



Research workflows and digital data sets:

traceability, reproducibility, comparability issues illustrated on spatio-acoustic data acquisition protocols

Iwona DUDEK

UPR 2002 CNRS MAP



FR-IT Bilateral Cooperation in Heritage Science – 7th Edition

DIVING INTO DIGITAL DATA FOR HERITAGE SCIENCE: DIAGNOSTICS, UNDERWATER HERITAGE, AND SOUND
Marseille, Centre National de la Recherche Scientifique (CNRS), Campus Joseph Aiguier
February 7th, 2024





Workflow descriptions imply choice of a language

“context-free” languages

the meaning is highly context-dependent

“HARD”
LANGUAGES

“SOFT”
LANGUAGES



A. KOBZYCKI, *The role of the language in a perceptual process*, New York 1951

S. LEM, *Tajemnica chińskiego pokoju*, Kraków 1996





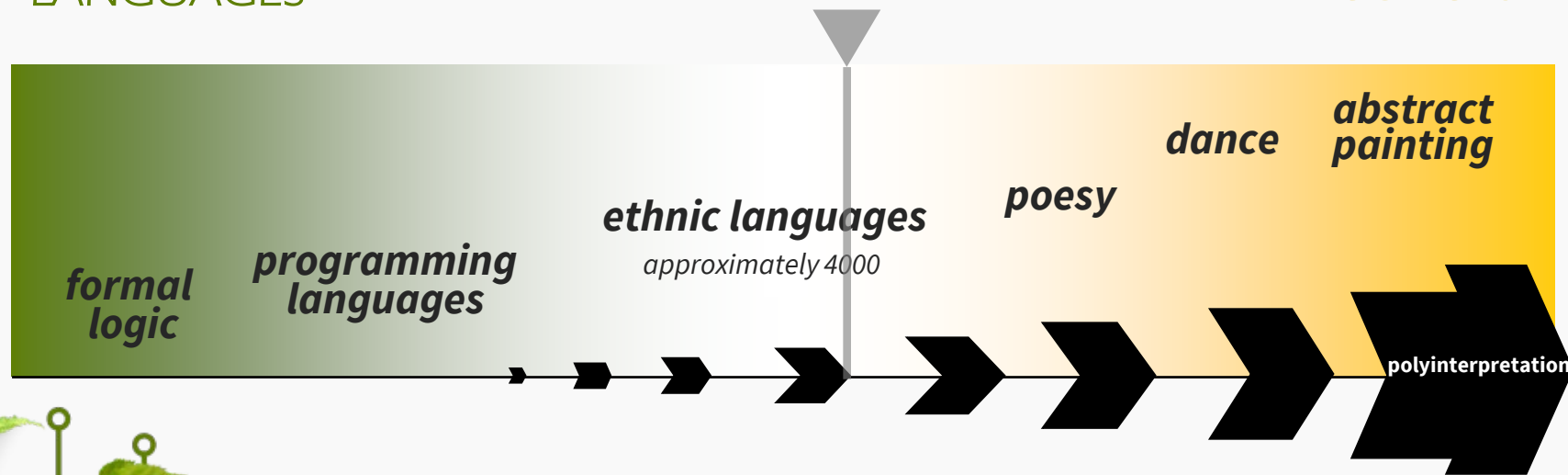
Workflow descriptions imply choice of a language

“context-free” languages

the meaning is highly context-dependent

“HARD”
LANGUAGES

“SOFT”
LANGUAGES



A. KOBZYCKI, *The role of the language in a perceptual process*, New York 1951

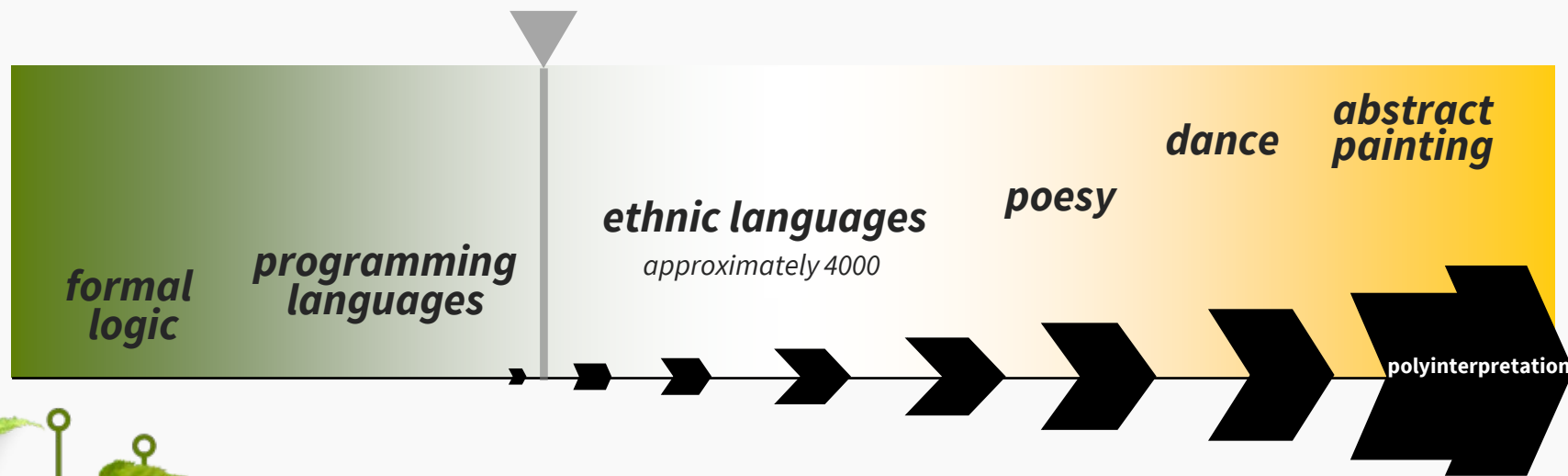
S. LEM, *Tajemnica chińskiego pokoju*, Kraków 1996





Workflow descriptions imply choice of a language

conceptual modelling
controlled vocabulary
diagrammatic representations based on drawing conventions



A schematic representation explains parts, operations and relationships between elements.



