



UiO : **University of Oslo**

Beginning with metadata: Case study RITMO

Kayla Burnim



The Research Council of Norway

RITMO Centre for Interdisciplinary Studies in Rhythm, Time and Motion

- Research Center of Excellence opened in fall 2018
- Organized "horizontally" within the University of Oslo, overlapping with the departments of:
 - Psychology (Faculty of Social Sciences)
 - Informatics (Faculty of Mathematics and Natural Sciences)
 - Musicology (Faculty of Humanities)
- Researchers are affiliated with one (or more) of the three departments
- Over 60 people employed at RITMO

RITMO at a glance

Rhythm is everywhere, from how we walk, talk, dance and play, to how we tell stories about the past and predict the future.

What are the basic structures and features of rhythmic phenomena?

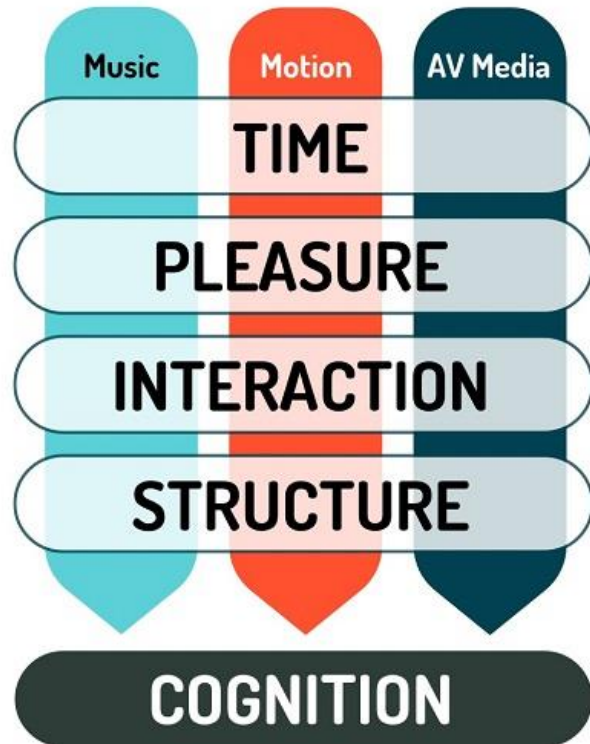
What are the mechanisms underlying human perception and cognition of rhythm?

The LINK!



Research Clusters RITMO

HUMAN RHYTHM



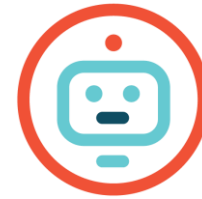
Structure & Time



Structure & Cognition

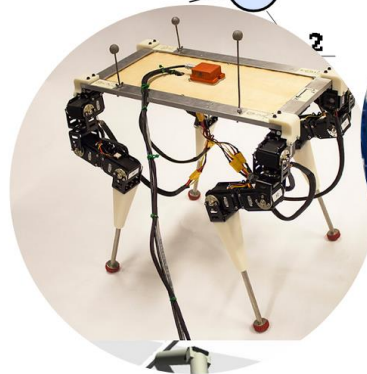
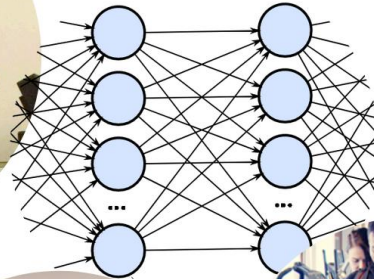
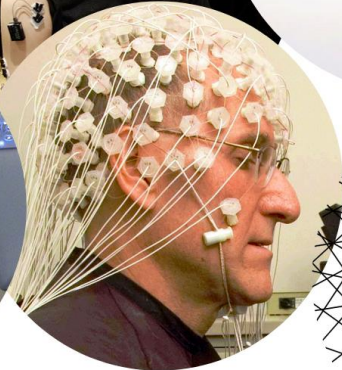
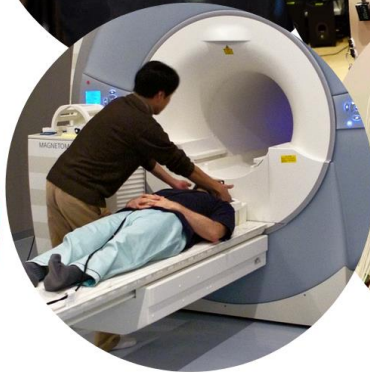
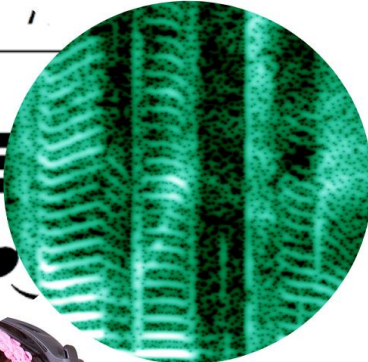
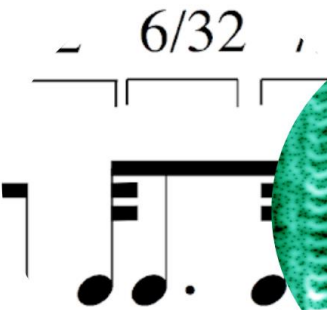


