



COLLABORATIVE OPEN PLANT OMICS

FELIX SHAW & TONI ETUK



- ▶ 3 Year Project
- ▶ PIs are Rob Davey and Vicky Schneider
- ▶ Collaborators in Uni. Oxford, Uni. Warwick and EBI
- ▶ Work carried out by Felix Shaw, Anthony Etuk, Alejandra Gonzales-Beltran, Phillipe Rocca-Serra and David Johnson



- ▶ COPO is for plant scientists to:
 - ▶ Easily share research objects
 - ▶ Get credit for doing so



- ▶ Funding bodies require data sharing
- ▶ Improved availability of data
- ▶ Citation for all your outputs



ISA REST RDF XML
SRA SOAP ENA
WTF!
JSON XSD
OWL

**...AVERAGE TIME FOR A
NOVICE TO GET DATA INTO
ENA? ...3 WEEKS!**

Kenneth Haug - EBI





- ▶ Data retrieval almost as difficult
 - ▶ Must have accession
 - ▶ People resort to copying accessions from publications



```
<SUBMISSION_SET>
<SUBMISSION alias="CCC_III_18" center_name="TGAC"
|·| submission_comment="CCC_III_18" submission_date="2014-12-16T00:00:00">
<TITLE>CCC_III_18_JonathanHeeney_chimpSIV_LEGACY</TITLE>
<CONTACTS>
<CONTACT name="Robert Davey"/>
</CONTACTS>
<ACTIONS>
<ACTION>
<ADD schema="experiment" source="SUB6_experiment_2014-12-16.rml.xml"/>
</ACTION>
<ACTION>
<ADD schema="study" source="SUB6_study_2014-12-16.xml"/>
</ACTION>
<ACTION>
<ADD schema="run" source="SUB6_run_2014-12-16.rml.xml"/>
</ACTION>
<ACTION>
<ADD schema="sample" source="SUB6_sample_2014-12-16.xml"/>
</ACTION>
</ACTIONS>
</SUBMISSION>
</SUBMISSION_SET>
```

METADATA EXAMPLES



```
<EXPERIMENT_SET>↵↵↵
<EXPERIMENT alias="EXP_AUTOGEN_STU180_Transcriptome Analysis_1-1" center_name="TGAC">↵↵↵
<TITLE>PR026 Illumina Transcriptome Analysis experiment (Auto-gen)</TITLE>↵↵↵
<STUDY_REF refcenter="TGAC" refname="CCC_III_18"/>↵↵↵
<DESIGN>↵↵↵
<DESIGN_DESCRIPTION>CCC_III_18</DESIGN_DESCRIPTION>↵↵↵
<SAMPLE_DESCRIPTOR refname="JH_S1_Chimpanzee" refcenter="TGAC">↵↵↵
</SAMPLE_DESCRIPTOR>↵↵↵
<LIBRARY_DESCRIPTOR>↵↵↵
<LIBRARY_NAME/>↵↵↵
<LIBRARY_STRATEGY>OTHER</LIBRARY_STRATEGY>↵↵↵
<LIBRARY_SOURCE>TRANSCRIPTOMIC</LIBRARY_SOURCE>↵↵↵
<LIBRARY_SELECTION>PCR</LIBRARY_SELECTION>↵↵↵
<LIBRARY_LAYOUT>↵↵↵
<PAIRED NOMINAL_LENGTH="640"/>↵↵↵
</LIBRARY_LAYOUT>↵↵↵
<POOLING_STRATEGY>multiplexed libraries</POOLING_STRATEGY>↵↵↵
</LIBRARY_DESCRIPTOR>↵↵↵
</DESIGN>↵↵↵
<PLATFORM>↵↵↵
<ILLUMINA>↵↵↵
<INSTRUMENT_MODEL>Illumina HiSeq 2000</INSTRUMENT_MODEL>↵↵↵
</ILLUMINA>↵↵↵
</PLATFORM>↵↵↵
<PROCESSING/>↵↵↵
</EXPERIMENT>↵↵↵
<EXPERIMENT alias="EXP_AUTOGEN_STU180_Transcriptome Analysis_1-2" center_name="TGAC">↵↵↵
<TITLE>PR026 Illumina Transcriptome Analysis experiment (Auto-gen)</TITLE>↵↵↵
<STUDY_REF refcenter="TGAC" refname="CCC_III_18_JonathanHeeney_chimpSIV_LEGACY"/>↵↵↵
```

METADATA EXAMPLES



```
<?xml version="1.0" encoding="UTF-8"?>
<SAMPLE_SET>
<SAMPLE alias="JH_S1_Chimpanzee" center_name="TGAC">
<TITLE>JH_S1_Chimpanzee</TITLE>
<SAMPLE_NAME>
<SCIENTIFIC_NAME>Pan troglodytes</SCIENTIFIC_NAME>
<TAXON_ID>9598</TAXON_ID>
</SAMPLE_NAME>
<DESCRIPTION>CA_x062 wk4</DESCRIPTION>
</SAMPLE>
<SAMPLE alias="JH_S6_Chimpanzee" center_name="TGAC">
<TITLE>JH_S6_Chimpanzee</TITLE>
<SAMPLE_NAME>
<SCIENTIFIC_NAME>Pan troglodytes</SCIENTIFIC_NAME>
<TAXON_ID>9598</TAXON_ID>
</SAMPLE_NAME>
<DESCRIPTION>CF_x176 wk52</DESCRIPTION>
</SAMPLE>
<SAMPLE alias="JH_S7_Chimpanzee" center_name="TGAC">
<TITLE>JH_S7_Chimpanzee</TITLE>
<SAMPLE_NAME>
<SCIENTIFIC_NAME>Pan troglodytes</SCIENTIFIC_NAME>
<TAXON_ID>9598</TAXON_ID>
</SAMPLE_NAME>
<DESCRIPTION>CH_x284 wk4</DESCRIPTION>
</SAMPLE>
<SAMPLE alias="JH_S8_Chimpanzee" center_name="TGAC">
```