Structure of the presentation

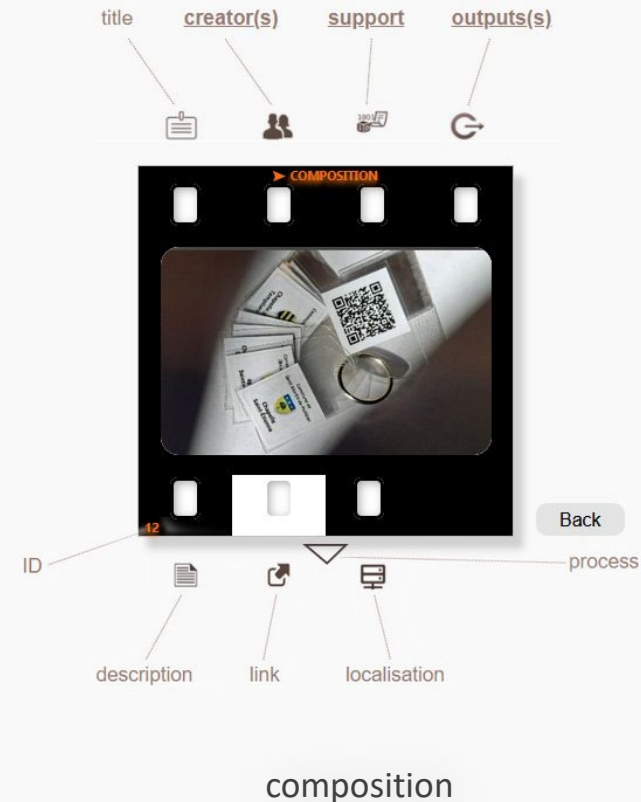
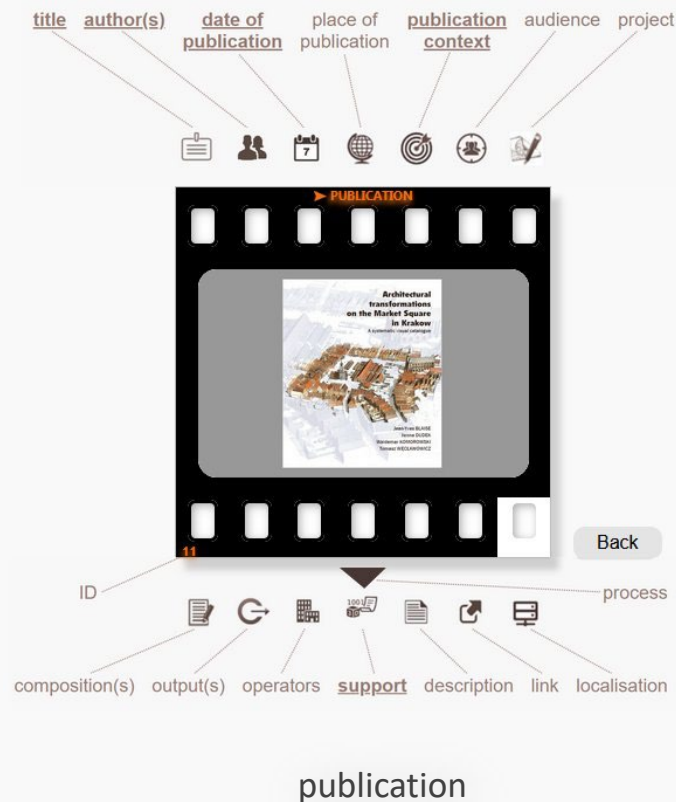
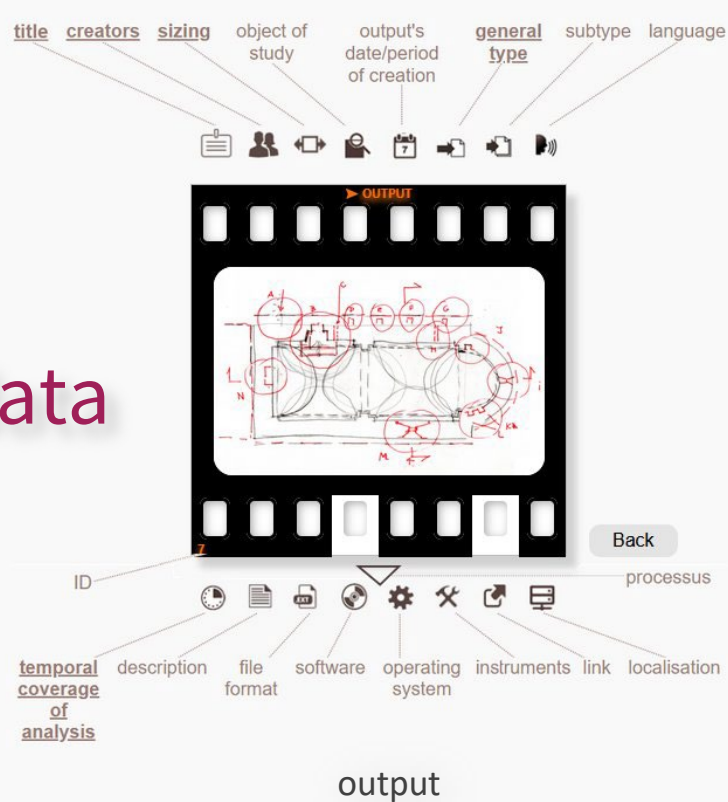
- brief outline of the Memoria IS objectives
- clarification of the underlying concepts
- illustration of these concepts by means of examples
- discussion about advantages and limits of the system



main objectives

allow for the description of research results and their production history

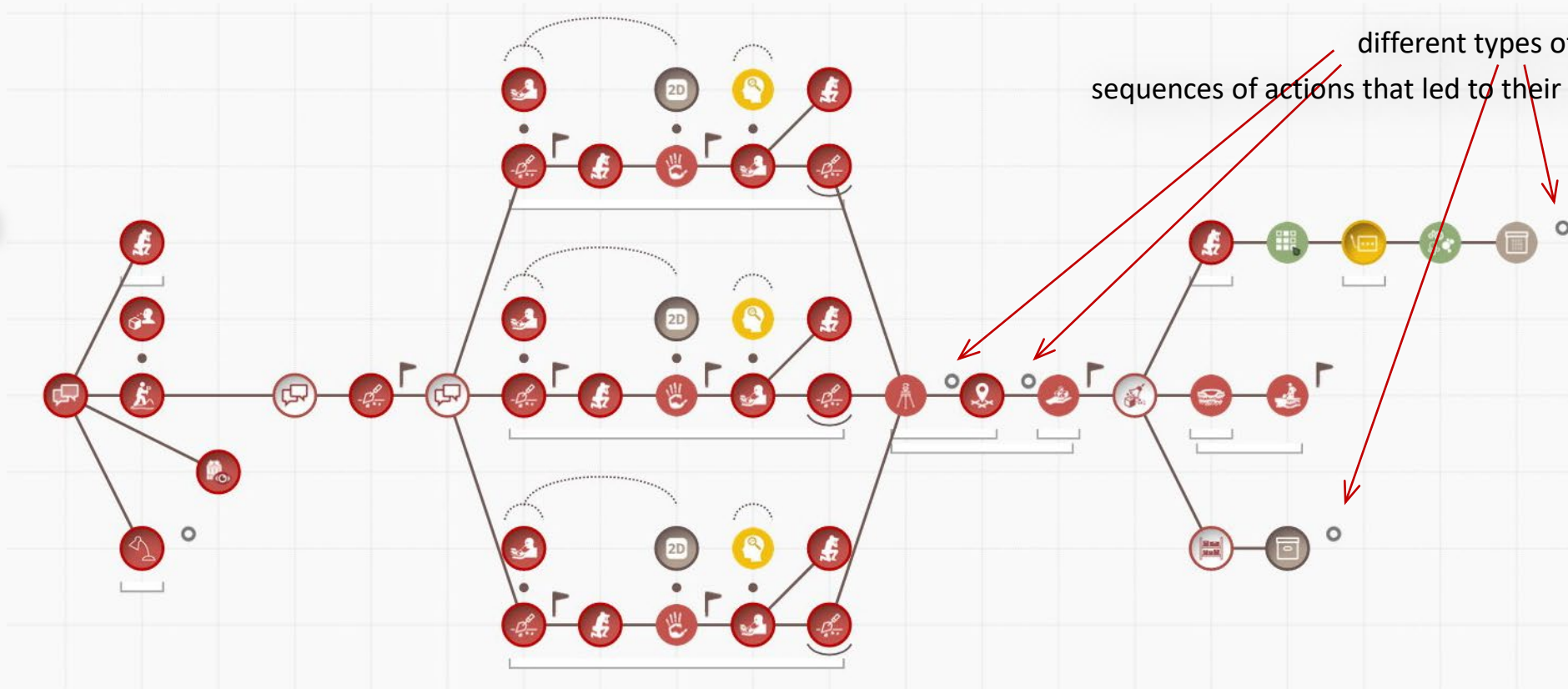
metadata



main objectives

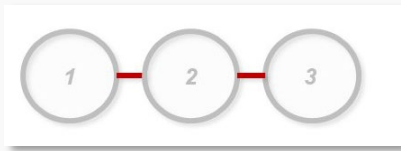
allow for the description of research results and their production history

