# THE SCIENTIFIC INFORMATION LIFECYCLE

(of "little science" research)

## hello, nice to meet you.

# I study scientific collaboration networks



## What is CENS?

- NSF Science & Technology Center 2002-2012
- Sensor network research
  - "We envision a world where researchers, students, industry and government routinely use distributed sensor and actuator networks to understand and control both natural and artificial systems."
- CENS community
  - 300+ researchers
  - 5 research institutions





### **CENS** research areas



# A CENS deployment



# Scholarly production at CENS

Manuscript type	
Conference proceedings	400
Journal article	189
Book chapter	18
Book	1
Number of authors	
1	59

	59
2	155
3	158
4	94
5	59
6	32
7	19
8	13
9	5
10+ (where 14 is the maximum number of authors found)	14

# How is CENS research different from AA & HEP?

- Big Science vs. little science
- Disciplinary heterogeneity
- Data heterogeneity
- No open access and data sharing culture
- Lack of technological infrastructure

# How do you build an information infrastructure from scratch?



# What kind of information?



# 3 specialized repositories

### CENS Deployment Center

CINS

Welcome to the CENS Deployment Center e CENS DEPLOYMENT CENTER (CENSDC) supports CENS research by enabling pre-deployment anning (task lists, equipment lists, participant lists, etc.) and the collection of post-deployment sons learned. Deployments are searchable by locations, technologies, equipment, and people The CENSDC provides a cross-linkable source of contextual information for CENS data stored elsewhere. If you do not have a login, please register with us. The CENSDC strives to be a repository of information to halp future dealowments.

### CENS Deployments Worldwide



### Contextual data

### SensorBase a project by CaNS

### Sensor Data - Slog. Share. Use.

Why SensorBase?

SensorBase provides you with a way to publish, share, and manage your sensor data much in the same way that you can publish, share, and manage journal entries in a blog. Also, data from many different sensor networks, with SensorBase, are centralized and no longer in the form of scattered ext files

### Slog Sensor Data

Similar to blogs, SensorBase provides an easy way to log sensor data, or rather, *slog*. We can think of sensor data as anything from numeric values to images, audio, and video. SensorBase allows you to slog these types of data both manually and automatically.

### Share Sensor Data

### Once data is slogged, you can share your data with collaborators or let all SensorBase users take a look. You can also elect to make your data private.

lise Sensor Data

Of course data isn't much good if you just set it aside to collect dust. SensorBase makes it easy to export data via the user interface or SOAP web services.

### Latest Blog Entry cing Notifications! (and some other fixes....)

U0-01-2009 10:3736 Hey all, So I know there hasn't been many updates in the last month, and that's mostly because I've been working on a big addition to Sensorbase, which as it gets expanded upon will be very useful. In fact, I think it's pretty useful as it is stands right now. Today, we intro ... Previous Post | Check out the Blog



Light level read by senso

Sensor battery voltage

0.42602

Sensor data

Number of photons in micromoles



eScholarship ersity of Califo

CINS

Mobile phone image feed

Paners

Home About Browse Publish Help My Items (0)

Search publications.

University of Calif

Center for Embedded Network Sensing

- Print Page

go

### **Publications**



# Case study: a CENS seismology deployment in Peru

- A joint UCLA-Caltech deployment in Southern Peru
- Study of scientific practices and artifacts
- Via ethnographic work and participant "insider" observation
- A narrative emerges











### What next

- I. Automatic and incremental generation of OAI-ORE aggregations via Annual Reporting system.
- 2. Construct a lightweight OAI-ORE platform to enable discoverability of newly created aggregations.

### Conclusion

- I. Contextual and raw data are as important to knowledge as are the published papers that report them.
- 2. Aggregation and linking mechanisms need to begin early in the lifecycle (work forward rather than backwards), i.e. capture data as cleanly as possible and as early as possible in the scientific life cycle.

## THANK YOU.

Alberto Pepe apepe@ucla.edu http://albertopepe.com http://twitter.com/albertopepe

This talk was adapted from:

From artifacts to aggregation: Modeling scientific lifecycles on the Semantic Web. Alberto Pepe, Matt Mayernik, Christine Borgman, Herbert Van de Sompel. Journal of the American Society for Information Science and Technology (JASIST). 61(3). 2010