METADATA EXAMPLES



```
<RUN_SET>
<RUN alias="L004:LDI2167:120327_SN790_0195_AD0TDRACXX" center_name="TGAC" run_center="TGAC" run_date="2012-04-08T00:00:00">
<EXPERIMENT_REF refcenter="TGAC" refname="EXP_AUTOGEN_STU180_Transcriptome Analysis_2-20"/>
<DATA_BLOCK sector="4">
<FILES>
<FILE ascii_offset="!" checksum_method="MD5" filename="Project_CCC-14/Run372_LIB2555_LDI2167_C8-S_CTTGTA_L004_R1.fastq.gz"
filetype="fastq" quality_encoding="ascii" quality_scoring_system="phred">
</FILE>
<FILE ascii_offset="!" checksum_method="MD5" filename="Project_CCC-14/Run372_LIB2555_LDI2167_C8-S_CTTGTA_L004_R2.fastq.gz"
filetype="fastq" quality_encoding="ascii" quality_scoring_system="phred">
</FILE>
</FILES>
</DATA_BLOCK>
</RUN>
<RUN alias="L004:LDI2165:120327_SN790_0195_AD0TDRACXX" center_name="TGAC" run_center="TGAC" run_date="2012-04-08T00:00:00">
<EXPERIMENT_REF refcenter="TGAC" refname="EXP_AUTOGEN_STU180_Transcriptome Analysis_2-19"/>
<DATA_BLOCK sector="4">
<FILES>
<FILE ascii_offset="!" checksum_method="MD5" filename="Project_CCC-14/Run372_LIB2553_LDI2165_C6_CAGATC_L004_R1.fastq.gz"
filetype="fastq" quality_encoding="ascii" quality_scoring_system="phred">
</FILE>
<FILE ascii_offset="!" checksum_method="MD5" filename="Project_CCC-14/Run372_LIB2553_LDI2165_C6_CAGATC_L004_R2.fastq.gz"
filetype="fastq" quality_encoding="ascii" quality_scoring_system="phred">
</FILE>
</FILES>
</DATA_BLOCK>
</RUN>
<RUN alias="L004:LDI2164:120327_SN790_0195_AD0TDRACXX" center_name="TGAC" run_center="TGAC" run_date="2012-04-08T00:00:00">
<EXPERIMENT_REF refcenter="TGAC" refname="EXP_AUTOGEN_STU180_Transcriptome Analysis_2-18"/>
<DATA_BLOCK sector="4">
<FILES>
<FILE ascii_offset="!" checksum_method="MD5" filename="Project_CCC-14/Run372_LIB2552_LDI2164_C5_GCCAAT_L004_R1.fastq.gz"
```



```
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema" xmlns:com="SRA.common">
  <xs:import schemalocation="SRA.common.xsd" namespace="SRA.common"/>

  <xs:complexType name="SubmissionType">
    <xs:annotation>
      <xs:documentation>
        A Submission type is used to describe an object that contains submission actions to be performed by the archiv
      </xs:documentation>
    </xs:annotation>

    <xs:sequence>
      <xs:element name="IDENTIFIERS" type="com:IdentifierType" maxOccurs="1" minOccurs="0"> </xs:element>
      <xs:element name="TITLE" type="xs:string" minOccurs="0" maxOccurs="1">
        <xs:annotation>
          <xs:documentation>
            Short text that can be used to define submissions in searches or in displays.
          </xs:documentation>
        </xs:annotation>
      </xs:element>
      <xs:element name="CONTACTS" minOccurs="0" maxOccurs="1">
        <xs:complexType>
          <xs:sequence minOccurs="1" maxOccurs="unbounded">
            <xs:element name="CONTACT">
              <xs:complexType>
                <xs:attribute name="name" type="xs:string" use="optional">
                  <xs:annotation>
                    <xs:documentation>
                      Name of contact person for this submission.
                    </xs:documentation>
                  </xs:annotation>
                </xs:attribute>
                <xs:attribute name="inform_on_status" type="xs:anyURI" use="optional">
                  <xs:annotation>
                    <xs:documentation>
                      Internet address of person or service to inform on any status changes for this submission.
                    </xs:documentation>
                  </xs:annotation>
                </xs:attribute>
                <xs:attribute name="inform_on_error" type="xs:anyURI" use="optional">
                  <xs:annotation>
                    <xs:documentation>
                      Internet address of person or service to inform on any errors for this submission.
                    </xs:documentation>
                  </xs:annotation>
                </xs:attribute>
              </xs:complexType>
            </xs:element>
          </xs:sequence>
        </xs:complexType>
      </xs:element>

      <xs:element name="ACTIONS" minOccurs="0" maxOccurs="1">
        <xs:complexType>
          <xs:sequence minOccurs="1" maxOccurs="unbounded">
            <xs:element name="ACTION">