paradata

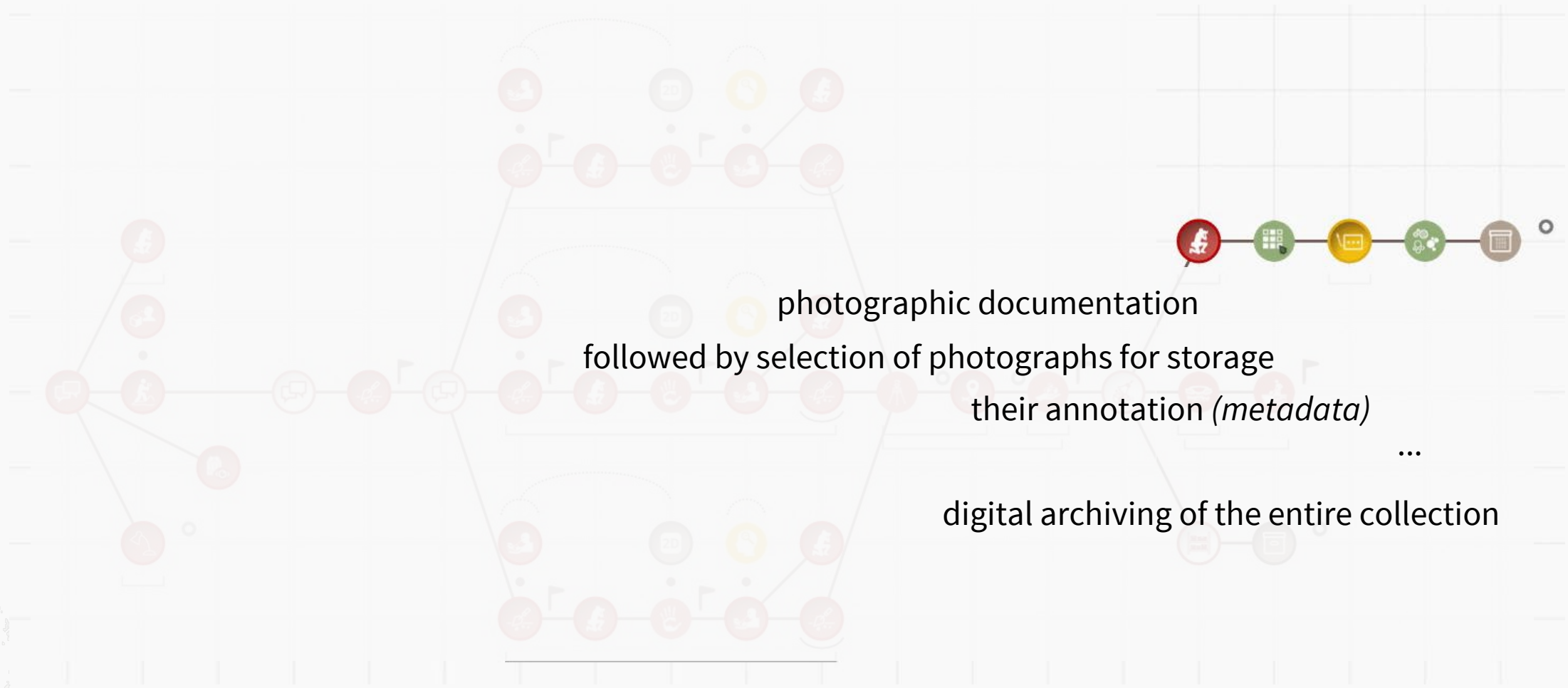




underlying concepts



chain - a succession of activities over time (displaying their order)



photographic documentation

followed by selection of photographs for storage

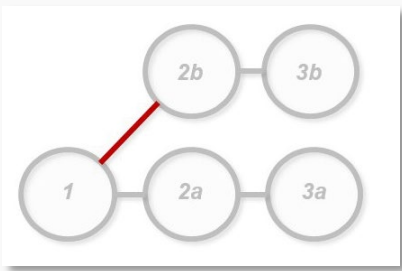
their annotation (*metadata*)

...

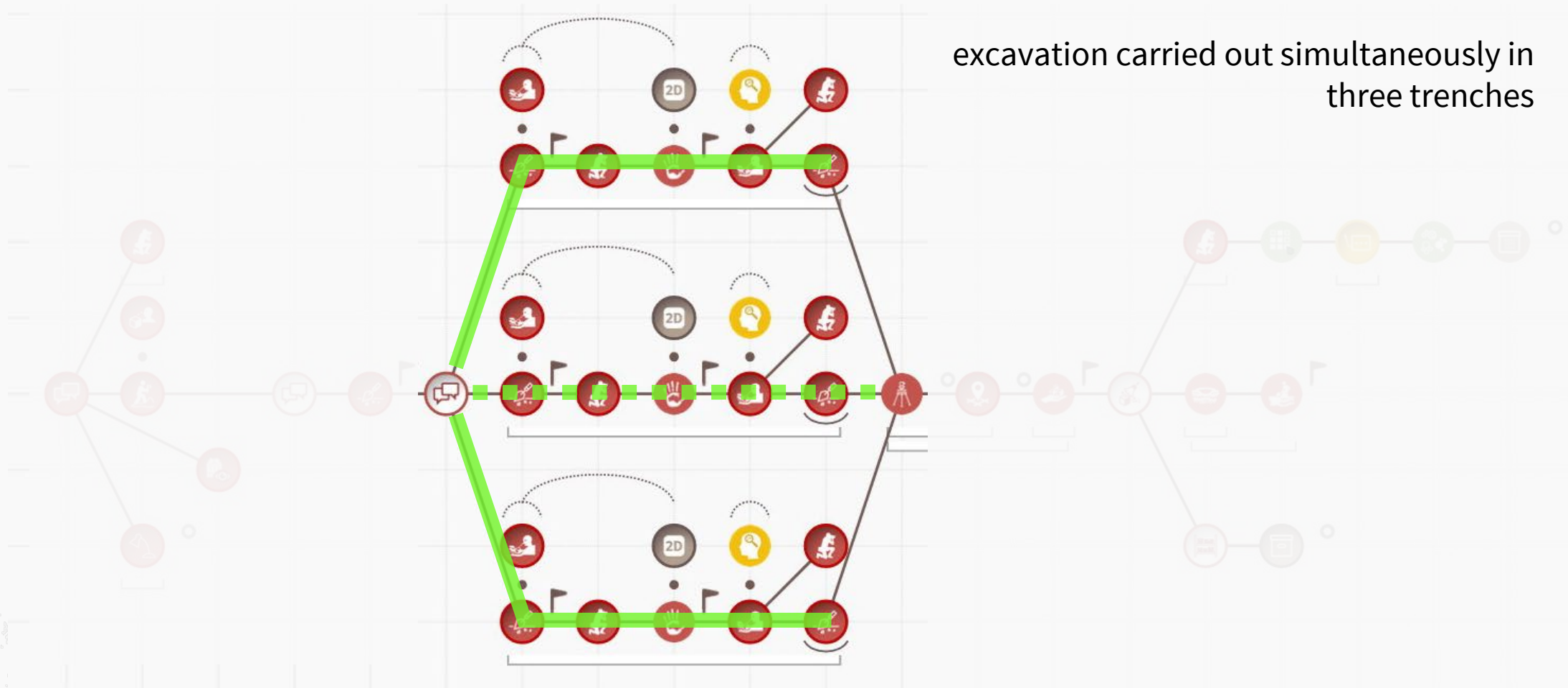
digital archiving of the entire collection



underlying concepts



parallel (sequences of) activities - activities that run concurrently in time

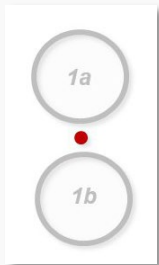
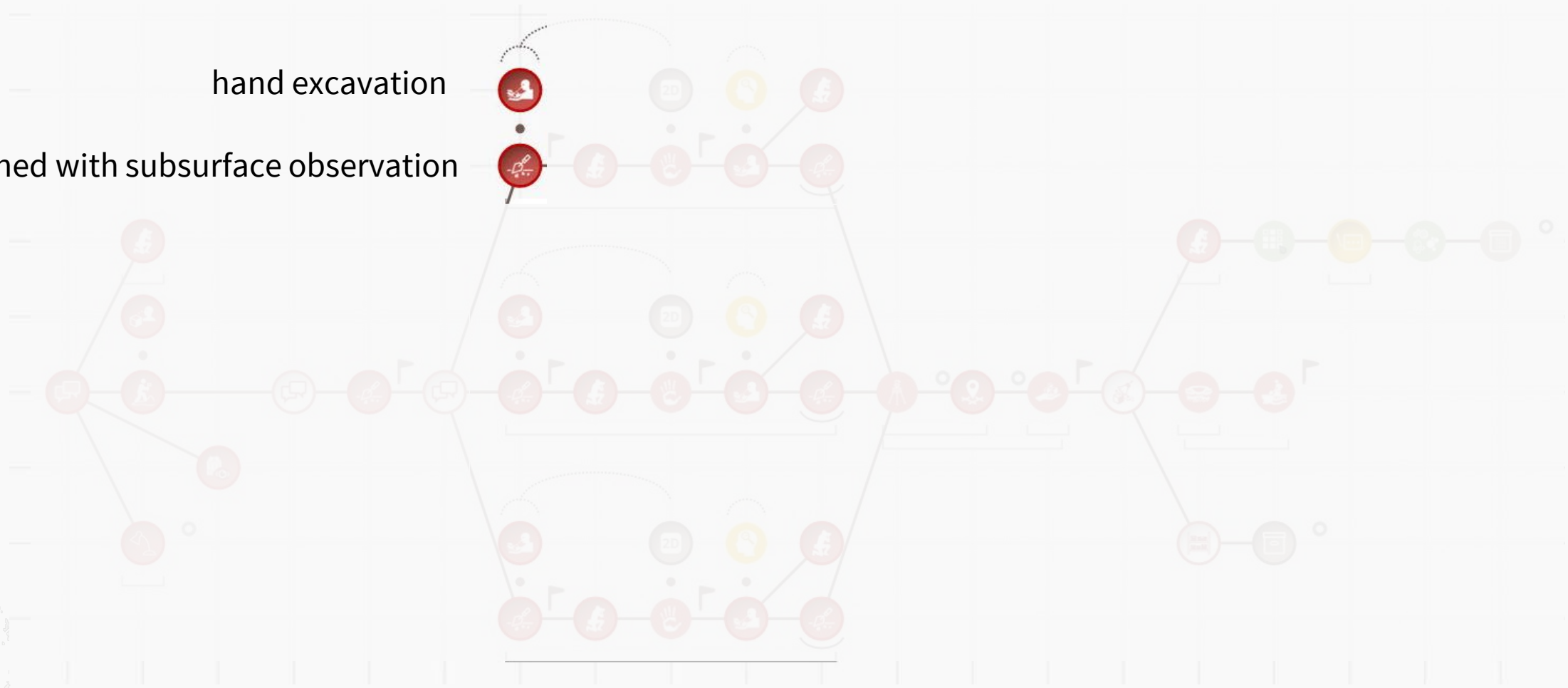