Interaction & Pleasure



Interaction & Robotics

RITMO Methods



Lab Structure at RITMO

RITMO's labs are run by two lab engineers, who also act as data managers.

FRONT Neurolab



fourMs Lab



Interaction (Robotics)



Data Management at the start of RITMO

Lab engineer started early 2019

GDPR implemented spring 2018

UiO website: general policies, guidelines, and information on research data management

New center -> no local procedures

What were researchers actually doing?



Data Management Status

Investigation into current state of data management



Researchers are from 3 different departments in 3 different faculty so prior to RITMO they had different:

- Research support – communication, training, etc
- IT departments – including e-infrastructure and storage
- Computer knowledge

Result of investigation:

- DM was unique to the individual
- Data is saved in many locations
- Most data was undocumented
- Researchers were not confident with data management

Starting Data Management

Two parallel paths:

- Increase data management knowledge and skills
- Create system to organize future data and clean up past data

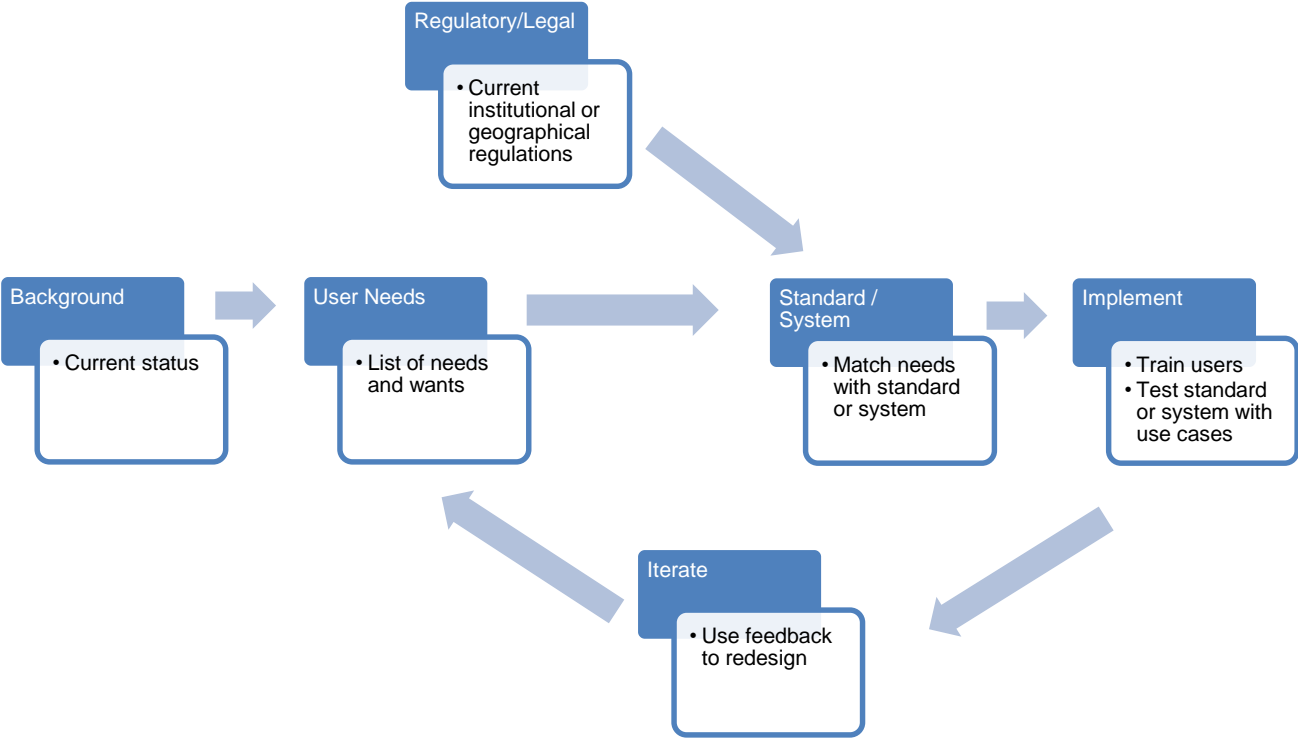


Find all the data and put it in the same place
Good time to add documentation

Next up Metadata!