```

ALTERNATIVES

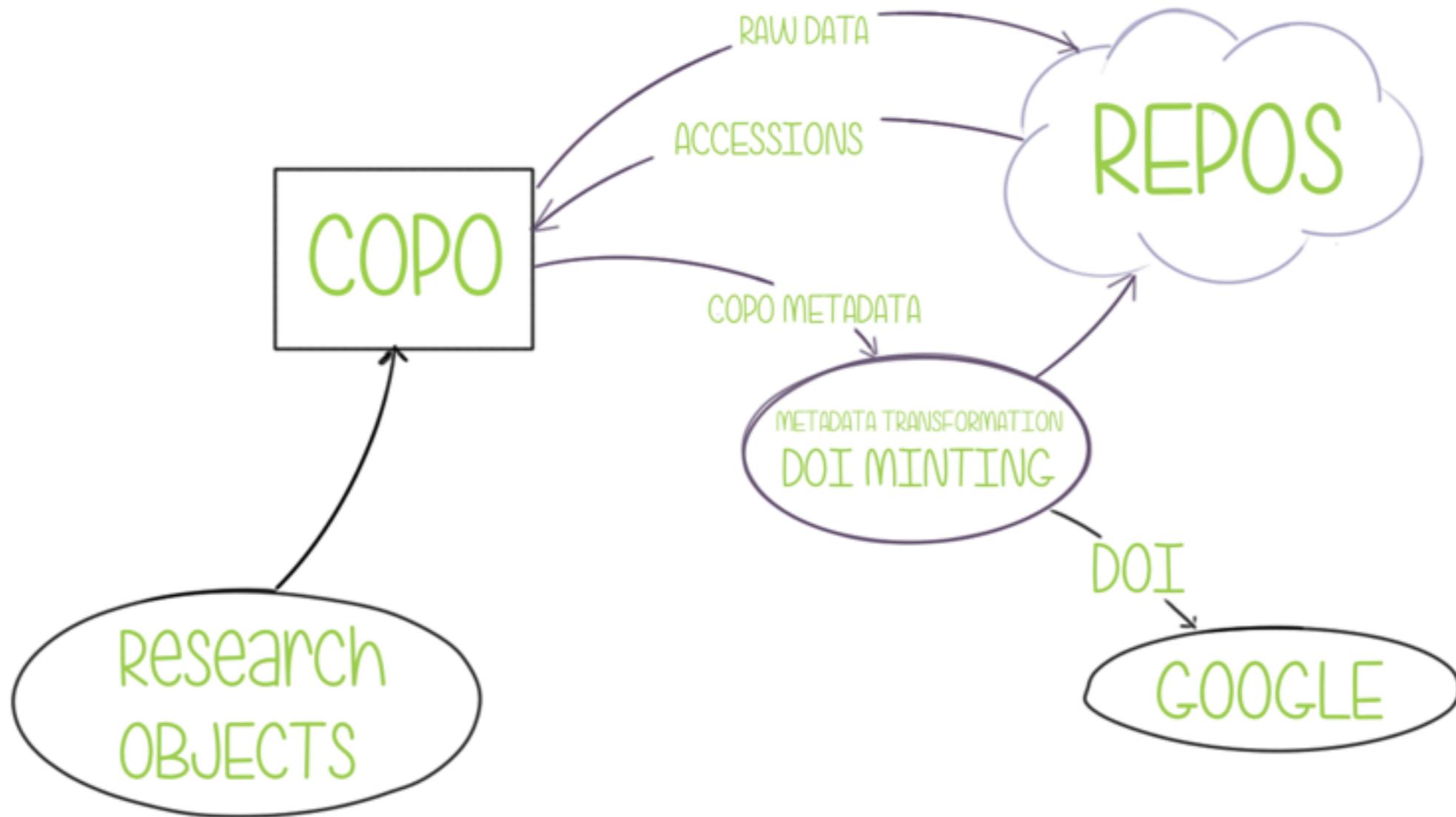


```
STUDY FACTORS~
Study Factor Name~"dose"~ "compound"~ "collection time"~
Study Factor Type~"dose"~ "chemical compound"~"time"~
Study Factor Type Term Accession Number~""~ ""~ ""~
Study Factor Type Term Source REF~""~ ""~ ""~
STUDY ASSAYS~
Study Assay File Name~"a_gilbert-assay-Gx.txt"~ "a_gilbert-assay-Tx.txt"~
Study Assay Measurement Type~ "metagenome sequencing"~"transcription profiling"~
Study Assay Measurement Type Term Accession Number~ ""~ "0000424"~
Study Assay Measurement Type Term Source REF~ "OBI"~"OBI"~
Study Assay Technology Type~"nucleotide sequencing"~"nucleotide sequencing"~
Study Assay Technology Type Term Accession Number~""~ ""~
Study Assay Technology Type Term Source REF~"OBI"~"OBI"~
Study Assay Technology Platform~"454 Genome Sequencer FLX"~ "454 Genome Sequencer FS"~
STUDY PROTOCOLS~
Study Protocol Name~"sample collection - standard procedure 1"~ "nucleic acid extraction - standard procedure 2"~ "mRNA extraction - standard procedure 3"~ "genomic DNA extraction - standard procedure 4"~"reverse
transcription - standard procedure 5"~ "pyrosequencing - standard procedure 6"~"sequence analysis - standard procedure 7"~
Study Protocol Type~"environmental material collection"~"nucleic acid extraction"~"RNA extraction"~ "DNA extraction"~ "reverse transcription"~"DNA sequencing"~ "DNA sequencing"~
Study Protocol Type Term Accession Number~""~ ""~ ""~ ""~ ""~ ""~
Study Protocol Type Term Source REF~""~ ""~ ""~ ""~ ""~ ""~
Study Protocol Description~ "Waters samples were prefiltered through a 1.6 um GF/A glass fibre filter to reduce Eukaryotic contamination. Filtrate was then collected on a 0.2 um Sterivex (millipore) filter which was
frozen in liquid nitrogen until nucleic acid extraction. CO2 bubbled through 11000 l mesocosm to simulate ocean acidification predicted conditions. Then phosphate and nitrate were added to induce a phytoplankton
bloom."~"Total nucleic acid extraction was done as quickly as possible using the method of Neufeld et al, 2007."~ "RNA MinElute + substrative Hybridization + MEGAclean For transcriptomics, total RNA was separated
from the columns using the RNA MinElute clean-up kit (Qiagen) and checked for integrity of rRNA using an Agilent bioanalyser (RNA nano6000 chip). High integrity rRNA is essential for subtractive hybridization.
Samples were treated with Turbo DNA-free enzyme (Ambion) to remove contaminating DNA. The rRNA was removed from mRNA by subtractive hybridization (Microbe Express Kit, Ambion), and absence of rRNA and DNA
contamination was confirmed using the Agilent bioanalyser. The mRNA was further purified with the MEGAcleanTM kit (Ambion). Reverse transcription of mRNA was performed using the SuperScript III enzyme (Invitrogen)
with random hexamer primers (Promega). The cDNA was treated with RiboShredderTM RNase Blend (Epicentre) to remove trace RNA contaminants. To improve the yield of cDNA, samples were subjected to random amplification