excavation carried out simultaneously in three trenches



underlying concepts

hand excavation
combined with subsurface observation

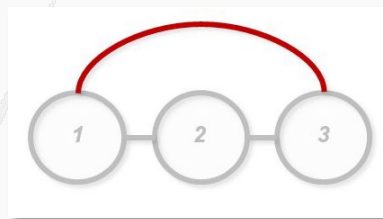


knot of activities - activities that take place simultaneously or occur without any apparent sequential order

hand excavation
combined with subsurface observation

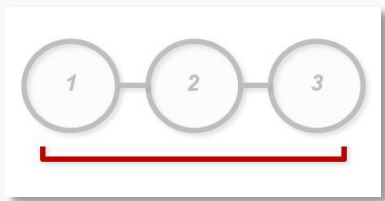
followed by photographic documentation
and inventory of stratigraphic units
(measurements and documentation)

The whole sequence is repeated
until an artefact is discovered

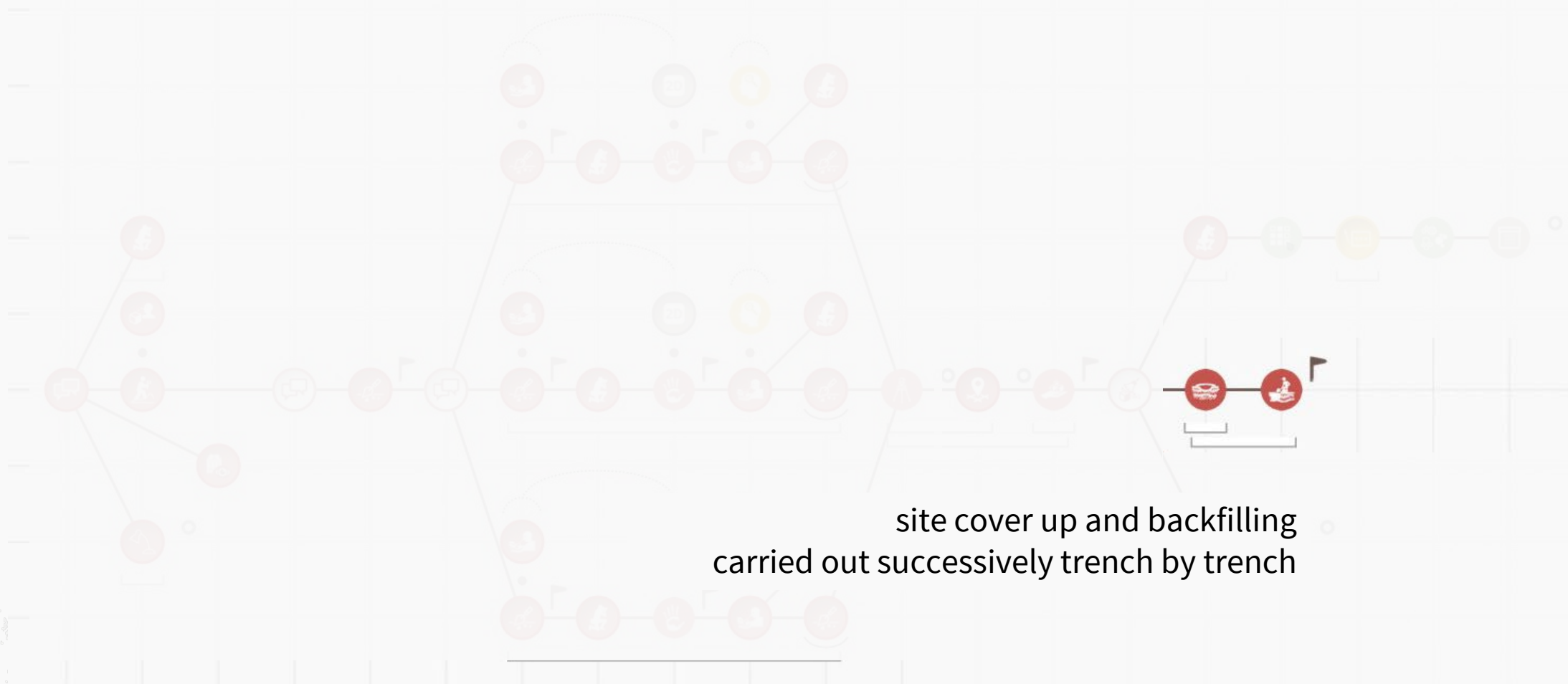


iterative activities - sequence of operations is **repeated on the same object**
until the desired result is obtained

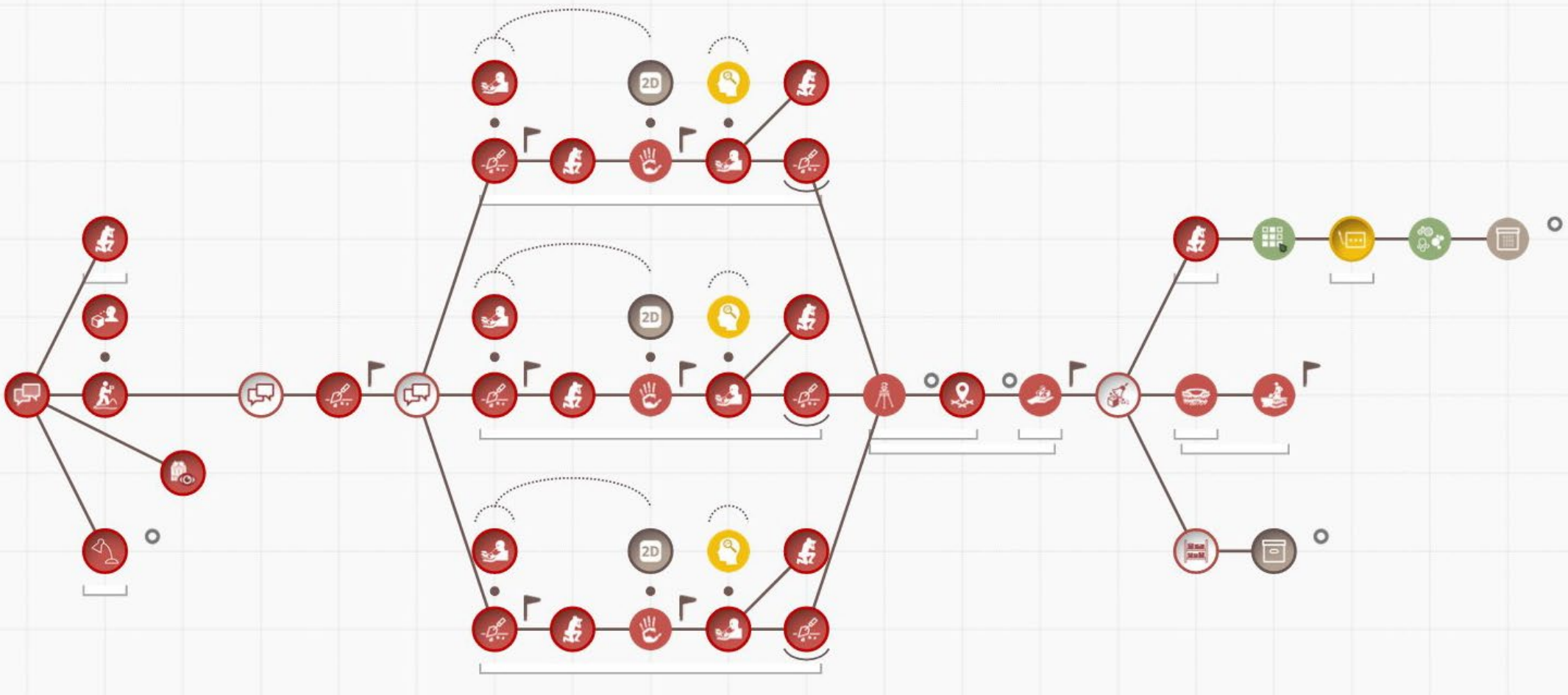
underlying concepts



repetitive activities - repetition of the same operations **on multiple objects**
(done once for each item)



underlying concepts



● data collection and acquisition
● data filtering and treatment

● data analysis
● added value procedural activities

● finalisation



Organising activities in a workflow diagram > drag-and-drop web interface



MEMORIA
MEMOrisation de Ressources numériques et d'Activités

Show > [Icons: +, 🔍, ✎, 👤, 🗣️]

