Understanding what it means to be a repository online – unpacking the definition and requirements of a "digital repository"

Adam Brin

Digital Antiquity

What makes a repository?

tDAR

Preservation – ensuring a safe environment for the materials deposited

Conservation – monitoring and caring for the objects

Context – organizing and documenting the objects

So, what about a digital repository?

tDAR

Preservation – ensuring that the information entered and provided is stored, managed, and usable in 100 years…

Conservation – testing files regularly and converting them as necessary to be readable and accessible.

So, what about a digital repository?

tDAR

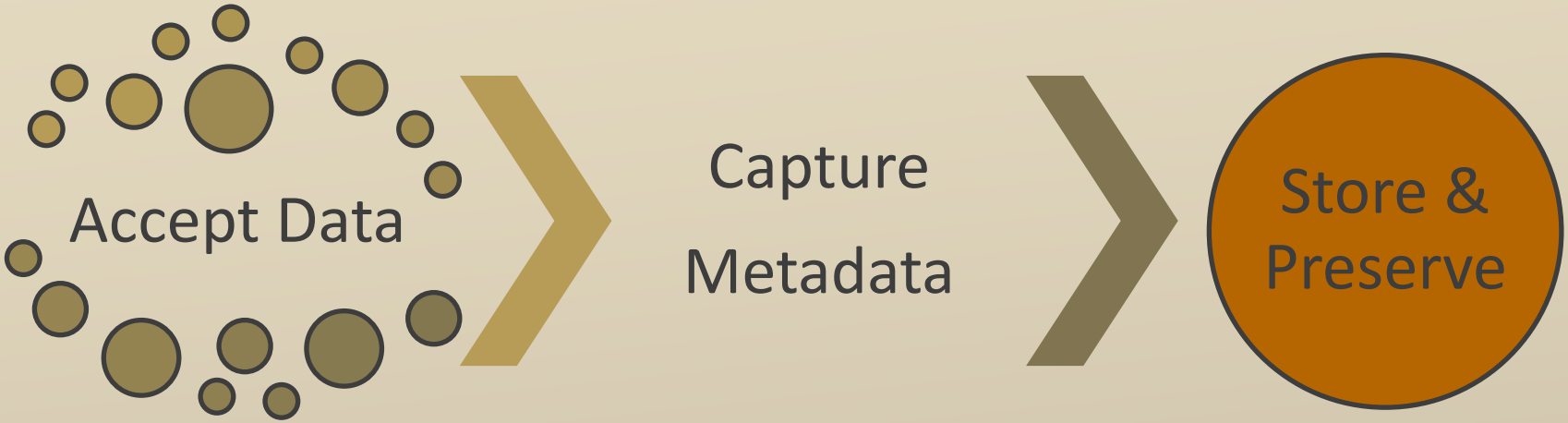
Context – Like physical collections, digital collections require metadata. What is the "minimum" information required to properly **read** and **use** a file?

Access – Repositories often work under two models: "dark" and "open."

Identity – All materials stored in the repository should be uniquely identifiable.