using the GenomiPhi V2 method (GE Healthcare). GenomiPhi technology produces branched DNA molecules that are recalcitrant to the pyrosequencing methodology. Therefore amplified samples were treated with S1 nuclease
using the method of Zhang et al.2006."~ ""~ "superscript+random hexamer primer"~"1. Sample Input and Fragmentation: The Genome Sequencer FLX System supports the sequencing of samples from a wide variety of starting
materials including genomic DNA, PCR products, BACs, and cDNA. Samples such as genomic DNA and BACs are fractionated into small, 300- to 800-base pair fragments. For smaller samples, such as small non-coding RNA or
PCR amplicons, fragmentation is not required. Instead, short PCR products amplified using Genome Sequencer fusion primers can be used for immobilization onto DNA capture beads as shown below."~ ""~
Study Protocol URI~ ""~ ""~ ""~ ""~ ""~ ""~
Study Protocol Version~ ""~ ""~ ""~ ""~ ""~ ""~
Study Protocol Parameters Name~ "filter pore size"~ ""~ ""~ ""~ ""~ ""~ ""~
Study Protocol Parameters Name Term Accession Number~ ""~ ""~ ""~ ""~ ""~ ""~
Study Protocol Parameters Name Term Source REF~ ""~ ""~ ""~ ""~ ""~ ""~
Study Protocol Components Name~ ""~ ""~ ""~ ""~ ""~ "454 GS-FLX"~ ""~
Study Protocol Components Type~ ""~ ""~ ""~ ""~ ""~ ""~ "DNA sequencer"~""~
Study Protocol Components Type Term Accession Number~ ""~ ""~ ""~ ""~ ""~ ""~ ""~
Study Protocol Components Type Term Source REF~ ""~ ""~ ""~ ""~ ""~ ""~ ""~
STUDY CONTACTS~
Study Person Last Name~ "Gilbert"~"Field"~"Huang"~"Edwards"~"Li"~ "Gilna"~"Joint"~
Study Person First Name~ "Jack"~ "Dawn"~ "Ying"~ "Rob"~"Weizhong"~ "Paul"~ "Ian"~
Study Person Mid Initials~"A"~""~ ""~ ""~ ""~ ""~ ""~
Study Person Email~ "jagi@pml.ac.uk"~ ""~ ""~ ""~ ""~ ""~ ""~
Study Person Phone~ ""~ ""~ ""~ ""~ ""~ ""~ ""~
Study Person Fax~ ""~ ""~ ""~ ""~ ""~ ""~ ""~
Study Person Address~ "Prospect Place, Plymouth, United Kingdom"~ "CEH Oxford, Oxford, United Kingdom"~ "San Diego State University, San Diego, California, United States of America"~"Argonne National Laboratory,
Argonne, Illinois, United States of America"~ "San Diego State University, San Diego, California, United States of America"~"San Diego State University, San Diego, California, United States of America"~"Prospect
Place, Plymouth, United Kingdom"~
Study Person Affiliation~ "Plymouth Marine Laboratory"~ "NERC Centre for Ecology and Hydrology"~"California Institute for Telecommunications and Information Technology"~ "Department of Computer Science, Mathematics
and Computer Science Division,"~ "California Institute for Telecommunications and Information Technology"~ "California Institute for Telecommunications and Information Technology"~ "Plymouth Marine Laboratory"~
Study Person Roles~ "principal investigator role"~"principal investigator role"~"principal investigator role"~"principal investigator role"~"principal investigator role"~"principal
investigator role"~
Study Person Roles Term Accession Number~ ""~ ""~ ""~ ""~ ""~ ""~ ""~
Study Person Roles Term Source REF~ ""~ ""~ ""~ ""~ ""~ ""~ ""~
```