[Icons: 📄, 📁, ↻, 🔄, 📋, ⓘ, 🗨️, 👤 idu, 🏠]

spatially constrained setup

Arranging instruments and operations with regards to specific spatial constraints, typically with regards to a specific reference frame used in the acquisition phase.
e.g., Setting up a grid used as a frame to position instruments.

Cancel

MEMORIA Information System > add mode > add new process

UMR 3495 CNRS/MCC MAP | Modified on : 04/01/21 | Legal notice



Description of individual activities

Once an activity is characterised, a background colour is added to its icon.



MEMORIA
MEMORISATION de Ressources numériques et d'Activités

Show > + 🔍 ✎ 🚶 (1)

inference (?)

- structure
 - simple reasoning
 - compound reasoning
- purpose
 - justifying reasoning
 - revealing reasoning
- causal direction
 - progressive reasoning
 - regressive reasoning
- reliable reasoning
 - deductive reasoning
 - induction by complete enumeration
 - eliminative induction
- cognitive skills
 - numerical reasoning
 - visual reasoning
 - verbal reasoning
 - spatial reasoning
- presumptive reasoning
 - induction by partial enumeration
 - reasoning by analogy
 - probabilistic reasoning
 - statistical reasoning
 - comparative reasoning
- type of premises
 - causal
 - structural
 - normative
 - semantic
 - empirical
 - logical

activity instance >

- activity >
- instance >
- history of Carthage: Carthage
- expedite >
- activity >
- collection of access data: Visual reconstruction hypothesis
- software >
- Adobe Illustrator
- operating system >
- Windows 97

Specify the start date (the beginning) of the activity: 06/03/2018

Specify the closing date of the activity: 02/04/2017

Specify a time unit that will be used to describe the duration of the activity: day | week | month | year

Specify the duration of the activity using the selected time unit: 5

Submit Cancel

Cancel Accept

MEMORIA Information System > add mode > add new process

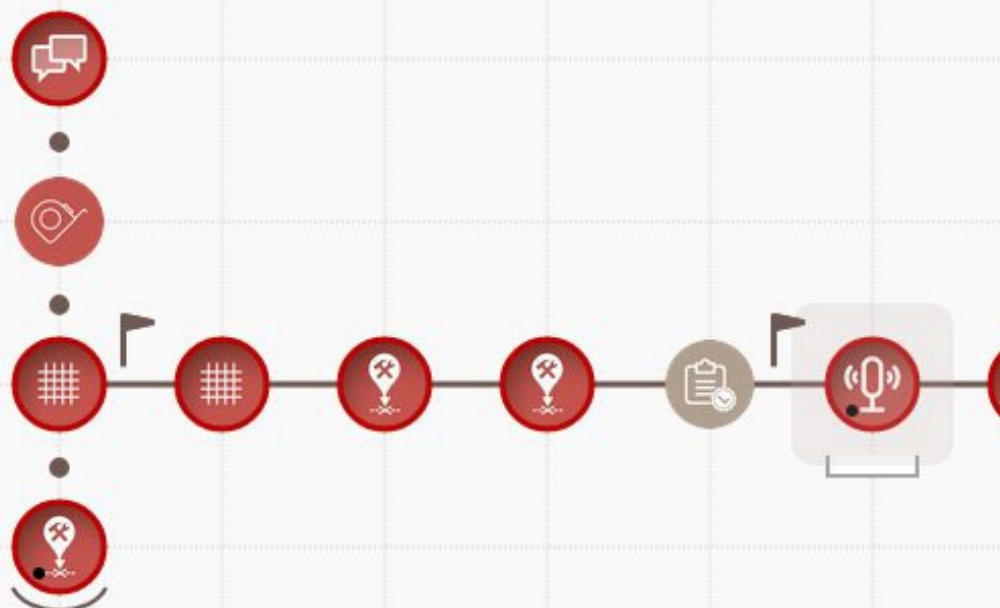
UMR 3495 CNRS/MCC MAP | Modified on : 03/11/20 | Legal notice



Paradata depicting each activity is stored within in the workflow structure
freely accessible, searchable, editable



116 - acquisition de données acoustiques - salle pleine, séminaire SESAMES (salle Jean Salusse)