Basic Repository Workflow

tDAR



- Validate
- Verify
- Extract Metadata
- Convert

- Technical
- Administrative
- Descriptive

- Object In repository
- Index for discovery
- Present

Most of a repository is below the surface…

tDAR

Repository Software Components

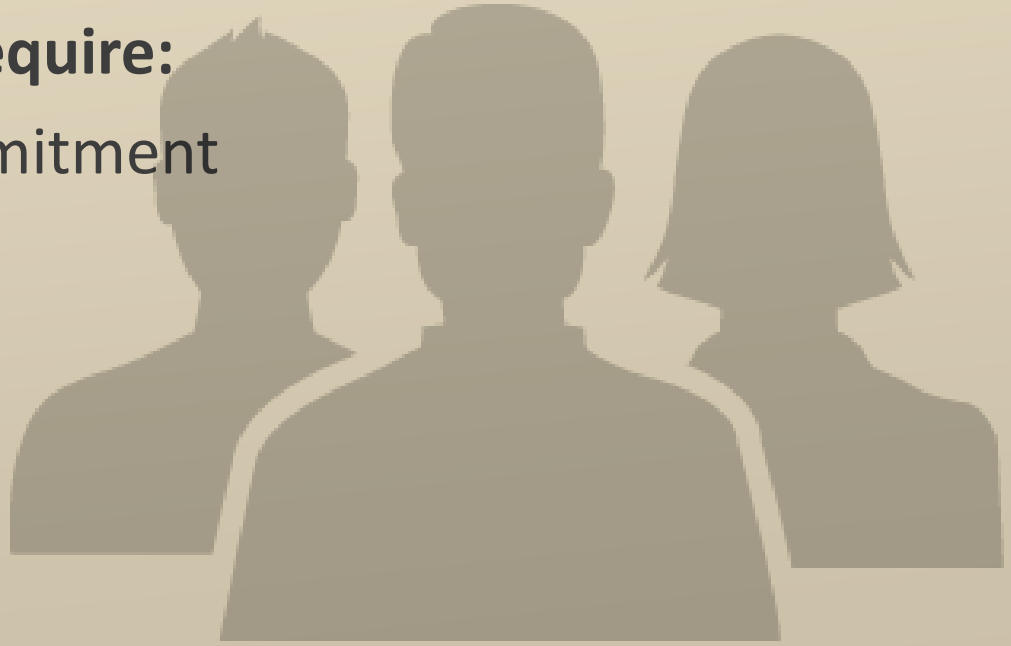
Preservation		Metadata			Access / Discovery		Curation
Fixity	Identity	Technical	Administrative	Descriptive	Browse	Search	Organize

But, it's not just software…

Probably the most important thing to understand about a digital repository is that it's not just software…

Digital Repositories require:

- Organizational commitment
- Monitoring
- Financial Support
- Long term Planning
- Infrastructure

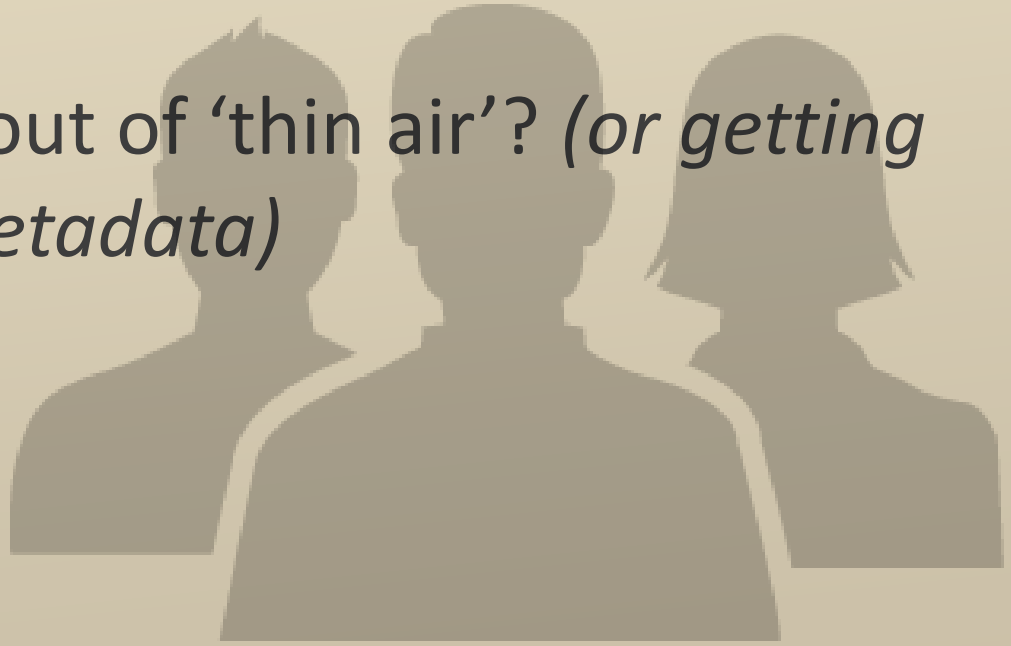


Technical Challenges…

- Developing and maintaining relationships between files
- Validating files
- Dealing with proprietary formats
- Migration
- Security

Social Challenges

- Integrating into the scholarly or professional workflow *(or getting users to contribute)*
- Making metadata out of ‘thin air’? *(or getting users to provide metadata)*
- Copyright



Why create tDAR?