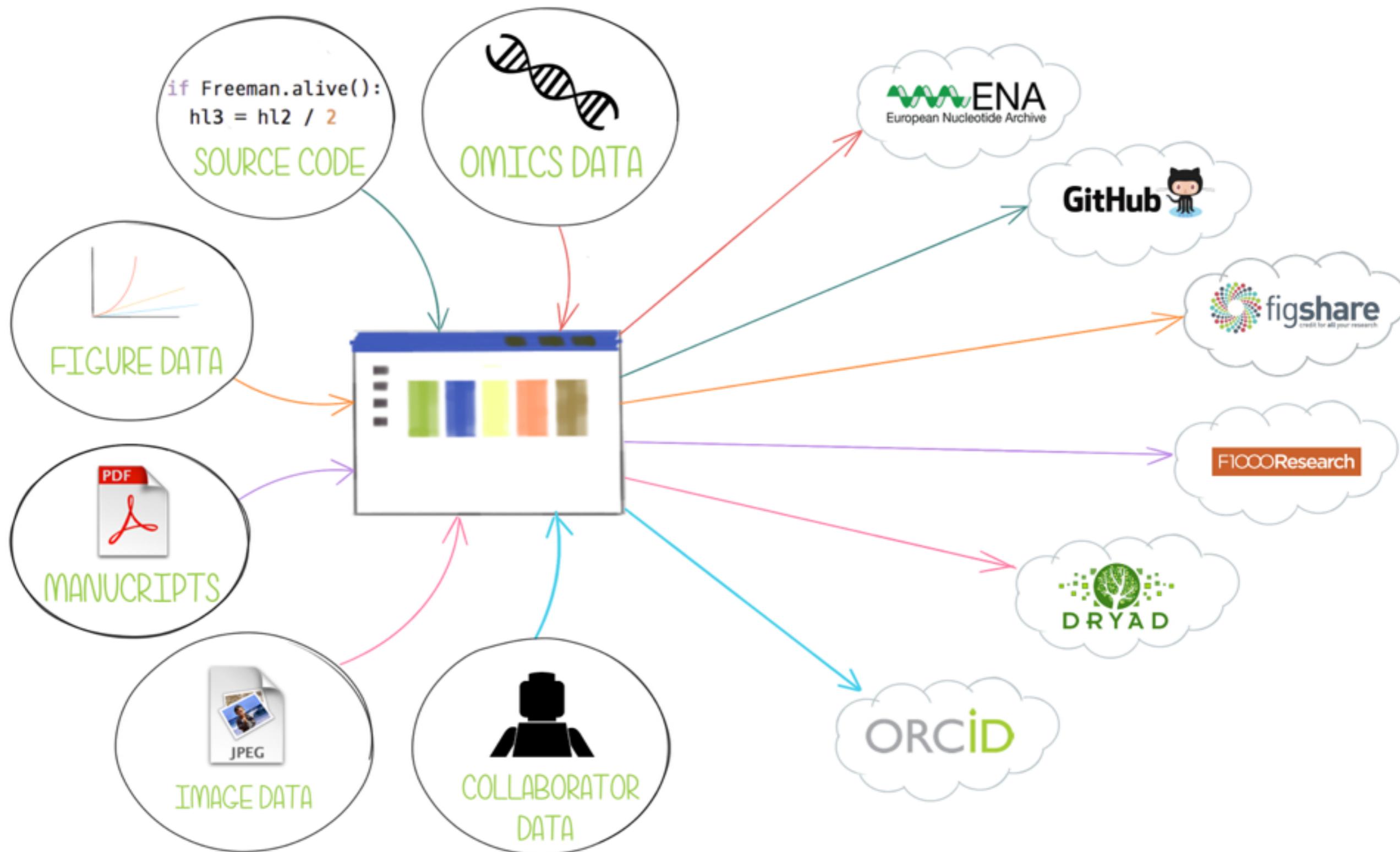



- ▶ Aggregation and Deposition Portal
 - ▶ Add research objects to profiles
 - ▶ Profiles deposited to public repositories
 - ▶ DOIs minted for logical groups
 - ▶ DOIs can be cross-referenced

COPO OVERVIEW



PROFILE OVERVIEW



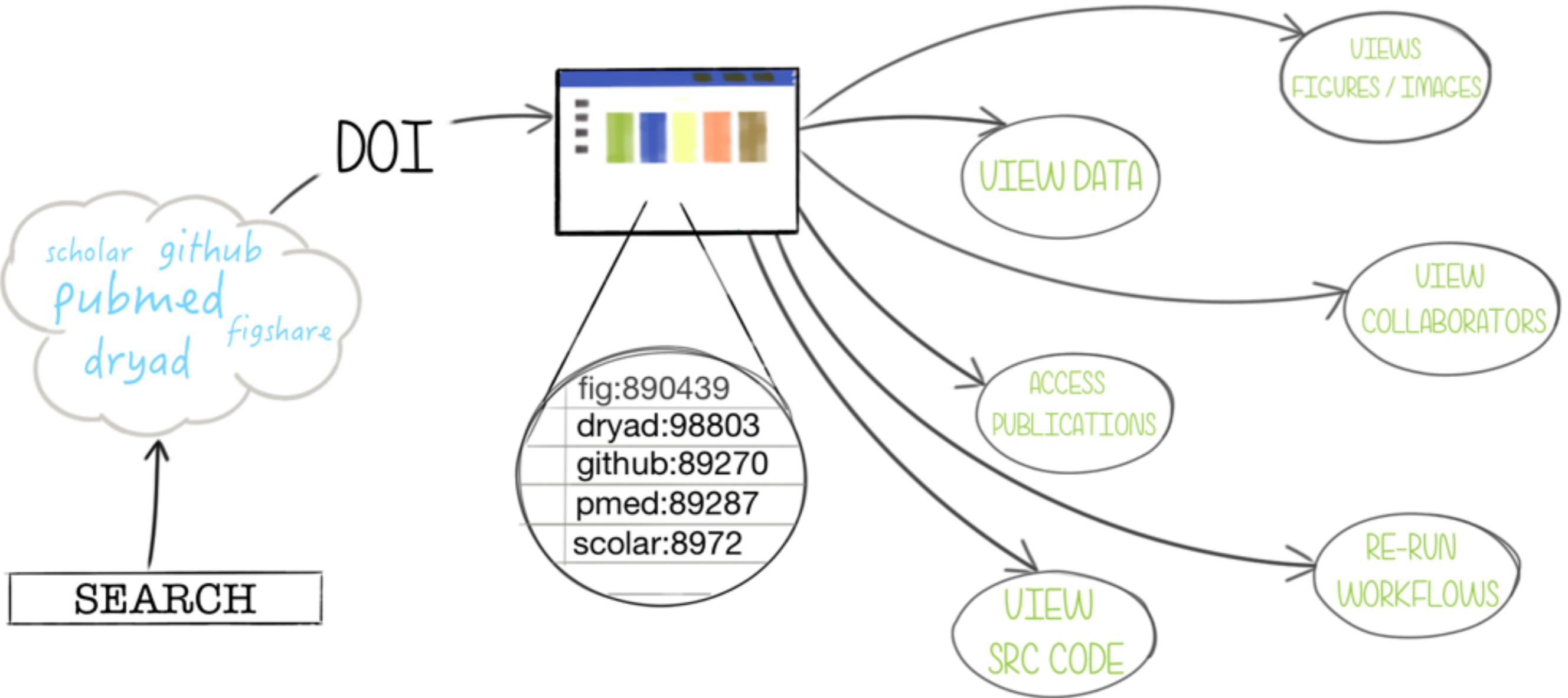
**...SCIENTISTS ARE LIKE
CATS, GIVE THEM MEAT
AND THEY WILL PROBABLY
TURN UP...PROBABLY**