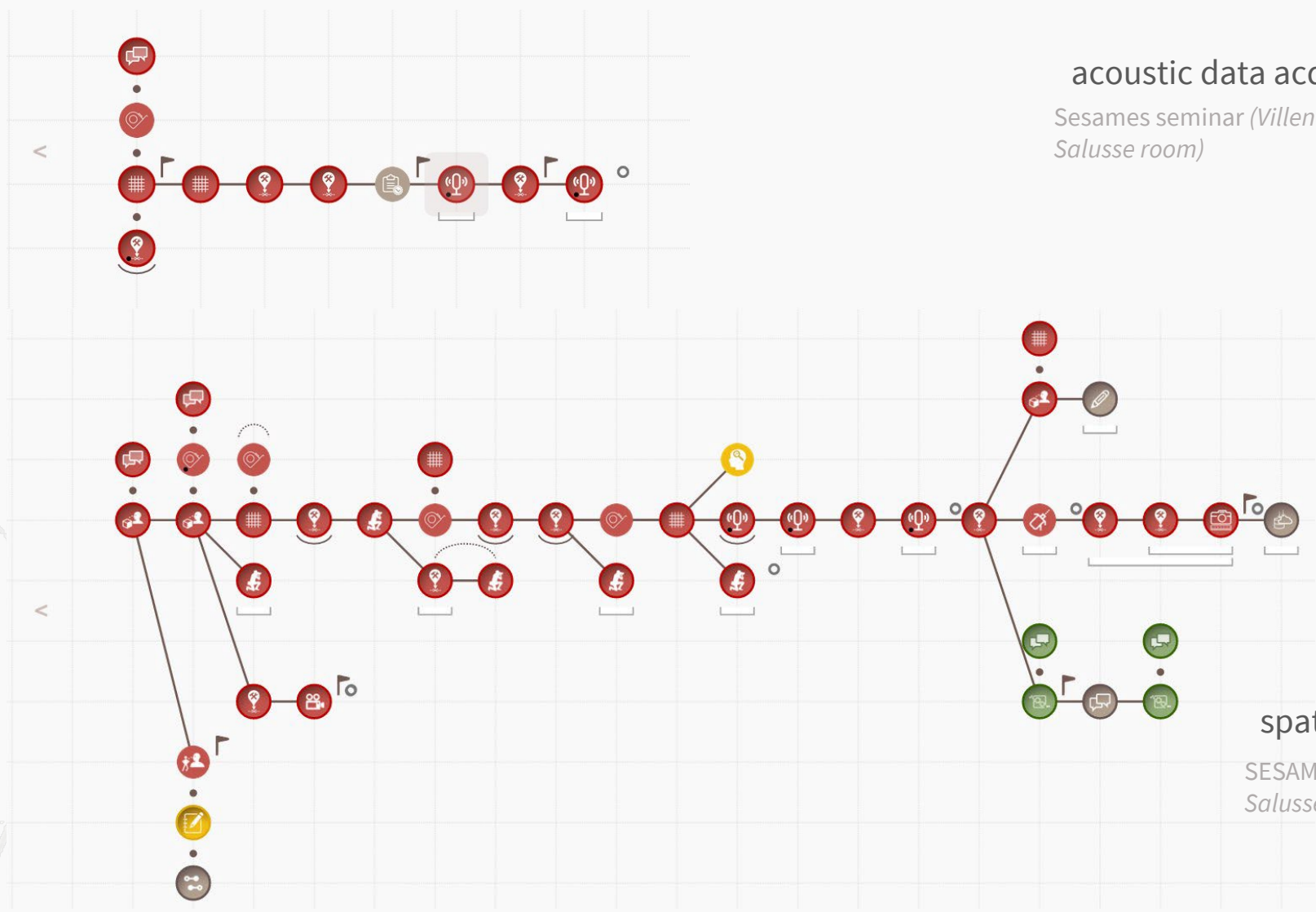
activity instance >

activity group >	data collection and acquisition
activity type >	audio recording
acoustic source >	preregistered sound
place of registration >	other indoor spaces
individuals >	Fargeot Simon > AMU Vidal Adrien > PRISM
collective bodies >	
sources >	
inputs >	sweep acquisition Sesames
compositions >	
publications >	
expertise >	
outputs >	
instrument(s) >	em32 Eigenmike (microphone) Genelec 8020C (speaker) MacBook Pro (mi-2014)- PRISM
software >	REAPER
operating system >	Mac OS X 10.14 (Mojave)
start date >	23/06/2022
end date >	23/06/2022
duration unit >	minute
remarks >	mesures acoustiques enregistrement du sweep 3x, 25 personnes dans la salle





Spatio-acoustic data acquisition



acoustic data acquisition

Sesames seminar (Villeneuve lez Avignon, Jean Salusse room)

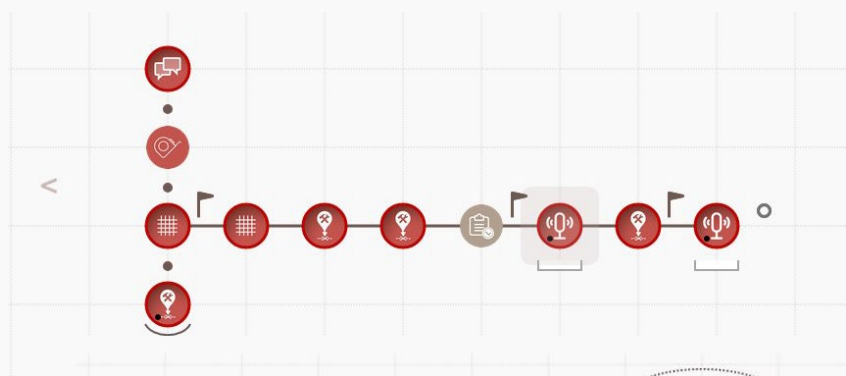
spatio-acoustic data acquisition

SESAMES project (Villeneuve lez Avignon, Jean Salusse room)



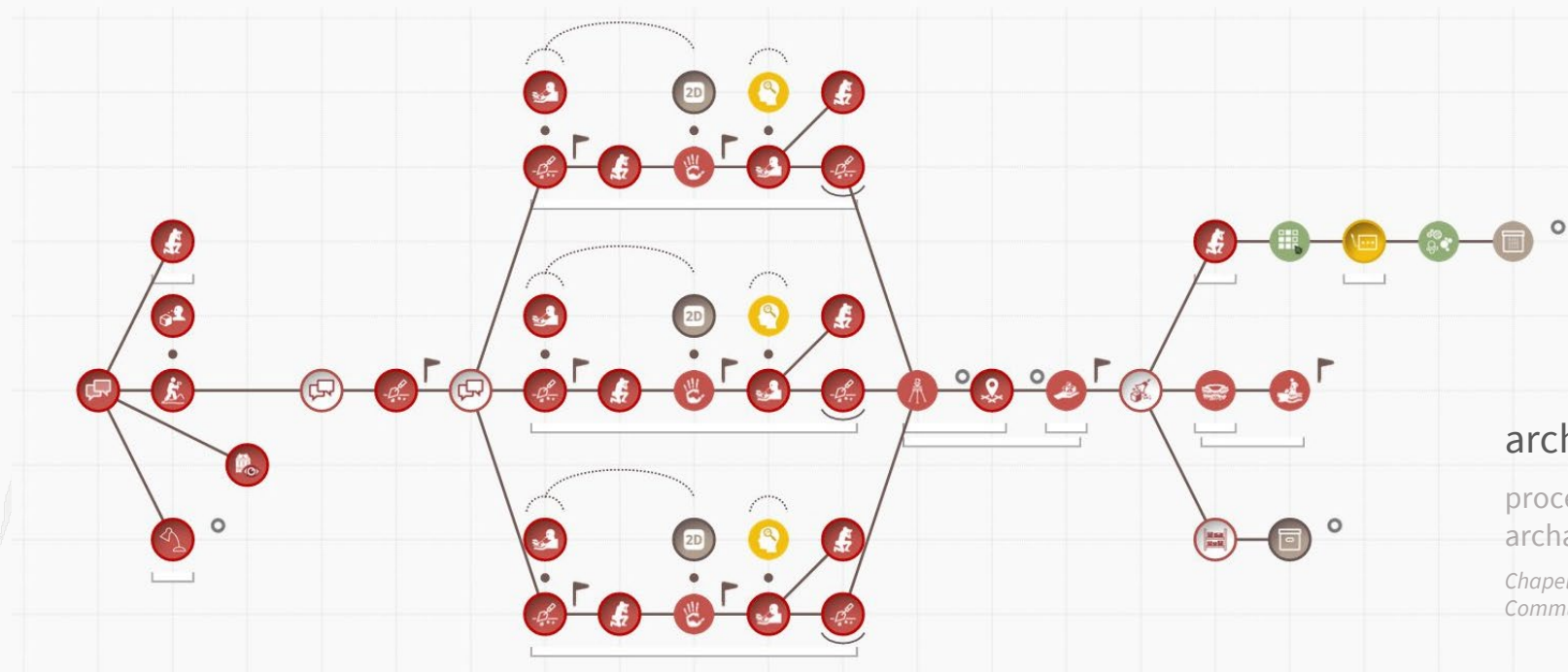


Spatio-acoustic data acquisition



acoustic data acquisition

Sesames seminar (Villeneuve lez Avignon, Jean Salusse room)