Digital Data are not accessible—

- Standard archaeological work flows do not move digital data into trusted repositories.
- Few institutions provide digital curation (access and preservation).
- Media on which data reside are treated as artifacts.

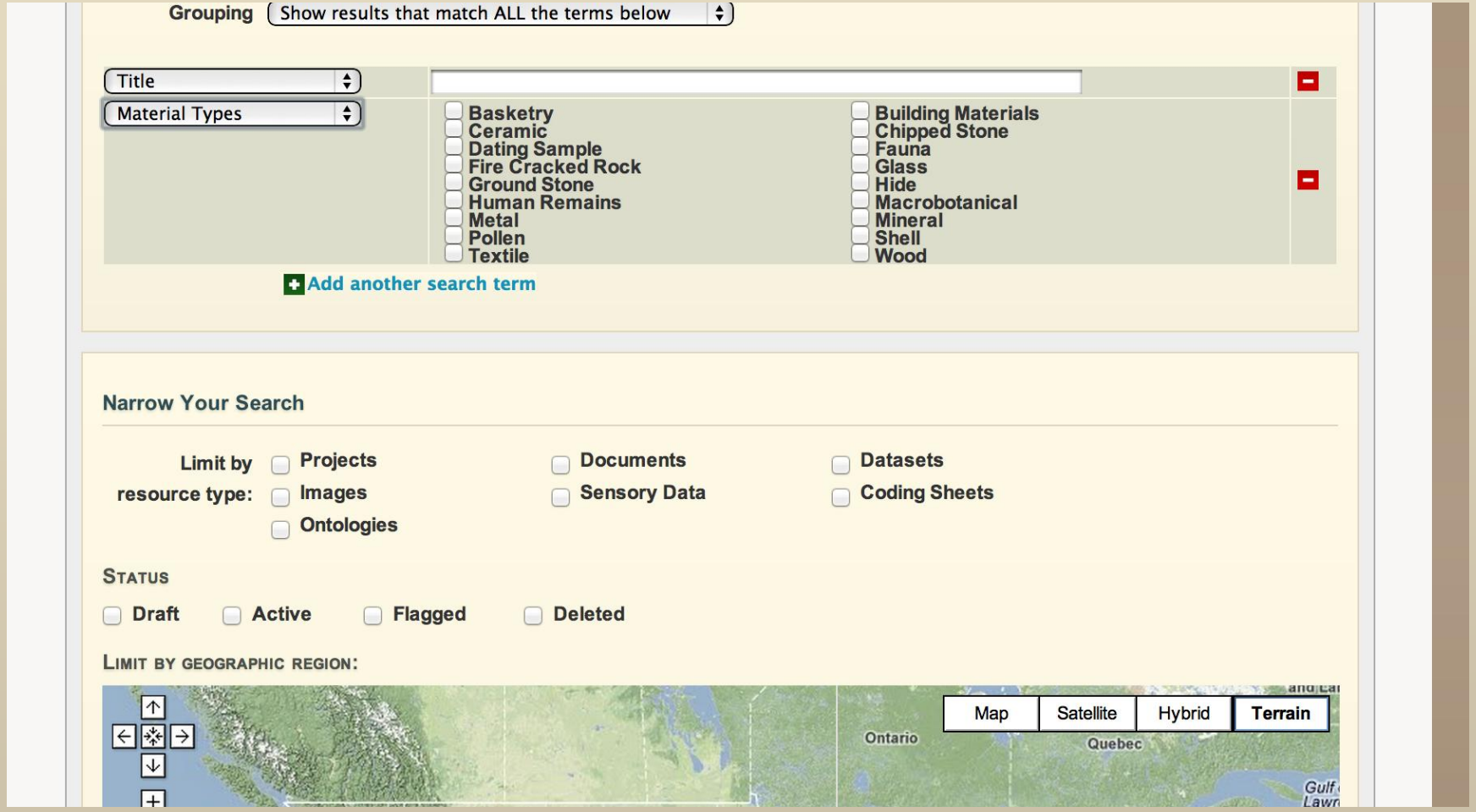
Digital Data are fragile and are not preserved—