Ari Asmi, University of Helsinki





- ▶ Tangible evidence of
 - ▶ Usages of their Data
 - ▶ Usages of their Source Code
 - ▶ Usages of Analysis Tools



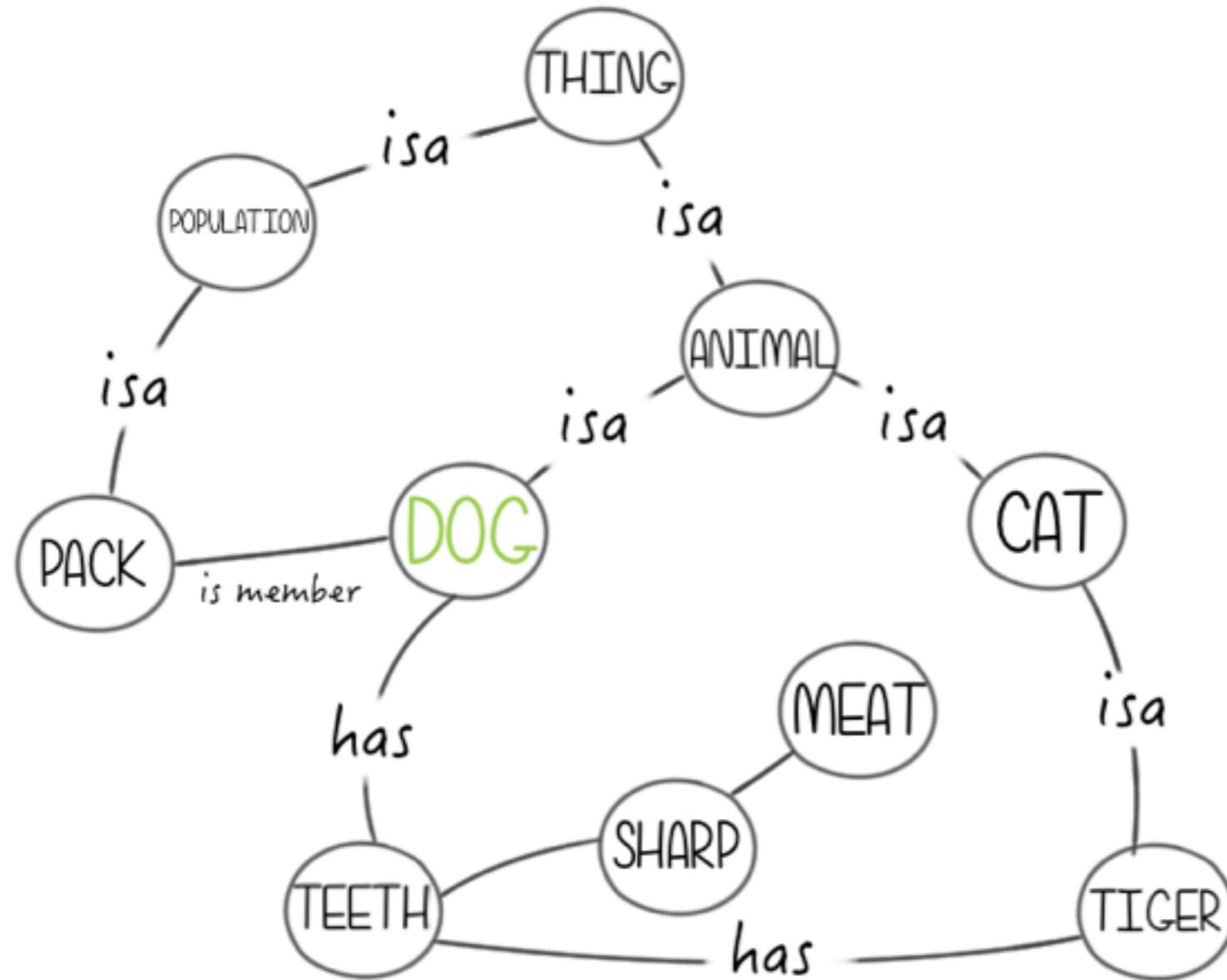


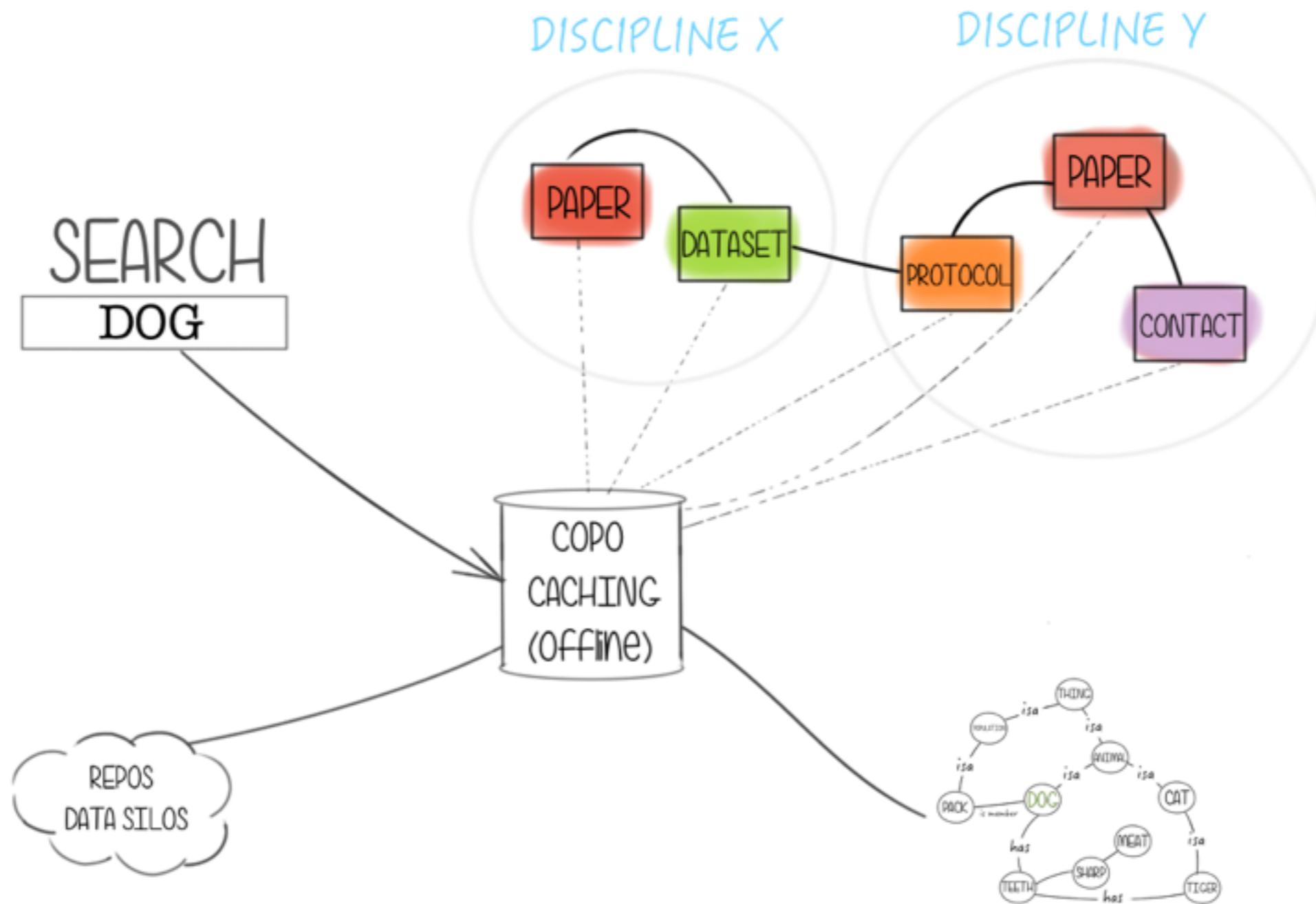
ORCID

 orcid.org/0000-0001-9649-5906

- ▶ One login for multiple services
- ▶ De-facto standard for services / repositories
- ▶ Provides access to full researcher profile
 - ▶ Work history, Publications, Grants etc.
- ▶ API provided

ONTOLOGIES







jQuery

MongoDB

Python

Django

COPO - Platform in Development

ISATools

Bootstrap

JSON-LD

Elasticsearch



Please sign in

Sign In

[Forgot Password?](#)

[sign up](#)

or

 Sign in with ORCID

 Sign in with Google



ORCID

COPO (Collaborative Open Plant Omics) ?

has asked for the following access to your ORCID Record



Get your ORCID iD



Allow this permission until I revoke it.

You may revoke permissions on your account settings page. Unchecking this box will grant permission this time only.

This application will not be able to see your ORCID password, or other private info in your ORCID Record. [Privacy Policy](#).

Don't have an ORCID iD? [Register](#).

Email or iD



Password



[Forgotten password?](#)

Deny

Authorize



COPO - Collaborative Open Plant Omics Status 3 Alerts Settings anthony

Home Search Orcid Profile Logout

http://orcid.org/0000_0001_8320_660X

Publications

TIDY: A Trust-Based Approach to Information Fusion through Diversity.
16th International Conference on Information Fusion
2013-07 | Journal_Article

Affiliations

Genome Analysis Centre, Norwich, GB
None, (None)
- to Present

University of Aberdeen, Aberdeen, GB
PhD, (Computing Science)
2010 - 10 to 2014 - 09



COPO - Collaborative Open Plant Omics

Status 2 Alerts Settings anthony

Home

Search

- Profiles
- Accessions
- Publications
- References

data	samples	submissions	publications	people
4 <i>files in profile</i>	2 <i>samples in profile</i>	2 <i>submissions in profile</i>	5 <i>publications in profile</i>	13 <i>people in group</i>
