

Open Science

Idafen Santana, Alexander García, Oscar Corcho,
Olga Giraldo, Carlos Badenes, Esteban González

openscience@delicias.dia.fi.upm.es
www.oeg-upm.net/openscience

Ontology Engineering Group
Universidad Politécnica de Madrid



The team



Idafen Santana
Research Fellow



Olga Giraldo
PhD Student



Oscar Corcho
Full Professor



Daniel Garijo
Research Fellow
(collaborator)



Esteban González
Research Assistant



Alexander García
Senior Research
Fellow



Carlos Badenes
Research Assistant



Open Science

What is Open Science?

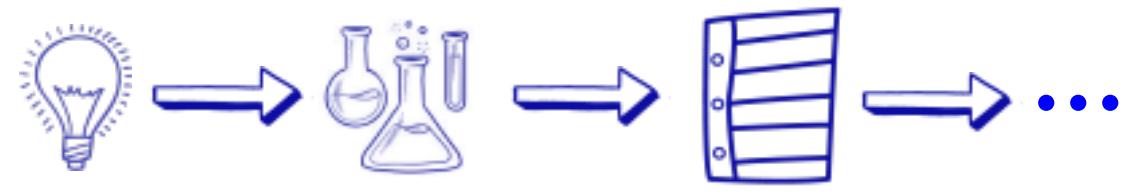


Opening the scientific process to a
broader audience, promoting **collaboration** and **participation**



Open Science

What is Open Science?



**Opening the scientific process to a
broader audience, promoting collaboration and participation**



Open Science

Opening the scientific process to a
broader audience, promoting **collaboration and participation**

Students Citizens
Scientists
Companies Educators



Open Science

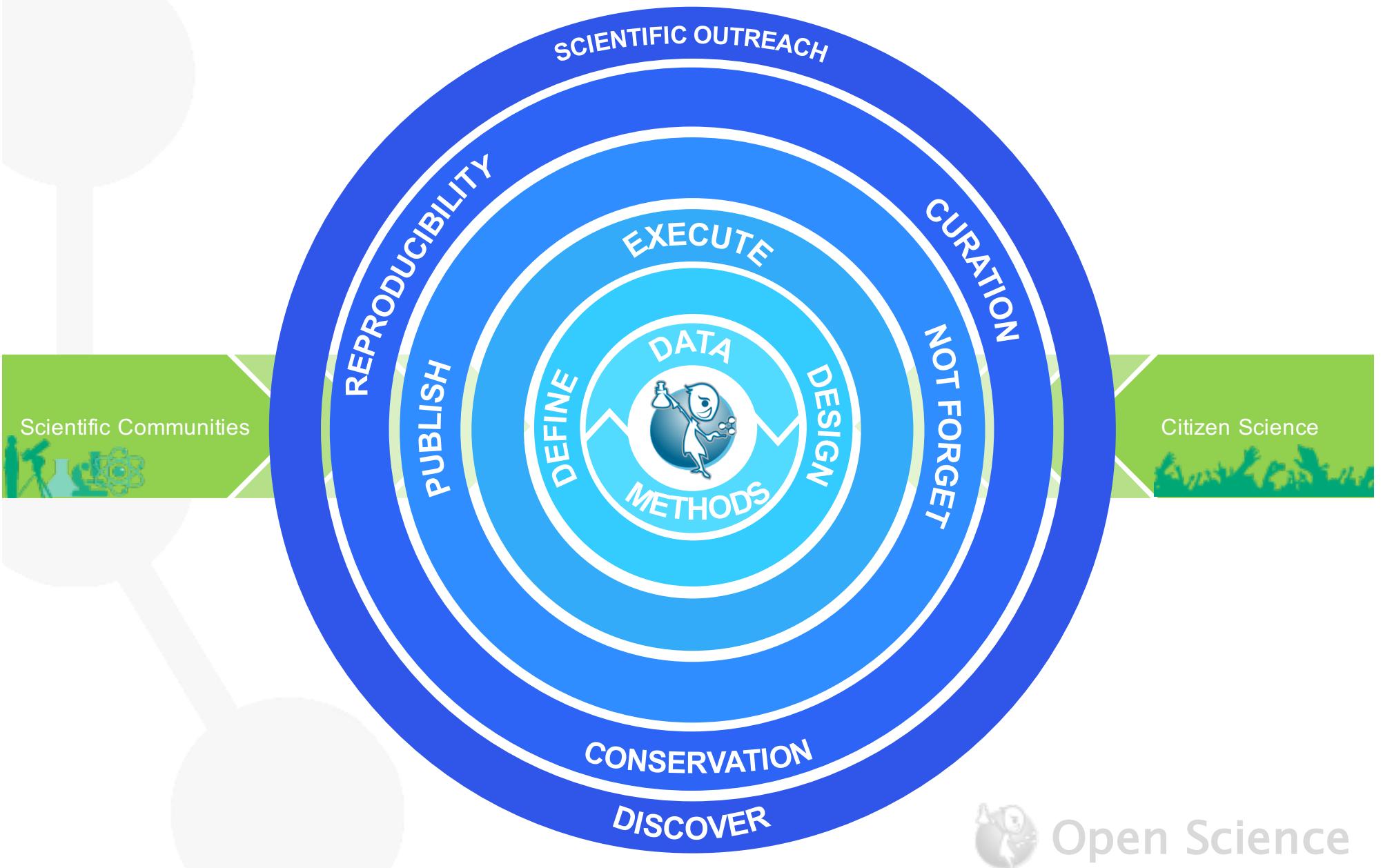
What is Open Science?