Systemic Approach to Metadata

Metadata is not one size fits all



Using researcher needs to drive a metadata solution

User Needs with Metadata

How can “data about data” help RITMO?

Experiment level:

Must have

- What is this data?
- Who is responsible for it?

Nice to have

- Why was it made?
- Is there copyright?
- Was there NSD approval?

Data Level:

Must have

- What was done to the data

Nice to have

- Who processed it
- How it was processed
- When it was processed

#1 K.I.S.S.



Other Input for Standard / System

Institution updates:

- New initiatives like FAIR @UiO -> Metadata system in development

What is going on locally?

- Other lab group starting to use BIDS (Brain Imaging Data Structure) data standard developed for fMRI and adapting it for additional use with EEG data.
 - Input of DICOM to generate json metadata files and other DM
- Good for consistent methods, equipment, etc but not a good fit



Conclusion: Need “temporary” system

Choosing a standard

Institution System?

• Not yet

Is there a data type, subject, or disciplinary standard?

• Research too varied

Do any of the general standards fit?

• Yes!

Metadata
RDA | Metadata Directory

View the standards
View the metadata
View the tools
View the use cases
Browse by subject area

Contributor
Add standards
Add collections
Add tools
Add use cases

github
@rdalliance
rdalliance
facebook

Arts and Humanities
Encoded Archival Description (EAD) [G 6.1](#)
A standard for encoding archival finding aids using XML, in archival and manuscript repositories, implementing the recommendations of the International Council on Archives (ICA). [Category: International Standards Archives Description](#)

TEI - Text Encoding Initiative [G 6.1](#)
The guidelines of the Text Encoding Initiative make recommendations about suitable ways of representing those features of textual resources which need to be specified explicitly in order to facilitate processing by computer programs. They specify a set of XML tags in order to mark the textual resources, but structure, relationships between images and textual objects and other features of interest. The guidelines tools describe the format TEI encoding scheme, and document the TEI markup vocabulary. In their 30 years community driven development they have developed to a de facto standard in the production of textual data in the humanities.

DOI Data Documentation Initiative [G 6.1](#)
A widely used international standard for describing data from the social, behavioral, and economic sciences. Two versions of the standard are currently maintained in parallel:

- DOI Collections (or DOI version 2) is the simpler of the two, and intended for documenting simple survey data for exchange or archiving. Version 2.0 was released in August 2014.
- DOI Lifecycle (or DOI version 3) is richer and may be used to document datasets at each stage of their lifecycle from conception through to publication and reuse. Its modular and extensible. Version 3.0 was published in March 2014.

Both versions are XML-based and created using XML Schemas. They were developed and are maintained by the DOI Alliance.

MARC (Machine-Readable Cataloging) [G 6.1](#)
MARC is a standard and serialization format for representing bibliographic metadata, originally designed as a way of exchanging bibliographic records between library catalogs. Various other services have been defined, mostly with national or regional scope, of which MARC 21 is probably the most widely used. There also exists an XML serialization of MARC 21, known as MARCXML.

MOA3 (Museum Archives) [G 6.1](#)
A British cultural heritage standard for recording information on buildings, archaeological sites, shipwrecks, parks and gardens, suburbs, areas of interest and artifacts.

Sponsored by the Forum on Information Standards in Heritage. ISG/3 Version 1.1 was released in October 2012.

Open Archival Information System (OAIS) [G 6.1](#)
The goal of these standards is to expose the rich content in aggregations of Web resources to applications that support archiving, deposit, exchange, visualization, reuse, and preservation. The standards support the changing nature of

Because good research needs good data

About
News
Events
Services
Guidance
Briefing Papers
How-to Guides
Case Studies
Policy Analysis
Metadata
Disciplinary Metadata
Curator Lifecycle Model
Data Management Plans

Home • Guidance • Metadata • Disciplinary Metadata

List of Metadata Standards

ABCD - Access to Biological Collection Data
The Access to Biological Collections Data (ABCD) Schema is an evolving comprehensive standard for the access to and exchange of data about specimens and observations (i.e., primary biodiversity data). The ABCD Schema attempts to be comprehensive and highly structured, supporting data from a wide variety of databases. It is compatible with several existing data standards. Parallel structures exist so that either (or both) atomised data and free-text can be accommodated.

Sponsored by Biodiversity Information Standards TDWG - the Taxonomic Databases Working Group, the current specification was last modified in 2007.