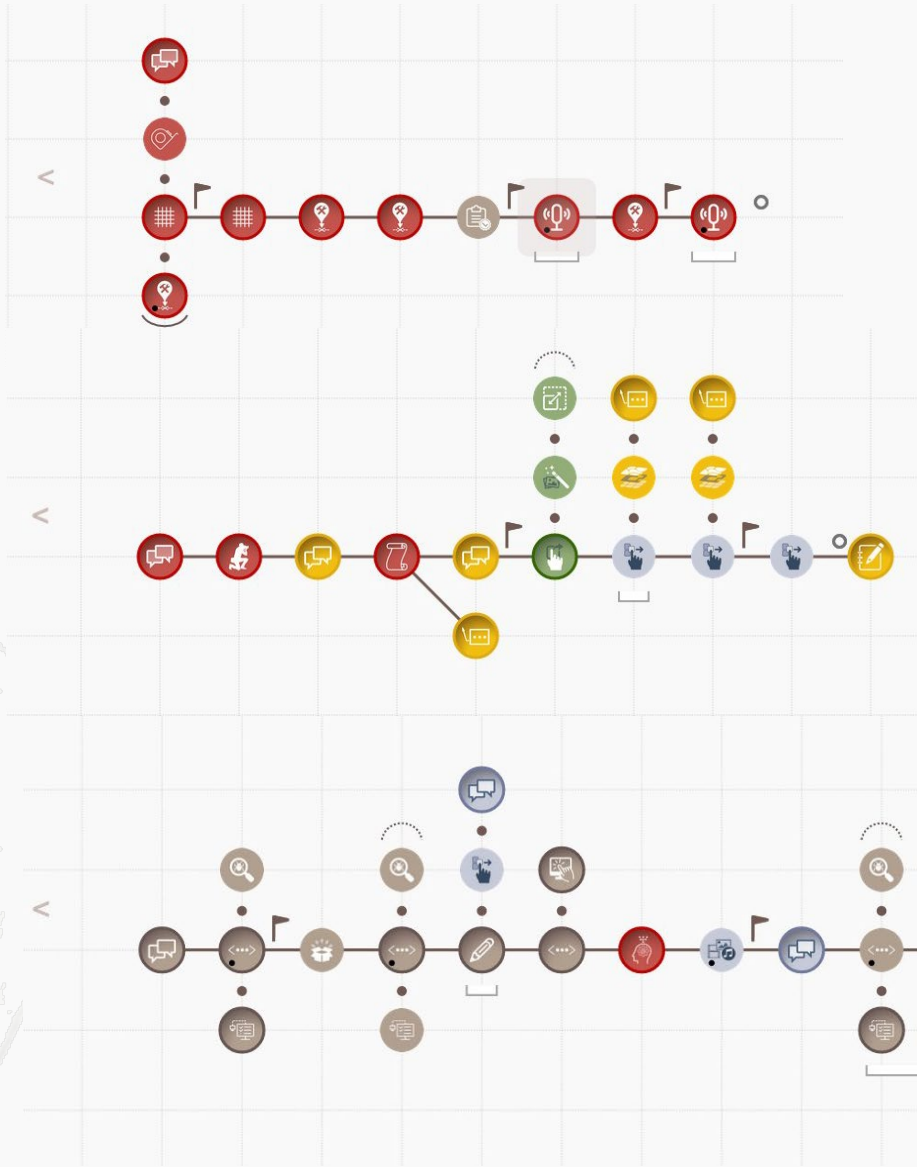
archaeological fieldworks

process described on the basis of an archaeological report

Chapelle Notre-Dame de Bethléem, Commune de Bras, Var, 2017



Spatio-acoustic data acquisition



acoustic data acquisition

Sesames seminar (Villeneuve lez Avignon, Jean Salusse room)

preparation of spatio-acoustic acquisition

Sesames seminar (Villeneuve lez Avignon)

design and production of
online panoramas (360°)
with sound tracks





Spatio-acoustic data acquisition



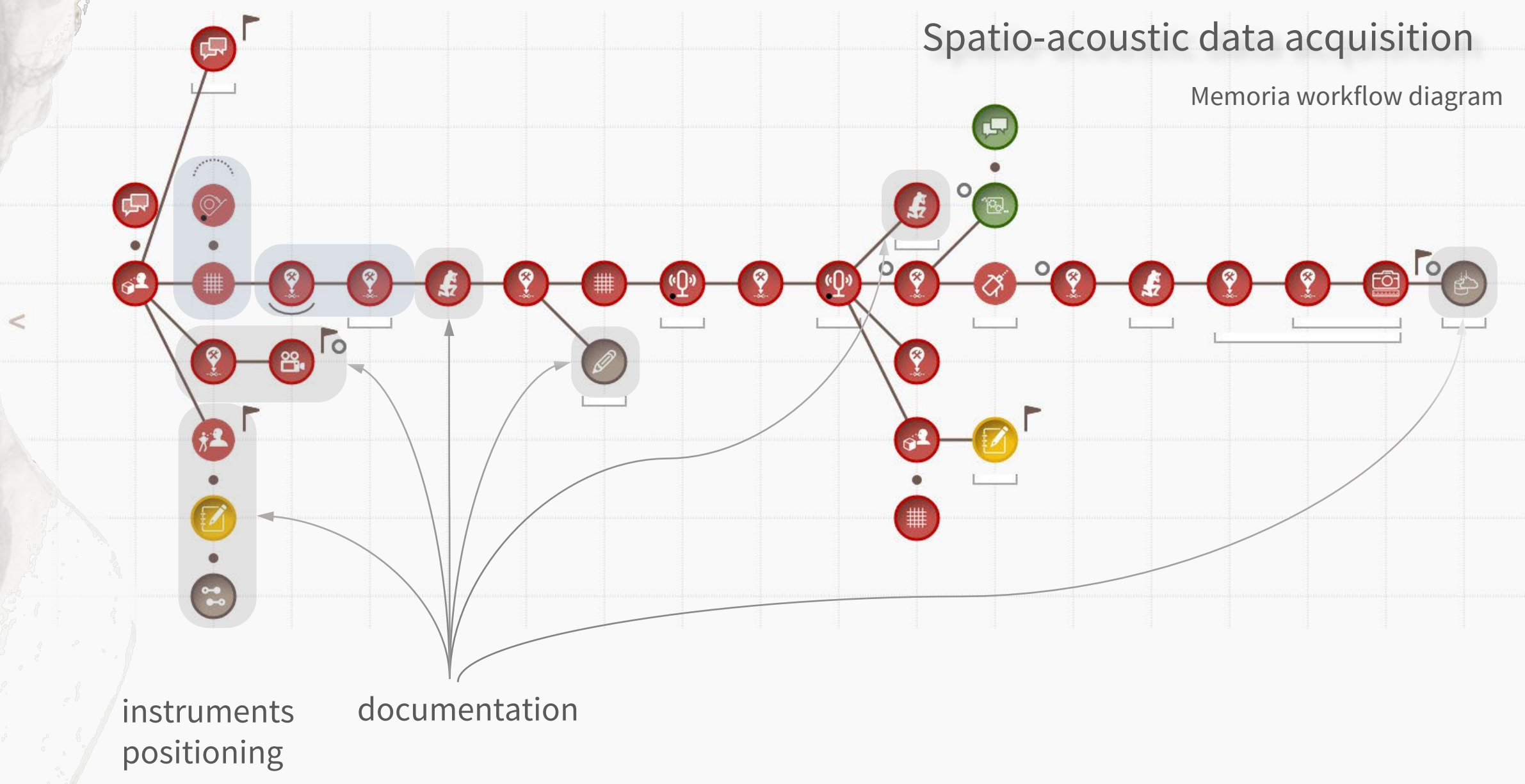
interdisciplinary acquisition protocol

actions

- instrument positioning
- acoustic data acquisition
- metric and visual data acquisition
- documentation