- media degradation
- software obsolescence
- loss of information about the data (metadata)

What's Special About tDAR?

- Built around archaeological metadata for cataloging and discovery
- Protects site-security and limits access, if necessary
- Supports commonly used file types in our field
- Enables the integration of disparate datasets within the tool

Archaeological Search



Data Access

tDAR

tDAR the Digital Archaeological Record

A SERVICE OF DIGITAL ANTIQUITY

Home | Search | Login | Help

DATASET Yellow Jacket Pueblo faunal dataset

Part of the Crow Canyon Archaeological Center fauna through 2008 project

Robert Muir (Creator) Jonathan Driver (Creator) Karen Schollmeyer (Creator)

Faunal data from Yellow Jacket Pueblo, excavated by Crow Canyon Archaeological Center 1995 through 1997. Faunal analysis was by Robert J. Muir and Jonathan C. Driver in 1999, using methodology described in the CCAC Manual for the Description of Vertebrate Remains. Context information was integrated by Karen G. Schollmeyer in 2011 using data supplied by CCAC. Fauna is from AD 1060 to the late 1200s. Additional information on Yellow Jacket Pueblo and its fauna is available on the web at http://www.crowcanyon.org/ResearchReports/YellowJacket/Text/yjpw_faunalremains.asp.

Basic Information

Resource language: English
Year: 2011
tDAR ID: 6126
External Id: doi:10.6067/XCV8H41Q2T

Download(s)