<ul style="list-style-type: none">upload Datalabel with metadatacreate bundles	<ul style="list-style-type: none">describe samplesrecycleshare with collaborators	<ul style="list-style-type: none">Upload to repositoriesShare with the worldGet cited	<ul style="list-style-type: none">share publicationspublish to journalsGet cited	<ul style="list-style-type: none">collaborateinteractshare



The screenshot displays a web application interface for 'Collaborative Open Plant Omics'. A modal window titled 'Add New Publication' is open, featuring a 'Help tips ON' toggle. A green callout box on the left contains the text: 'Context-based help tips and suggestions e.g., exploiting work history, similar work profiles, etc.' The modal form includes input fields for 'Author List', 'DOI', 'PubMed ID', and 'Status'. A 'Cancel' button and a 'Save' button with a download icon are at the bottom. In the background, a table lists publications with columns for 'PubMed ID' and 'Status'. One entry shows '3727177' with the status 'Indexed for PubMed'. A tooltip titled 'Author List' is visible over the 'Author List' input field, containing the text: 'Enter authors for the publication. Press the enter key to add to the list.'



The identification of genes / ENA-Study-Collection

Samples

[Add Sample](#)

Studies

[Add New Study](#) [Clone Existing Study](#)

Study Reference	Study Type	#Samples	Actions
Study-3	Metagenomics	0	Edit Delete
Study-1	Whole Genome Sequencing	0	Edit Delete
Study-2	Whole Genome Sequencing	0	Edit Delete

DESCRIBE YOUR SAMPLES



Add Study Sample ? ×

Sample Name

Source Name

NCBI Taxon

Sample Attributes

<input type="text" value="host"/>	<input type="text" value="value"/>	<input type="text" value="n/a"/>	<input type="checkbox"/>	<input type="button" value="Delete"/>
<input type="text" value="environment (feature)"/>	<input type="text" value="value"/>	<input type="text" value="n/a"/>	<input type="checkbox"/>	<input type="button" value="Delete"/>
<input type="text" value="environment (material)"/>	<input type="text" value="value"/>	<input type="text" value="n/a"/>	<input type="checkbox"/>	<input type="button" value="Delete"/>

Assign Sample to...