AgMES - Agricultural Metadata Element Set
A semantic standard developed by the Food and Agriculture Organization (FAO) of the United Nations. AgMES enables description, resource discovery, interoperability and data exchange of different types of information resources in all areas relevant to food production, nutrition and rural development.

Sponsored by the UN/IMS - Agricultural Information Management Standards, the current standard was issued in November 2010.

AVM - Astronomy Visualization Metadata
The AVM scheme supports the cross-searching of collections of print-ready and screen-ready astronomical

Search all of Fairsharing Standards | Databases | Policies | Collections | Add/Claim Content | Stats | Log In or Register

FAIRsharing.org

Search Standards

Showing records 1 - 50 of 1326.

Registry	Name	Abbreviation	Type	Subject	Domains	Terminology	Related Databases	Related Standard	Related Policy	In Collection/Recommendation	Status
International Union of Pure and Applied Chemistry	IUPAC	IUPAC	Standard	Chemistry	Chemistry	Chemistry	Chemistry	None	None	None	11
Access to Biological Collections Data	ABCD	ABCD	Standard	Biodiversity	Biodiversity	Biodiversity	Biodiversity	None	None	None	11
Access to Biological Collections Data	ABCD	ABCD	Standard	Biodiversity	Biodiversity	Biodiversity	Biodiversity	None	None	None	11
Access to Biological Collections Data	ABCD	ABCD	Standard	Biodiversity	Biodiversity	Biodiversity	Biodiversity	None	None	None	11
Access to Biological Collections Data	ABCD	ABCD	Standard	Biodiversity	Biodiversity	Biodiversity	Biodiversity	None	None	None	11
Access to Biological Collections Data	ABCD	ABCD	Standard	Biodiversity	Biodiversity	Biodiversity	Biodiversity	None	None	None	11
Access to Biological Collections Data	ABCD	ABCD	Standard	Biodiversity	Biodiversity	Biodiversity	Biodiversity	None	None	None	11
Access to Biological Collections Data	ABCD	ABCD	Standard	Biodiversity	Biodiversity	Biodiversity	Biodiversity	None	None	None	11
Access to Biological Collections Data	ABCD	ABCD	Standard	Biodiversity	Biodiversity	Biodiversity	Biodiversity	None	None	None	11
Access to Biological Collections Data	ABCD	ABCD	Standard	Biodiversity	Biodiversity	Biodiversity	Biodiversity	None	None	None	11

- <https://rd-alliance.github.io/metadata-directory/standards/>
- <https://www.dcc.ac.uk/guidance/standards/metadata/list>
- <https://fairsharing.org/standards/>

Dublin Core matches both wants and needs

General Metadata Roll Out

Working it into concurrent data management work



Email with link
to Dublin Core
Website

- 1) Increase DM knowledge and skills
 - i. Local meetings, workshops, seminars
 - ii. Relay University seminars and resources

- 2) Create system and clean up past data
 - i. Create metadata templates within folder structure
 - Blank template added to all new or moved data
 - ii. Update handbook; summary info, useful links, etc
 - Including instructions with examples

Experiment Metadata - Current

We have a .txt based template with instructions added to new or moved experiment folders. Template and instructions are also available on our handbook.

UiO RITMO Centre for Interdisciplinary Studies in Rhythm, Time and Motion

Menu

← Research ← Groups ← fourMa Lab ← Handbook ← Data Management

Metadata

Archiving

Copyright

Data Management Plan

Ethics

File Naming

Metadata

Privacy

RITMO storage

Metadata is data on data.

Experiment Metadata

We are using a template based on the Dublin Core Metadata Initiative for project level metadata. DCMI is a widely used, simple, and flexible standard consisting of 15 elements. These elements might seem vague or confusing so examples have been included as well as the link to the DCMI element descriptions. The simple file will answer what the project is and who is responsible for it.

1. TITLE (Your Project Name)
 - Element Description: The name given to the resource
 - Additional Guidelines:
 - Examples: Lion King, Bohemian Rhapsody, MusicLab 4, Dance Dance Revolution (DDR)
2. CREATOR (Project Lead)