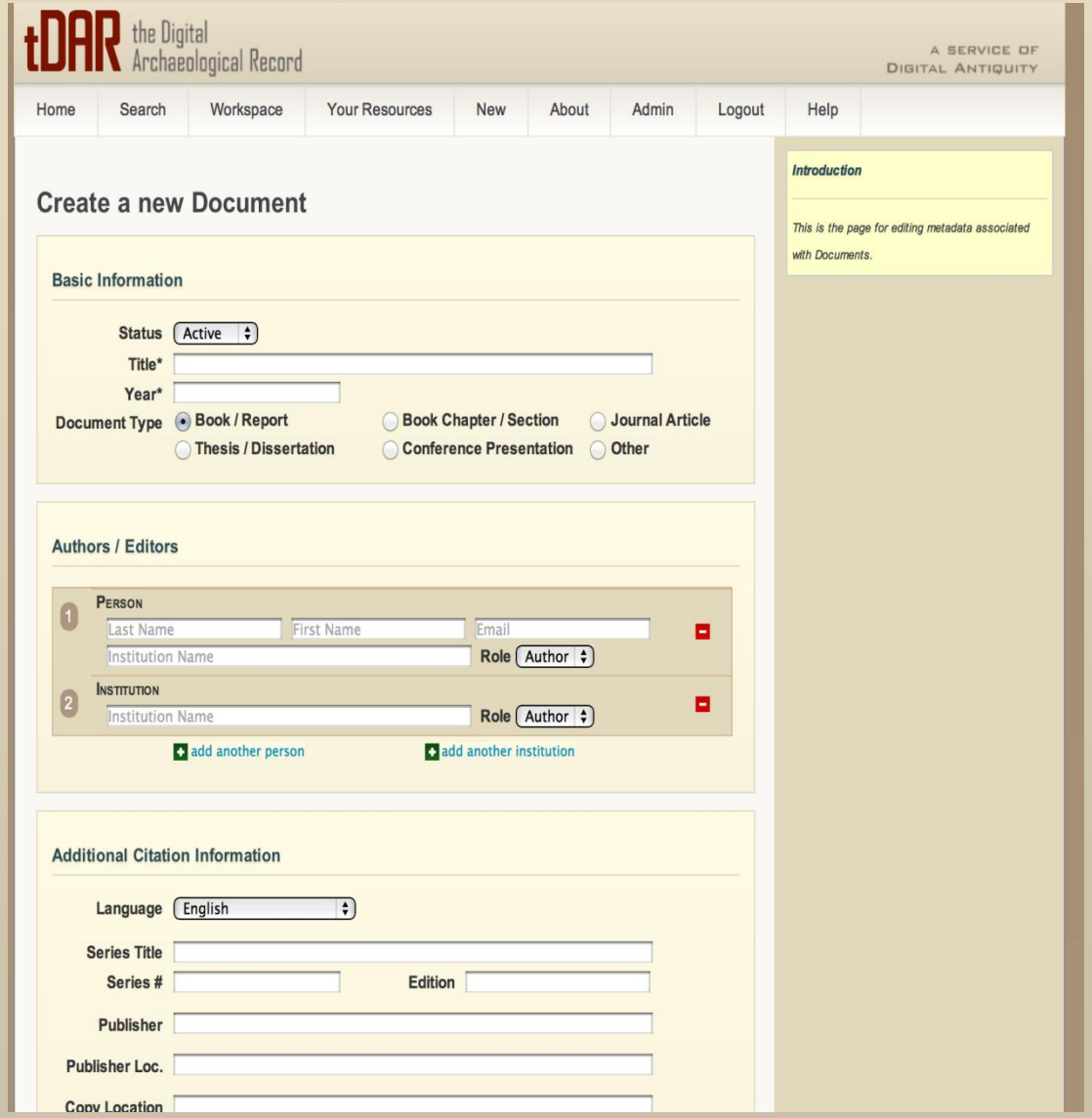
File #1 yellow-jacket-fauna-may2011.xlsx (1.16mb) [downloaded 20 times]

Keywords

Site Name: Yellow Jacket Pueblo
Site Type: Domestic Structure or Architectural Complex • Non-Domestic Structures
Culture: Ancestral Puebloan
Material: Fauna
Investigation Types: Data Recovery / Excavation

Spatial Coverage

Easy Capture & Entry



Documenting Datasets

YELLOW-JACKET-FAUNA-MAY2011 TABLE

Column Name	Data Type	Type	Category	Coding Sheet	Ontology
Site	VARCHAR	Uncoded Value	Provenience and Context : Horizontal Location	none	none
PD	BIGINT	Uncoded Value	Provenience and Context	none	none
FS	BIGINT	Uncoded Value	Provenience and Context	none	none
RecID	BIGINT	Uncoded Value	Fauna	none	none
Taxon	VARCHAR	Coded Value	Fauna : Taxon	**Taxon coding sheet for Crow Canyon Archaeological Center fauna through 2008**	**Outdated: Fauna Taxon - Southwest US Ontology (3028)**
Element	VARCHAR	Coded Value	Fauna : Element	**Element coding sheet for Crow Canyon Archaeological Center fauna through 2008**	**Fauna Element Ontology**
Part	BIGINT	Coded Value	Fauna : Breakage	**Part coding sheet for Crow Canyon Archaeological Center fauna through 2008**	none
Side	VARCHAR	Coded Value	Fauna : Side	**Side coding sheet for Crow Canyon Archaeological Center fauna through 2008**	none
ProxAntFusion	VARCHAR	Coded Value	Fauna : Fusion	**Fusion coding sheet for Crow Canyon Archaeological Center fauna through 2008**	none
DistPostFusion	VARCHAR	Coded Value	Fauna : Fusion	**Fusion coding sheet for Crow Canyon Archaeological Center fauna through 2008**	none
		Coded		**Breakage coding sheet for Crow Canyon**	

