

# Research Data Services: *Can Publishers Improve the Preservation and Usage of Research Data?*

Anita de Waard

VP Research Data Collaborations



*Elsevier*

Research Data Services

*stm*

International Association of  
Scientific, Technical & Medical Publishers  
The Voice of Academic and Professional Publishing

# Elsevier Research Data Services: Goals

## A. Increase Data Preservation:

Help increase the **amount and quality** of data preserved and shared

## B. Improve Data Use:

Help increase the **usage** of the data shared by increasing interoperability

## C. Develop Sustainable Models:

Help measure and deliver **credit** for shared data, and enable **sustainable** infrastructures.

# Elsevier Research Data Services: Principles

All data stays **open**

URLs, front end etc. stay where they are

Collaboration **tailored** to repositories' needs

Transparent business models –  
consultancy/service model

Very small, **nimble** department

**2013: Funding pilots to support a  
feasibility study: is this a market for us?**

# Elsevier Research Data Services: Goals

## A. Increase Data Preservation:

Help increase the amount and quality of data preserved and shared

## B. Improve Data Use:

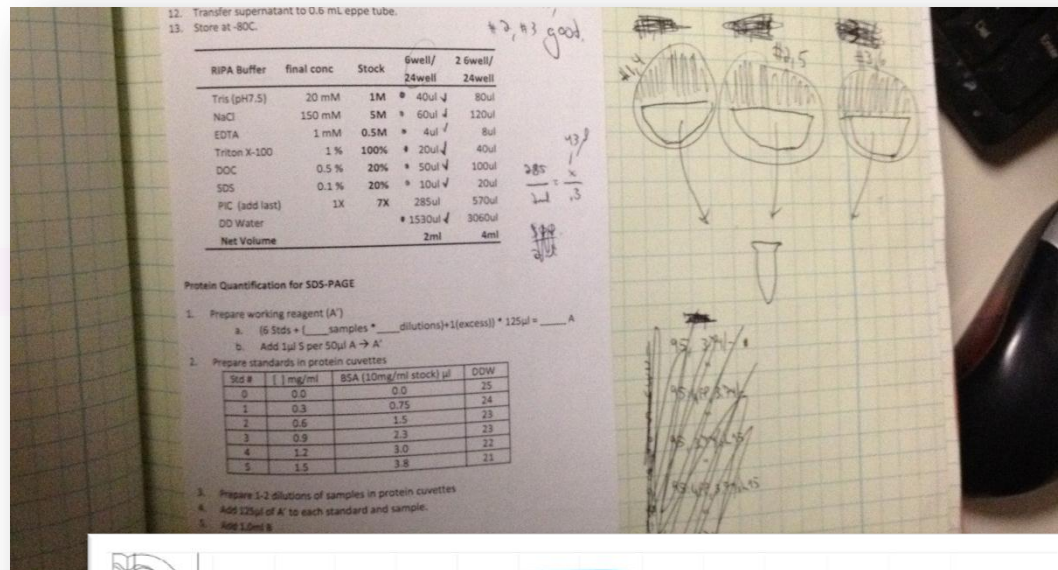
Help increase the **usage** of the data shared by increasing interoperability

## C. Develop Sustainable Models:

Help measure and deliver **credit** for shared data, and enable **sustainable** infrastructures.

# Pilot # 1: Metadata storage @ CMU

From paper notebooks



to app-driven metadata:

