waagmeester - Maastricht
- Sune Askjaer - Lundbeck
- Steve Pettifer - Manchester
- Lee Harland - Pfizer/CD
- Carina Haupt - Fraunhofer
- Colin Batchelor - RSC
- Miguel Vazquez - CNIO
- José María Fernández - CNIO
- Jahn Saito - Maastricht
- Andrew Gibson (Outside Expert) - Amsterdam
- Louis Wich - DTU