Study Reference	Study Type	# Samples	<input checked="" type="checkbox"/>
Study-3	Metagenomics	0	<input checked="" type="checkbox"/>
Study-1	Whole Genome Sequencing	0	<input checked="" type="checkbox"/>
Study-2	Whole Genome Sequencing	0	<input checked="" type="checkbox"/>

UI Forms partly driven by ISA specification. ISA formats and software suite enable experimental metadata attribution, management and conversion between metadata formats.



Clone Existing Study ?

×

Enter Study Reference

Existing Studies	View Pane	Cloned Pane
<ul style="list-style-type: none">Study-3 (Metagenomics)<ul style="list-style-type: none"><input checked="" type="checkbox"/> Study Type<input checked="" type="checkbox"/> Study Title<input type="checkbox"/> Comment[Study Funding Agency]<input type="checkbox"/> Study DescriptionSamples<ul style="list-style-type: none"><input checked="" type="checkbox"/> 0PublicationsContactsStudy-1 (Whole Genome Sequencing)Study-2 (Whole Genome Sequencing)	<p>Study Type: Metagenomics</p> <hr/> <p>Study Title:</p> <hr/> <p>Comment[Study Funding Agency]:</p> <hr/> <p>Study Description:</p> <hr/> <p>Samples</p> <p>Sample Name: test sample</p> <p>Source Name: test source</p> <hr/> <p>Publications</p> <hr/> <p>Contacts</p>	<p>Study Type: Metagenomics</p> <hr/> <p>Study Title:</p> <hr/> <p>Samples</p> <p>Sample Name: test sample</p> <p>Source Name: test source</p>

BRING ON YOUR DATA!



Study Samples

Assign Sample

Sample Name	Source Name	NCBI Taxon	
test sample	test source	test taxon	

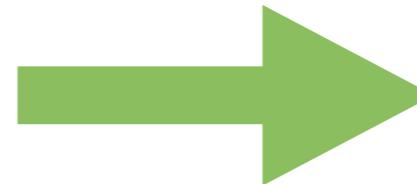
Chunked uploading breaks a large file into small chunks

Study Data Files

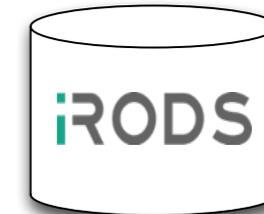
Select files...

1035_LIB8552_LDI7261_CGATGT_L007_R5

1149.91 MB of 2388.23 MB @ 205.89 MB/sec



*<https://www.irods.org/>



Study Publications

Add Publication

OR

Enter DOI or PubMed ID

Resolve

...allowing for access/sharing across multiple applications e.g., WheatIS.

BRING ON YOUR DATA!/2



Select files...

File	Sample	Actions
1035_LIB8552_LDI7261_CGATGT_L007_R5.fastq.gz ⓘ	<input type="text" value="Attach Sample..."/>	<div style="border: 1px solid #ccc; padding: 2px;">   </div>

Study Publications Send to Dropbox

OR

Study Contacts [E]



Study Data Files

Select files...

File	Sample	Actions
1035_LIB8552_LDI7261_CGATGT_L007_R5.fastq.gz  	<input type="text" value="Attach Sample..."/>	

transfer metadata

```
{
  "_id" : ObjectId("5637995f91ccd65fcafe4f94"),
  "transfer_status" : "success",
  "transfer_completed" : "2015-11-02 17:13:22.725923",
  "elapsed_time" : "82.34s",
  "request_user" : "toni",
  "file_path" :
  "/Users/etuka/Desktop/copo_uploads/chunked_uploads/
  2015_11_02/a7ef653936254c50bc10536987485d02.part",
  "file_size(bytes)" : "962686341",
  "transfer_commenced" : "2015-11-02 17:11:59.063049",
  "transfer_rate" : "93.53Mbps",
  "bytes_lost" : "0.00",
  "pct_completed" : "100"
}
```

Study Publications

Add Publications

Aspera (FASP™)-powered file transfer to remote repositories (e.g., European Nucleotide Archive)

Log/Queue Tools Help

Browse Change

Name	Size	Type	Date
<ul style="list-style-type: none"> / <ul style="list-style-type: none"> copo <ul style="list-style-type: none"> report 			
a7ef653936254c50bc1...	940,124 KB	part	Nov 2 17:13

DESCRIBE IT...WITH EASE!



Autocompletion and suggestions, again, by exploiting work history, similar work profiles, etc.

Assign File Details ? ×

1. Library layout 2. Library strategy 3. Library source

Library source

- GENOMIC
- GENOMIC SINGLE CELL
- TRANSCRIPTOMIC
- TRANSCRIPTOMIC SINGLE CELL
- METAGENOMIC
- METATRANSCRIPTOMIC
- SYNTHETIC
- VIRAL RNA
- OTHER

Next

File	Sample
1035_LIB8552_LDI7261_CGATGT_L007_R5.fastq.gz ?	Attach Sample...



Link to analysis services such as **Galaxy** and **iPlant**, as well as share/reuse metadata with other systems such as **WheatIS**

Exploit **linked metadata** to attach semantic meaning to research artefacts

**EXCLUSIVE TO
PLANT SCIENCE?**

Support for more repositories and publishing platforms e.g., **GigaScience**, **F1000**, **Scientific Data**, **Dryad**

Semantic inferences over artefacts to provide a richer search experience



COPO Workshop

23 & 24 June, 2015 @ TGAC

1st Quarter of 2016

ACKNOWLEDGEMENTS



Rob Davey

Nicola Soranzo



Susanna-Assunta Sansone

David Johnson

Philippe Rocca-Serra

Xingdong Bian

Vicky Schneider



Simon Tyrell

Tim Stitt

Alejandra Gonzalez

Alfie Abdul-Rahman



Ruth Bastow



Paul Kersey