**Matt Geramita** CMU Urban Labs

ID: D359945A-CCF5-4983-0C60-DCC3A4B24F97::S1::C1::E1::R1

**Electrode ID: E1**

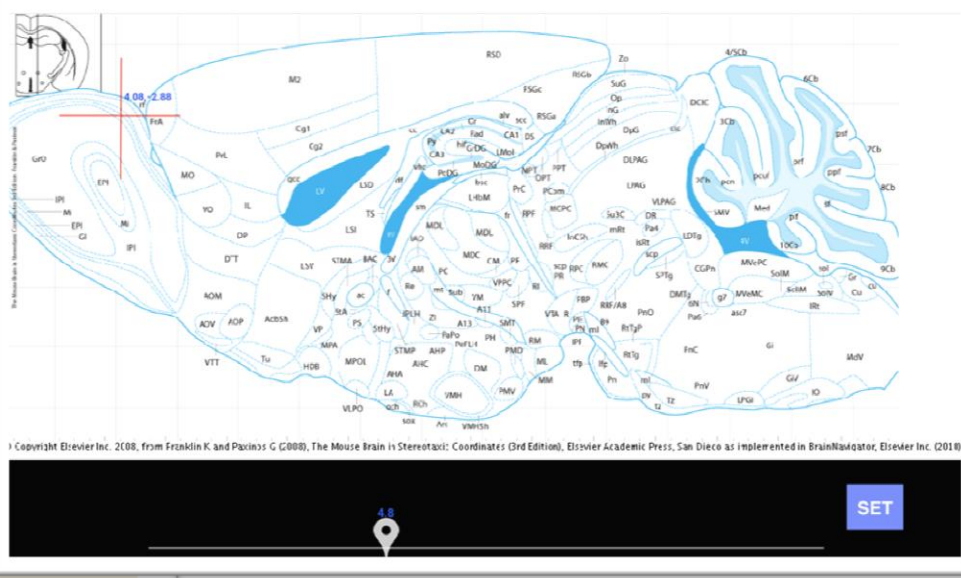
Single-cell soma Current clamp

C1 Voltage clamp

55 ohms Solution in electrodes

0.1% Mg in Ringers

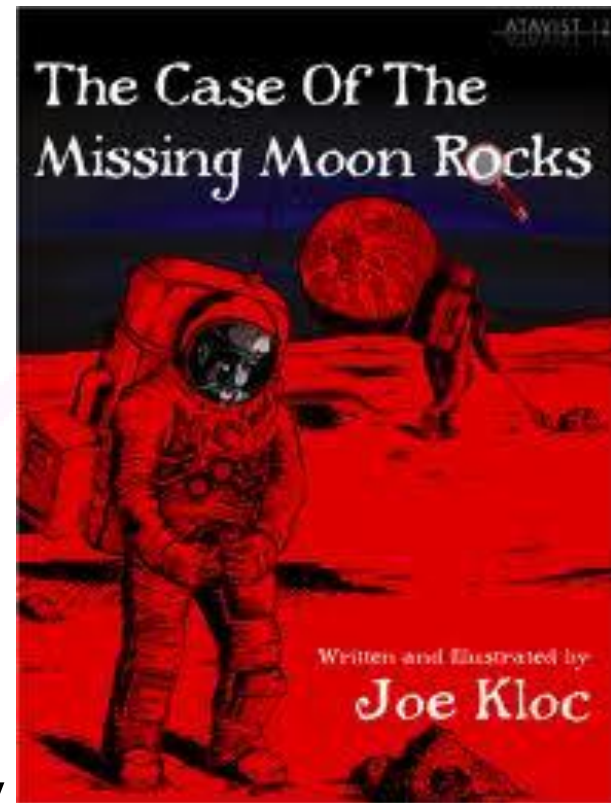
Electrode comments



# STM

# Pilot # 2: Data Preservation w/NASA

- Moon rock sample data:  
now stored as PDFs!
- With IEDA:  
add to PetDB database
- Train curators:  
from papers to DB entry
- Teaching scientists to be curators!  
What skills are needed?



# Pilot #3: Data Rescue Challenge (with IEDA)

Inviting data rescue projects in geosciences!

- Create overview of projects
- Alert geoscientists to our interest
- Start to create a community re. rescue