**Opening the scientific process to a
broader audience, promoting collaboration and participation**

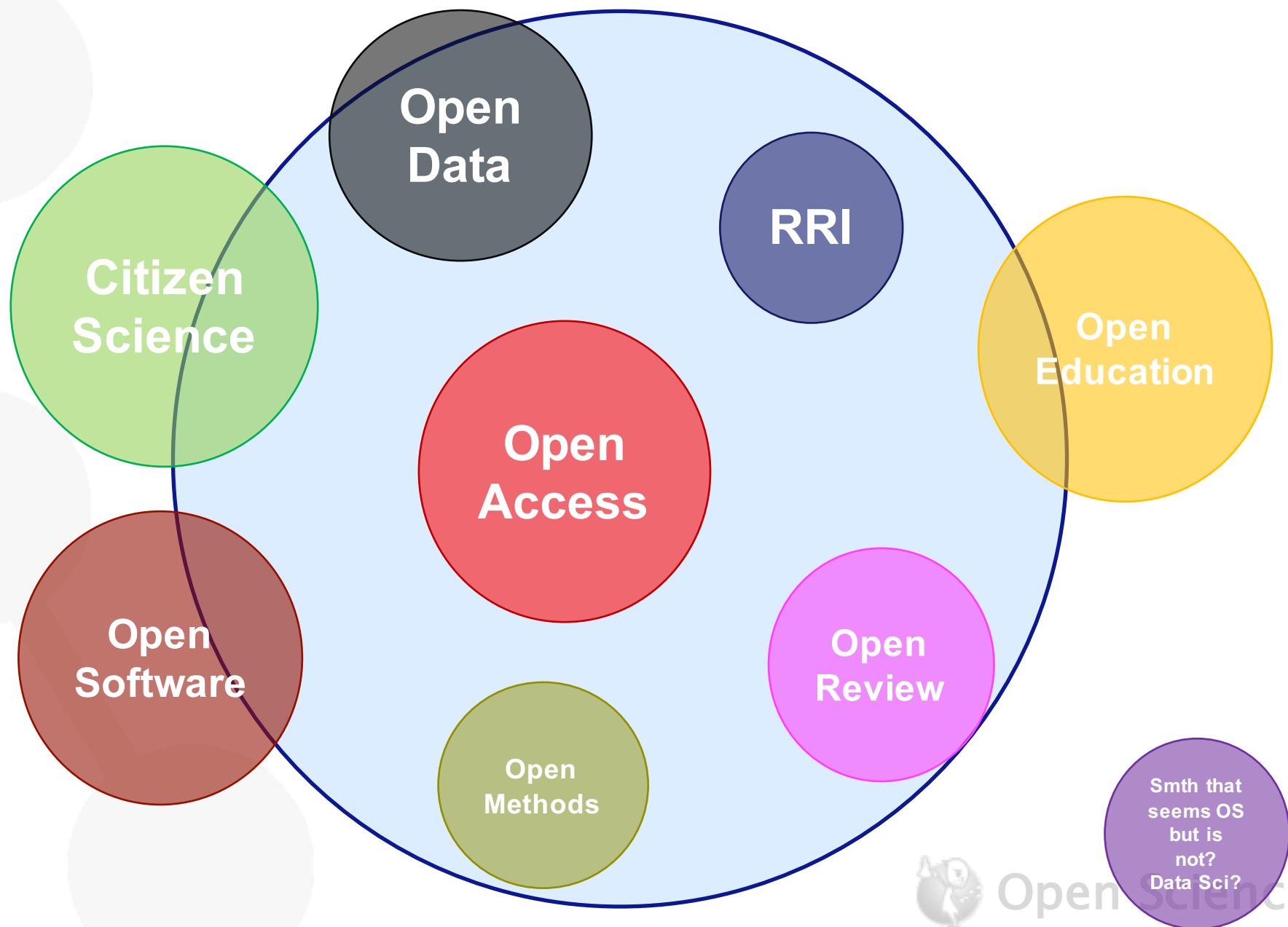


Open Science

Open Science Lifecycle



Open Science and other terms?



Contributions

 **W3C® PROV-O: The PROV Ontology** Provenance representation. W3C working group (<https://www.w3.org/TR/prov-o/>)



Scientific results publication on the web
(<http://www.researchobject.org/>)



Detection of common motifs among scientific workflows
(<https://github.com/dgarijo/FragFlow>)



Reproducibility of scientific execution environments
(<http://vocab.linkeddata.es/wicus/wicus/>)



Information retrieval from laboratory protocols
(<http://vocab.linkeddata.es/SMARTProtocols/>)



Analysis of large collections of scientific documents
(<http://librairy.github.io/>)



A Collective Awareness Platform for Promoting Dark Skies in Europe (<http://www.stars4all.eu/>)





Open Science



- Meteor Counter
 - Objective: Measure the meteor activity during a meteor shower
 - Features:
 - Web Application
 - Responsive Interface
 - Use of GPS
 - Use of accelerometer (optional)



Open Science

- Consensus tool
 - Objective: Create / Adapt a citizen collaboration tool where users can propose ideas to reduce light pollution
 - Features:
 - Based on Consul Application (Decide Madrid)
 - Users can propose initiatives or create debates
 - These initiatives can be voted.