Integration of Datasets

tDAR

Create your Integration Table:

EACH COLUMN BELOW WILL BE A COLUMN IN EXCEL

☐ **Auto-select integratable columns** ☐ **Clear all columns**

Column 1: integration (Fauna Age Ontology)	Column 2: integration
T1. BAGE	T1. BSIDE
T2. Age	T2. Side

Add a new Column

Next: filter values

Select Variables

Display Variable **Integration Variable with mapped Ontology** **Measurement Variable** **Count Variable**

▾ 1: ALEXANDRIA FAUNA DATASET

SITENO - Alexandria / Spitalfields Faunal Date Ontology	**CN**	**FEATURE**	**Artifact Type**	**SHERDCT**	**MNV**
CATNO	**BCLASS**	**BCLASSID**	**BNAME**	**BNAMEID**	**Taxon** - Fauna Taxon Ontology - TAG (UK)
BELEMENT - Outdated: Fauna Element Ontology (5863)	**BELEMENTID**	**BAGE** - Fauna Age Ontology	**BAGEID**	**BSIDE** - Fauna Side Ontology	**BSIDEID**
BPORTION	**BPATHOL**	**BPATHOLID**	**BCONDITION**	**BCONDCUT**	**BCONDGNAW**
BCONDBURN	**BFUNCGRP**	**BDIET**	**BCUT**	**COMMENTS**	**REFERENCE**
LOCATION	**BOXNO**				

▾ 2: SPITALFIELDS FAUNAL MAIN INFORMATION DATASHEET WITH CONTEXTUAL DATA

Site code	**Site-context**	**Context**	**Basic INT EXP**	**Context Notes**	**Period**
Period text	**Site-bone**	**Bone ID**	**Species Common name**	**Bone Common name**	**Side** - Fauna Side Ontology
Part Common	**No**	**FUS PROX**	**FUS DIS**	**Prop**	**Age** - Fauna Age Ontology
Sex - Fauna Sex Ontology	**Carcas**	**No.teeth with tooth wear**	**No. measurements taken**	**No. Modifications recorded**	**Notes**

Reconcile Ontologies

tDAR

Filter Ontology Values

You can filter data values for the datasets listed below. Only checked values mapped to an ontology will be reported below. Select checkboxes next to the values that you would like to be included or aggregated to that level. Checkboxes are automatically checked if values are present in ALL datatables. Indented unchecked values are aggregated to the next higher level that is checked. Unchecked values at the top (leftmost) level are ignored, along with any unchecked subdivision categories. Values that occur in each dataset are indicated with blue checks, absent values are indicated with red x's.

Ontology labels from Fauna Age Ontology [BAGE, Age] (Select All | Select All With Shared Values | Clear All | Hide Unmapped)	bage (Alexandria fauna dataset)	age (Spitalfields faunal main information datasheet with contextual data)
adult	✓	✓
indeterminate	✗	✗
juvenile (all | clear)	✓	✓
fetal	✗	✓
other juvenile	✗	✓
unknown	✗	✗

Ontology labels from Fauna Side Ontology [Side, BSIDE] (Select All | Select All With Shared Values | Clear All | Hide Unmapped)	side (Spitalfields faunal main information datasheet with contextual data)	bside (Alexandria fauna dataset)
Axial OR Not Applicable	✗	✓
Both Left and Right	✓	✗
Indeterminate	✗	✗
Left	✓	✓
Not Recorded	✗	✗
Right	✓	✓

Thanks

NATIONAL ENDOWMENT FOR THE
Humanities

The Andrew W. Mellon Foundation

JISC

Attributions

- The Noun Project:
  - File Cabinet - **Alex Hartmann**
  - Safe - **Simon Child**
  - Community - **Rémy Médard**
  - Tag - **Ian Hamilton**
  - Microchip - **Rémy Médard**