*Geosciences  
International  
Data Rescue  
Award 2013*



\$5000 reward



**STM**

# Elsevier Research Data Services: Goals

## A. Increase Data Preservation:

Help increase the **amount and quality** of data preserved and shared

## B. Improve Data Use:

Help increase the usage of the data shared by increasing interoperability

## C. Develop Sustainable Models:

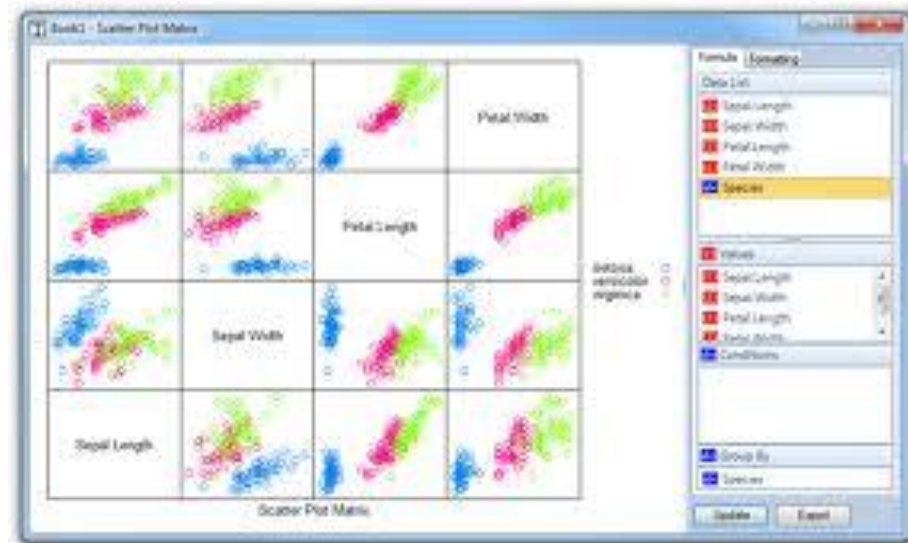
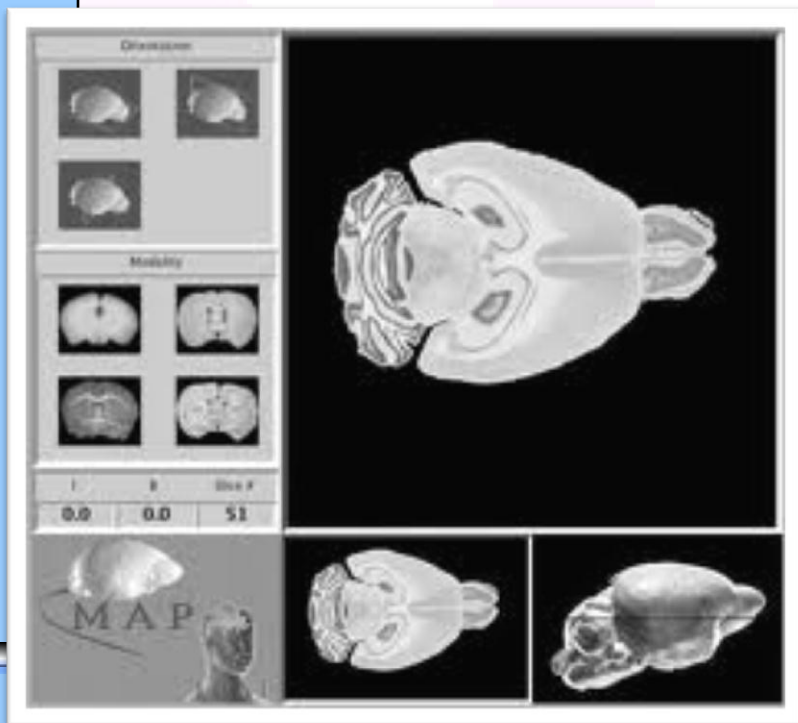
Help measure and deliver **credit** for shared data, and enable **sustainable** infrastructures.



# Pilot #4: Data Dashboard @ CMU

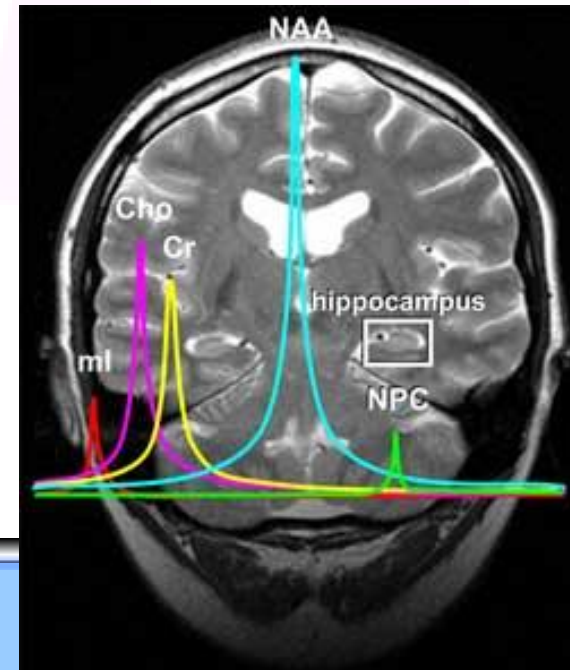
Once metadata is added: allow access to multiple labs' data, play with all data -

**Enable demonstrably better science!**



# Pilot #5: ImageVault, with Duke CIVM

- Get 3D image data into common format, resolution, annotated to allow comparison
- View other image data sets & do analytics
- Create funding for 3D image sets: free for raw data – subscriptions of analytics.