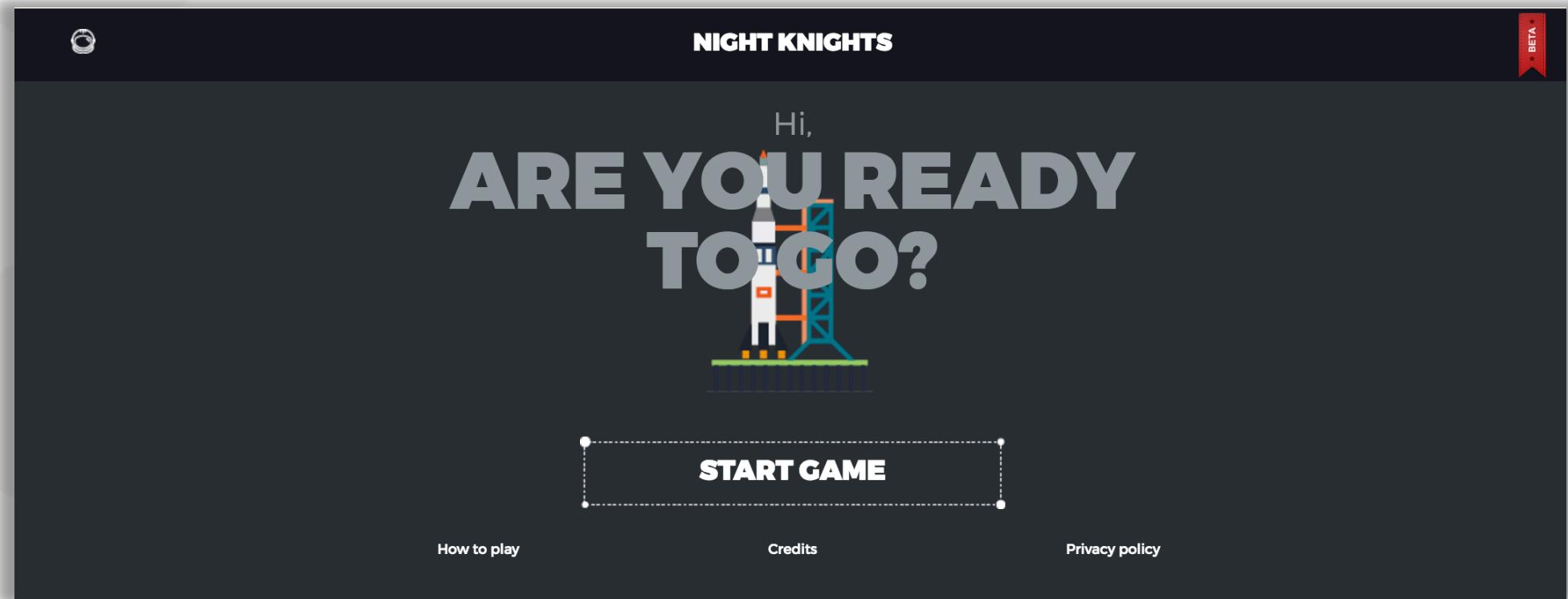


Open Science

- Lost At Night
 - Objective: Geolocation of cities using NASA images
 - Features:
 - Based on Lost At Night application
 - Web application
 - Responsive interface



Open Science



2010 Apr 7:00

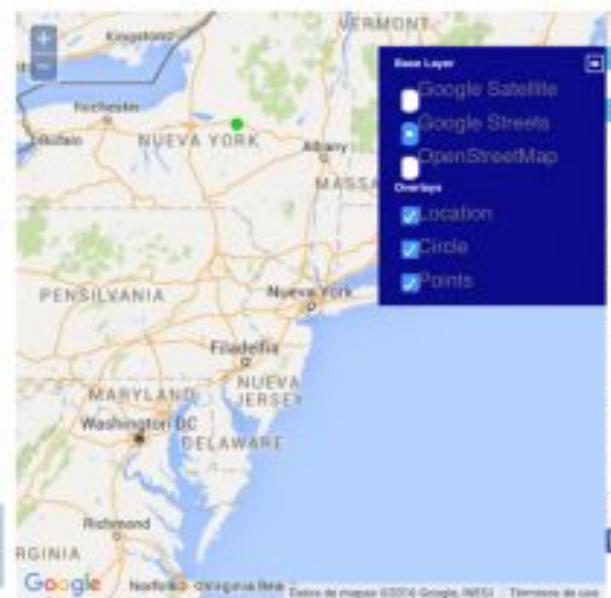


Open Science

[Save](#) [No photo](#) [I don't know](#)

Please, be patient with large ISS Images. [Image Sharpness](#): [Bad](#) [Medium](#) [Fair](#) [Cloudy](#) [None](#) [Less than 5%](#) [More than 5%](#)

Green point indicates ISS location at picture time. Use the [+] and [-] buttons to Zoom in and Zoom Out. Keep Shift + Left mouse button pressed while moving the mouse will rotate the image. Please do not just guess at a location; that wastes time when verifying results. Please only place a mark if you are absolutely sure that its location is accurate and can be later geo-referenced point-by-point, else press the I-Don't-Know button. [ISS picture data - Earth Science and Remote Sensing Unit, NASA-Johnson Space Center, "The Gateway to Astronaut Photography of Earth."](#)



[Tweet](#)



Open Science

Building a Gold Standard for
Experimental Protocols

Metodology and lessons learned



Open Science

Questions about a Gold Standard Corpus (GSC) of protocols

What?

A collection of protocols in electronic form manually annotated by experts.

Why?

For the evaluation of gazetteers and rules designed to annotate automatically protocols.

Is there?



How to build a GSC?

- Are there Good practices for building a GSC?
somebody knows any?
- How to determine the optimal size of a corpus?
- Does the quality of a GSC depend on the experience of the annotators?