DC Project Metadata Template.txt - Notepad

File Edit Format View Help

```
<<For instructions on the use of this template see the handbook https://www.uio.no/ritmo/eng
<<Template last updated 26Aug2020>>
```

```
1. TITLE (Your Project Name)
The name given to the resource
2. CREATOR (Project Lead)
The person primarily responsible for the intellectual content of the resource; the author.
3. SUBJECT (Main Topic)
The topic of the resource written in keyword format
4. DESCRIPTION
A textual description of the content of the resource
5. PUBLISHER (Probably UiO)
The entity responsible for making the resource available
6. CONTRIBUTORS
Additional person(s) or organization(s) responsible for making contributions to the project
7. DATE (Beginning of project)
The date associated with the beginning of the project.
8. TYPE
The nature or genre of the content of the resource.
9. FORMAT
The physical or digital format of the resource
10. IDENTIFIER (Optional)
A string or number used to uniquely identify the resource.
11. SOURCE (Optional)
The reference from which the resource is delivered (if applicable).
12. LANGUAGE
The language(s) of the intellectual content of the resource.
13. RELATION (Links to other projects)
A reference to a related resource
14. COVERAGE
The spatial and/or temporal scope of the resource
15. RIGHTS MANAGEMENT (If applicable)
Information about rights held in and over the resource.
```

Data Level Metadata - Background

Why this is needed?

- Most of the software used to process our data does not have a “track changes” style log
- Turnover with researchers, can't always just ask
- Reduce repeat processing stage work

Are there standards for this journal document?

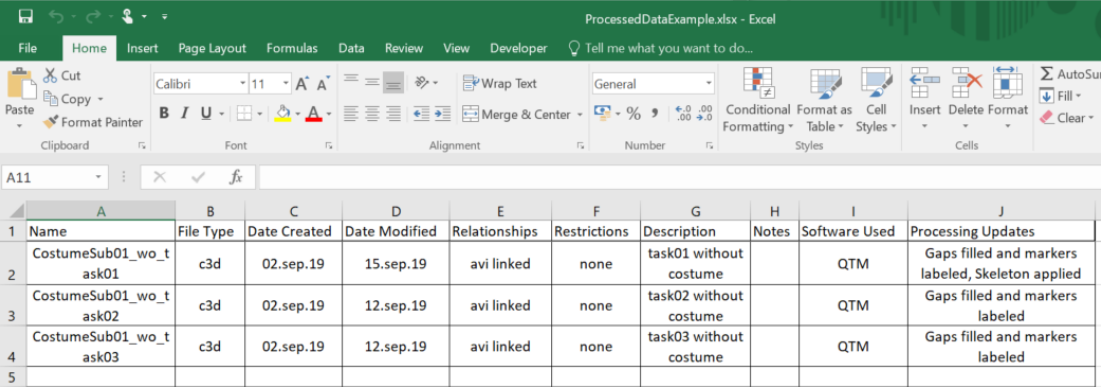
Electronic Lab Notebook, not super simple

Ok, so now what?

Make something up and see how it works!

Data Level Metadata – First Try

First iteration of “processing journal” was .csv template



The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D	E	F	G	H	I	J
1	Name	File Type	Date Created	Date Modified	Relationships	Restrictions	Description	Notes	Software Used	Processing Updates
2	CostumeSub01_wo_t ask01	c3d	02.sep.19	15.sep.19	avi linked	none	task01 without costume		QTM	Gaps filled and markers labeled, Skeleton applied
3	CostumeSub01_wo_t ask02	c3d	02.sep.19	12.sep.19	avi linked	none	task02 without costume		QTM	Gaps filled and markers labeled
4	CostumeSub01_wo_t ask03	c3d	02.sep.19	12.sep.19	avi linked	none	task03 without costume		QTM	Gaps filled and markers labeled
5										

Why we tried something else:

- Spreadsheet was cumbersome and took too long to fill out
- Matrix style was not the right fit
- Template did not capture what we really wanted



Data Level Metadata – Current Template

```
ProcessDataJournal.txt - Notepad
File Edit Format View Help
Processing Journal for Experiment XX

---Original Files---
Created by:
Date(s) Created:

Restrictions:

Notes:

---Processing Updates----
Files(s):
Edited by:
Date:
Software Used:
What was done:
```

Current iteration is .txt
processing journal

Why this is working better:

It allows for simple and easy updates whenever the data is altered noting what was done, by whom, when it was done, and how it was done.



What now?

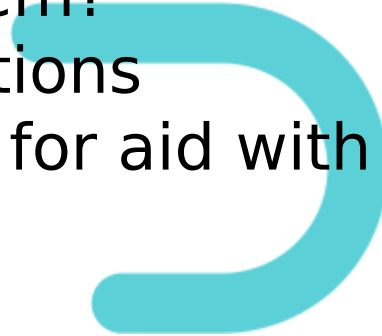
Currently in the “Implement” phase

Active assessment

- How are people using the system?
- Hands on office hours for questions
- Scheduled individual meetings for aid with bigger projects

Prepare for University system

- Machine readable system coming
- Tags and keywords



Breakout Room

Systemic approach more in depth
Interactive component to follow along
More in depth RITMO examples