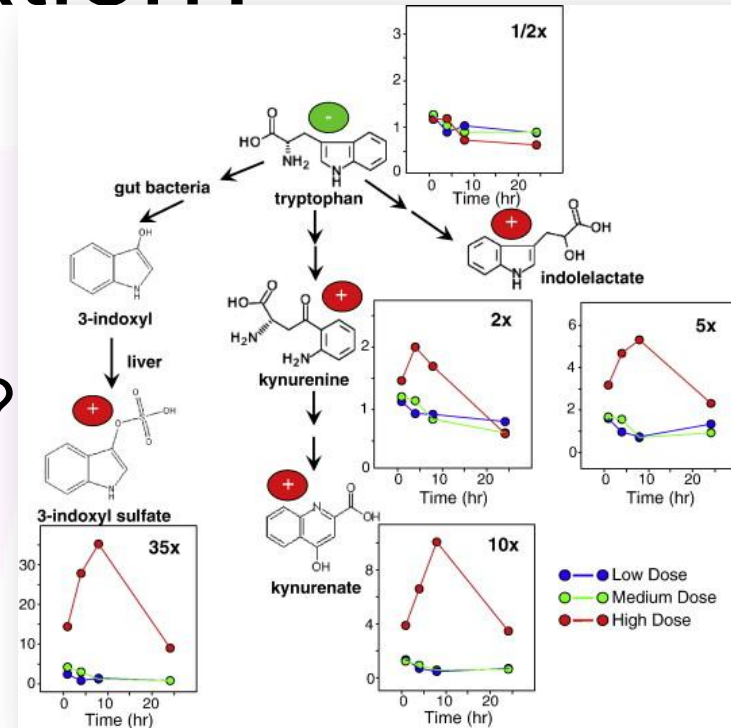
# Pilot # 6: Metabolomics Data Integration?

## Metabolomics:

- Look at processes in cell
- How does chemistry work?

## Still early days:

- How to combine cellular & metabolomics data?
- Where & how is compound information needed?
- Where is the data and how can we connect it?



# Elsevier Research Data Services: Goals

## A. Increase Data Preservation:

Help increase the **amount and quality** of data preserved and shared

## B. Improve Data Use:

Help increase the **usability** of the data shared by increasing interoperability

## C. Develop Sustainable Models:

Help measure and deliver credit for shared data, and enable sustainable infrastructures.

# Three Big Questions:

## 1. Where do you store your data?

- Generic repositories (Dryad, DataCite, Dataverse etc)
- Domain-specific repositories (PDB, EarthChem, TAIR, ...)
- Institutional repositories?

2. Who pays for all this?

3. What role can publishers play?

# Three Big Questions:

1. Where do you store your data?

**2. Who pays for all this?**

- Agencies demote funding infrastructure
- As more data is added, work increases!
- What are incentives? Merit, attribution?

3. What role can publishers play?

# Three Big Questions:

1. Where do you store your data?

2. Who pays for all this?

**3. What role can publishers play?**

- We do high-quality/-volume annotations
- We know scientific information
- We are good at long-term preservation and sustainable models.



*Elsevier*

# Research Data Services

## A) Data Preservation:

Pilot #1: Add metadata during experiment

Pilot #2: Rescue lunar sample data

Pilot #3: Data rescue challenge

## B) Data Usage

Pilot #5: Data Dashboard

Pilot #6: Image Vault

## C) Sustainable models:

1. Where do you store your data?

2. Who pays for all this?

3. What role can publishers play?



# Time's Up!

## Questions?

Anita de Waard

Elsevier Research Data Services

[a.dewaard@elsevier.com](mailto:a.dewaard@elsevier.com)

@anitawaard

<http://researchdata.elsevier.com/>



*Elsevier*

Research Data Services

**STM**