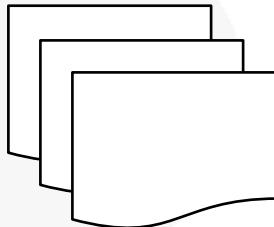
Time, tactic (method) and Money



Open Science

*Calculate the overall agreement amongst people annotating entities related to **Samples**, **Instruments**, **Reagents** and **Objective**.*

Materials



- Bio-Protocols = 9
- Nature = 19
- CSH-Protocols = 7
- **TOTAL = 35**

+



- Curso BIOS 2016, Colombia - 8 annotators
- Universidad del Valle, Colombia – 11 annotators
- Japon (Database Center for Life Science (DBCLS), Robotic Biology Institute (RBI), Spiber, Yachie-Lab, Universidad de Tokyo). - 14 annotators
- **TOTAL = 33**

+



The SMART Protocols Annotation Tool

<http://labs.linkingdata.io:9000/dist/dev/#/login>



- Guidelines about What and Who annotate
- User guide of the SMART Protocols Annotation Tool

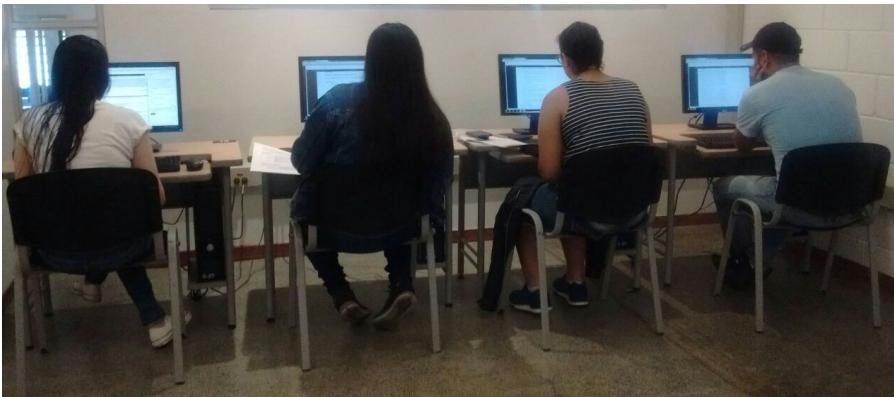


Open Science

Methods



Training



Annotation session



Gift

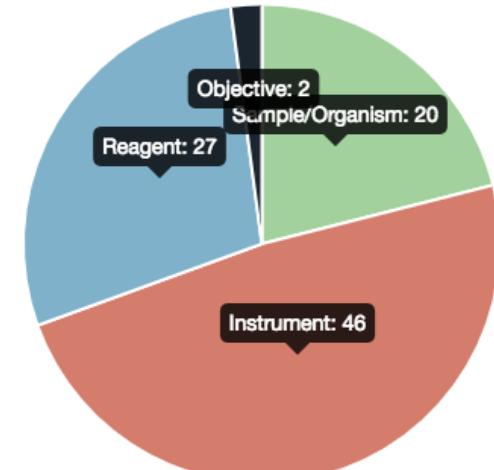
What to annotate?

- Sample / Organism
- Instruments
- Reagents
- Objective

How to annotate

- Read the document.
- Mark each occurrence of an entity.
- Reduce the noise in the annotation.
- Add a comment indicating an annotation or decision that was hard to make.
- Time to solve doubts.
- Notify the finalization of the annotation task.

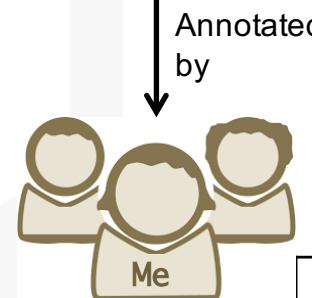
Tag distribution



<http://labs.linkingdata.io:9000/dist/dev/#/dashboard/charts/chartjs>

Pre-processing the data

... (poxvirus, herpesvirus, mimivirus, etc).....
.....
TAE buffer (Sambrook and Russell, 2001), and 5x Agarose Gel loading dye (QIAGEN, catalog number: 239901).
Thermal-cycler for PCR.



	annot1	annot2	annot3
DNA virus	X	✓	✓
poxvirus	✓	✓	✓
herpesvirus	✓	✓	✓
mimivirus	✓	✓	✓
TAE buffer	✓	✓	✓
5x Agarose gel	✓	✓	✓
Thermal-cycler	✓	✓	✓

	Entities	sample	instrument	reagent	objective
Sample	Neural cell	3	0	0	0
	neural stem cells (NSCs)	3	0	0	0
Instrument	Cell culture centrifuge	0	3	0	0
	cell culture incubator	0	3	0	0
	Microscope	0	3	0	0
	Millicell culture plate inserts 8-?m pore size	0	3	0	0
reagent	B27 supplement	0	0	3	0
	DMEM/F12	0	0	3	0
	FGF2 neutralizing antibody	0	0	3	0
	glucose	0	0	3	0
objective	Here we describe two migration assays, a matrigel migration assay and a Boyden chamber migration assay, which allow the in vitro assessment of neural migration under defined conditions (Ladewig, Koch and Brüstle, 2014).	0	0	0	3

Preliminary results

	Entities	sample	instrument	reagent	objective
Sample	Neural cell	3	0	0	0
	neural stem cells (NSCs)	3	0	0	0
Instrument	Cell culture centrifuge	0	3	0	0
	cell culture incubator	0	3	0	0
	Microscope	0	3	0	0
	Millicell culture plate inserts 8-?m pore size	0	3	0	0
reagent	B27 supplement	0	0	3	0
	DMEM/F12	0	0	3	0
	FGF2 neutralizing antibody	0	0	3	0
	glucose	0	0	3	0
objective	Here we describe two migration assays, a matrigel migration assay and a Boyden chamber migration assay, which allow the in vitro assessment of neural migration under defined conditions (Ladewig, Koch and Brüstle, 2014).	0	0	0	3

Fleiss Kappa for 3 raters = 1.0

	entities	sample	instrument	reagent
Reagent - Sample/Organism	Ac-omega viral DNA	1		2
	baculoviral	1		2
	DNA insert	2		1
	I-Sce I meganuclease	1		2
Sample/Organism	Insect cells	3		
Instrument	spinner		3	
	Centrifuge		3	
	Flask		3	
Reagent	IPL-41 powdered Liposome formulation Phenol:chloroform			3

Fleiss Kappa for 3 raters = 0.755

k	Interpretation
< 0	Poor agreement
0.01 – 0.20	Slight agreement
0.21 – 0.40	Fair agreement
0.41 – 0.60	Moderate agreement
0.61 – 0.80	Substantial agreement
0.81 – 1.00	Almost perfect agreement



Open Science

Good guidelines Including What and Who annotate

➡ Training time

↗ Number of annotated protocols

↗ Number of entities annotated

≡ Degree agreement



Open Science

- Annotation event at the Universidad Santiago de Cali, Colombia (Nov. 17, 2016)
- Annotation event at the Universidad Icesi, Cali, Colombia.
- Calculate *the overall agreement by using Fleiss Kappa*
- *Publication of the gold standard*
- evaluate the gazetteers and rules designed to annotate automatically protocols (precision, error rate and recall)



Open Science

WILEY

Consultancy for the Wiley Online Library

- is a global publishing company
- Founded 1807
- Headquarters location, Hoboken, New Jersey
- Academic publications in several subjects

My tasks:

- Support in the enrichment of the controlled vocabulary
- Research, specification of use cases, documentation, discussions and decision-making process
- Contract (Oct. 2016 – Jan. 2016).



Open Science



Open Science



- Semantic annotation
 - Tags coming from an ontology
- Tagging
 - Adding a label to an object
- Collaborative
 - Crowdsourcing
 - Tasks



Open Science

Is it that simple?

- No...
- You need
 - Documentation for annotators
 - Measurable targets
 - What are the annotations going to be used for?
 - How are you going to control the process?
 - Crowd vs no crowd?
 - A crowd of one?



Open Science

- More than 150 annotation tools
- More than a 150 annotation formats
- More than a 150 ways to annotate
- Less than 150 usability studies
- Less than 150 reusable annotation experiences
- Less than 150 documented experiences
- 150 experiences confusing annotation with commenting



Open Science

Working with <https://hypothes.is>

- Why?
 - Simplicity
 - A well funded project, for now
 - An incipient community, for now
 - Documentation
- Saves time
 - PDF
- Adaptable
 - To a point...



Open Science

Working with <https://hypothes.is>

- Semantic Annotations
 - Bioportal
- Targeted annotations
 - Specific “semantic tags”
 - Minimal amounts of Information models
- Adapting it to our experiments
- Stripping it off “commenting”



Open Science

- PLOS
 - Annotating specific facets
 - “Crowd”
 - Corpus of identifiable elements for further NLP, deep learning and dive learning experiments
- ZPID
 - Psychology paper of the future
- Bio-Protocols
 - SIRO elements
 - Feeds the onto development process



Open Science

ZPID the Psychology paper of the future

- With all the Jazz of all papers of the future



Open Science

Fri, 20 Dec 2013 in *The European Journal of Counselling Psychology*

Assessment of Family Functionality Among the Elderly With Chronic Illness

Claudia Balula Chaves Odete Pereira Amaral Paula Alexandra Nelas

Emilia Carvalho Coutinho Rui Manuel Dionisio

DOI: 10.5964/ejcop.v2i2.31

Abstract

The family APGAR scale was developed by Smilkstein, Ashworth, and Montano (1982). The satisfaction assessment of the elderly with chronic illness regarding family is essential. This study aims to describe the socio-demographic and clinical profile of elderly people with chronic illness and correlate with perceived family support. This is a cross-sectional,

kingdata.io:5001/static/jats/index.html?url=http://labs.linkingdata.io:5001/files/ejcop.v2i2.31.xml



Fri, 20 Dec 2013 in *The European Journal of Counselling Psychology*

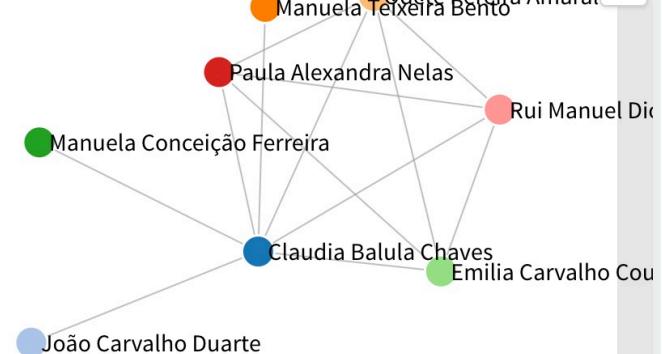
Assessment of Family Functionality Among the Elderly With Chronic Illness

Claudia Balula Chaves Odete Pereira Amaral Paula Alexandra Nelas
Emilia Carvalho Coutinho Rui Manuel Dionisio

DOI: 10.5964/ejcop.v2i2.31

Abstract

The family APGAR scale was developed by Smilkstein, Ashworth, and Montano (1982). The satisfaction assessment of the elderly with chronic illness regarding family is essential. This study aims to describe the socio-demographic and clinical profile of elderly people with chronic illness and correlate with perceived family support. This is a cross-sectional,



Name	Claudia Balula Chaves	Odete Pereira Amaral	Paula Alexandra Nelas	Emilia Carvalho Coutinho
Claudia Balula Chaves	2	1	1	1
Emilia Carvalho Coutinho	1	1	1	0

Fri, 20 Dec 2013 in *The European Journal of Counselling Psychology*

Assessment of Family Functionality Among the Elderly With Chronic Illness

Claudia Balula Chaves Odete Pereira Amaral Paula Alexandra Nelas

Emilia Carvalho Coutinho Rui Manuel Dionisio

DOI: [10.5964/ejcop.v2i2.31](https://doi.org/10.5964/ejcop.v2i2.31)

Abstract

The family APGAR scale was developed by Smilkstein, Ashworth, and Montano (1982). The satisfaction assessment of the elderly with chronic illness regarding family is essential. This study aims to describe the socio-demographic and clinical profile of elderly people with chronic illness and correlate with perceived family support. This is a cross-sectional,

Contents Fig Public

DOI
The article was received on Fri, 20 Dec 2013.

Copyright & License
This is an open-access article Attribution License, which medium, provided the orig

Author
Claudia Balula Chaves
Portugal
For correspondence
claudiachaves21@gmail.com

Author
Odete Pereira Amaral
Portugal

zpid_showcase 1 months ago
The family APGAR scale was developed by Smilkstein, Ashworth, and Montano (1982). The satisfaction assessment of the elderly with chronic illness regarding family is essential.
this is a test by alex
CHRONIC SCALE FAMILY
Edit Delete Reply Link
<http://labs.linkingdata.io:32768/a/AVczwza>

zpid_showcase 1 months ago
The family APGAR scale was developed by Smilkstein, Ashworth, and Montano (1982). The satisfaction assessment of the elderly ... More
Nice
SCALE FAMILY PEOPLE AGE FUNCTIONAL CHRONIC GROWTH
FUNCTIONALITY FEMALE QUESTIONNAIRE DATA MEAN
Edit Delete Reply Link
<http://labs.linkingdata.io:32768/a/AVczxAjE>

zpid_showcase 1 months ago
Portugal, diagnosed with chronic illness (77.9% cardiovascular);