results





Spatio-acoustic data acquisition

Memoria workflow diagram

instruments
positioning

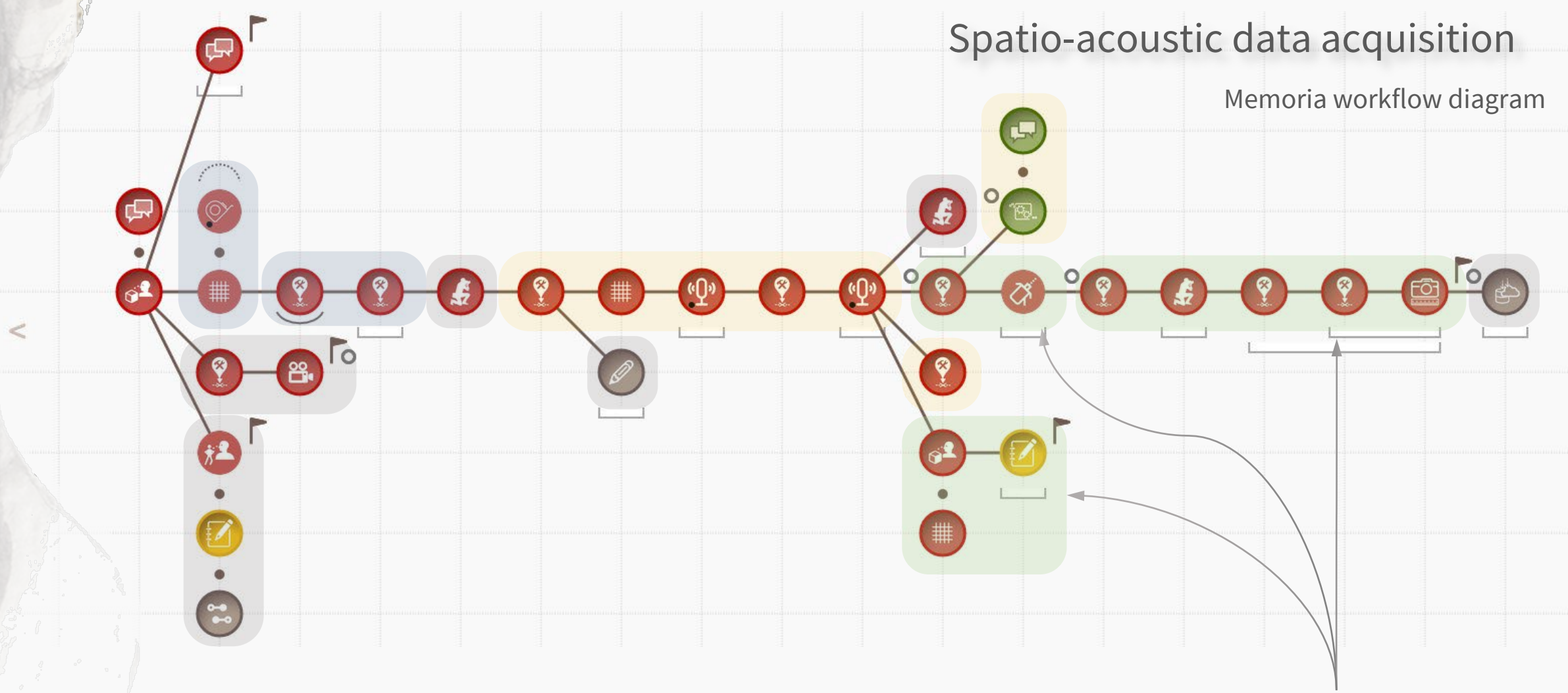
documentation

Λ

Λ

Λ

Λ



Spatio-acoustic data acquisition

Memoria workflow diagram

instruments
positioning

documentation

acoustic data acquisition
and pre-processing

laser distance measurement
photogrammetric acquisition

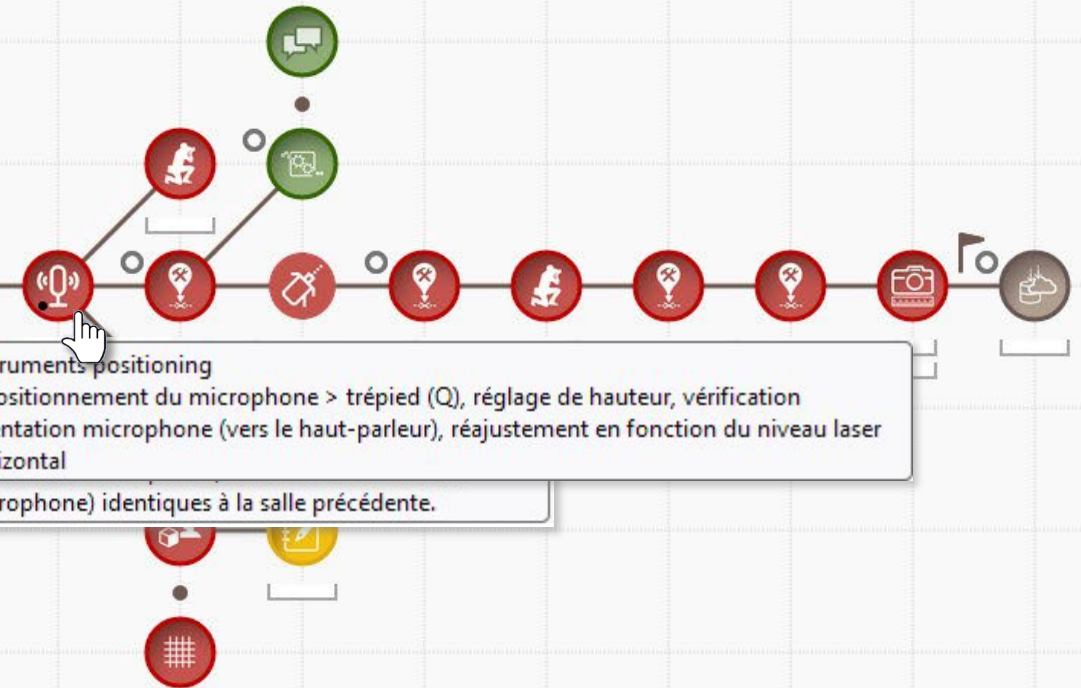


activity instance >



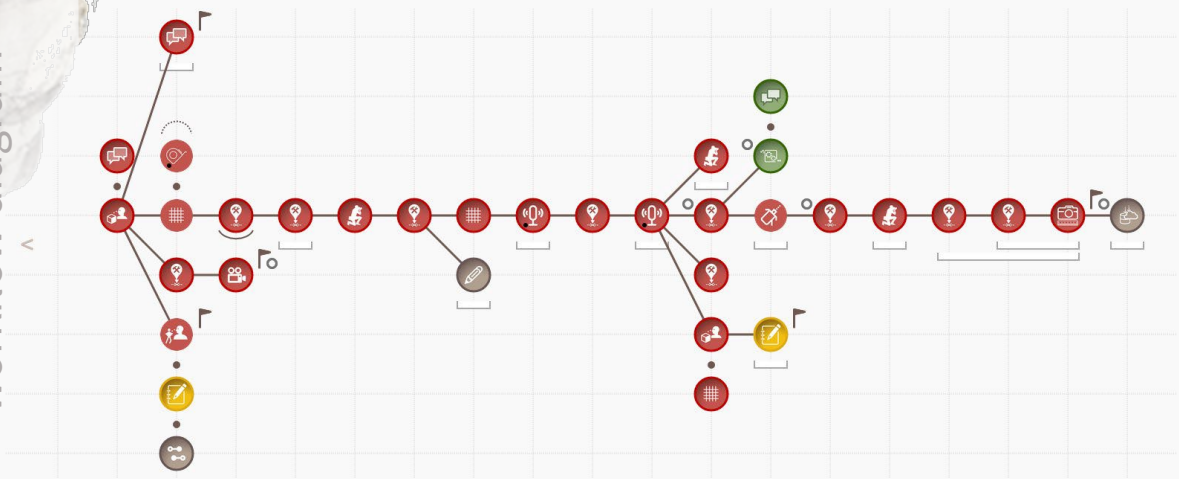
place of registration >	other indoor spaces
iterative character >	no
repetitive character >	yes
individuals >	Fargeot Simon > AMU Vidal Adrien > PRISM
collective bodies >	
sources >	
inputs >	sweep acquisition Sesames
compositions >	
publications >	
expertise >	
outputs >	
instrument(s) >	em32 Eigenmike (microphone) Genelec 8020C (speaker) MacBook Pro (mi-2014)- PRISM
software >	REAPER
operating system >	Mac OS X 10.14 (Mojave)
start date >	18/05/2022
end date >	18/05/2022
duration unit >	minute
duration value >	1
remarks >	mesures acoustiques enregistrement du sweep 3x (trépied Q) personne dans la pièce, pièce vide, porte fermée

Spatio-acoustic data acquisition



instruments positioning
 positionnement du microphone > trépied (Q), réglage de hauteur, vérification
 orientation microphone (vers le haut-parleur), réajustement en fonction du niveau laser
 horizontal
 microphone) identiques à la salle précédente.

workflow diagram



acoustic recordings

disto measurements

photogrammetric image set

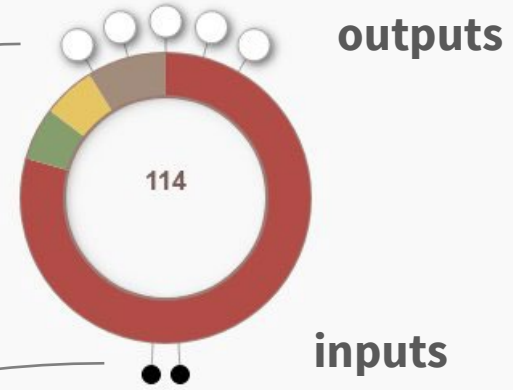
photographic documentation

video report

sweep production

protocol preparation

Information Visualisation



activities' proportion ring



acquisition de données acoustiques/métriques, protocole préparatoire au séminaire SESAMES (salle Gothique)

2022-05-18 (discussion/consultation) - 2022-06-10 (data/content feeding)

instruments:

- camera accessories > unidentified tripod
- computer > undefined tablet computer
- measuring tool > Leica DISTO S910 (avec : Trépied TRI 70, cible GZM3, adaptateur pour trépied FTA 360-S)
- measuring tool > mètre magnétique 8m DEXTER
- camera accessories > Manfrotto compact advanced (tripod)
- computer > DELL Precision 7740 (Astrolabe)
- audio equipment > em32 Eigenmike (microphone)
- audio equipment > Genelec 8020C (speaker)
- camera > Insta360's ONE R
- camera > YCaméra 360° YI 360
- camera > Samsung Galaxy A41
- computer > Bookeen Notéa V1
- camera accessories > Trépied (Manfrotto) 055 avec colonne horizontale
- computer > Crosscall Action X3
- computer > Macbook pro 16 inch 2019 (Porta)
- remote sensing accessories > coded targets of 12-bit
- remote sensing accessories > cibles 3D pour relevé des positions de trépieds (Sesames)
- remote sensing accessories > Huepar
- computer > MacBook Pro (mi-2014)- PRISM

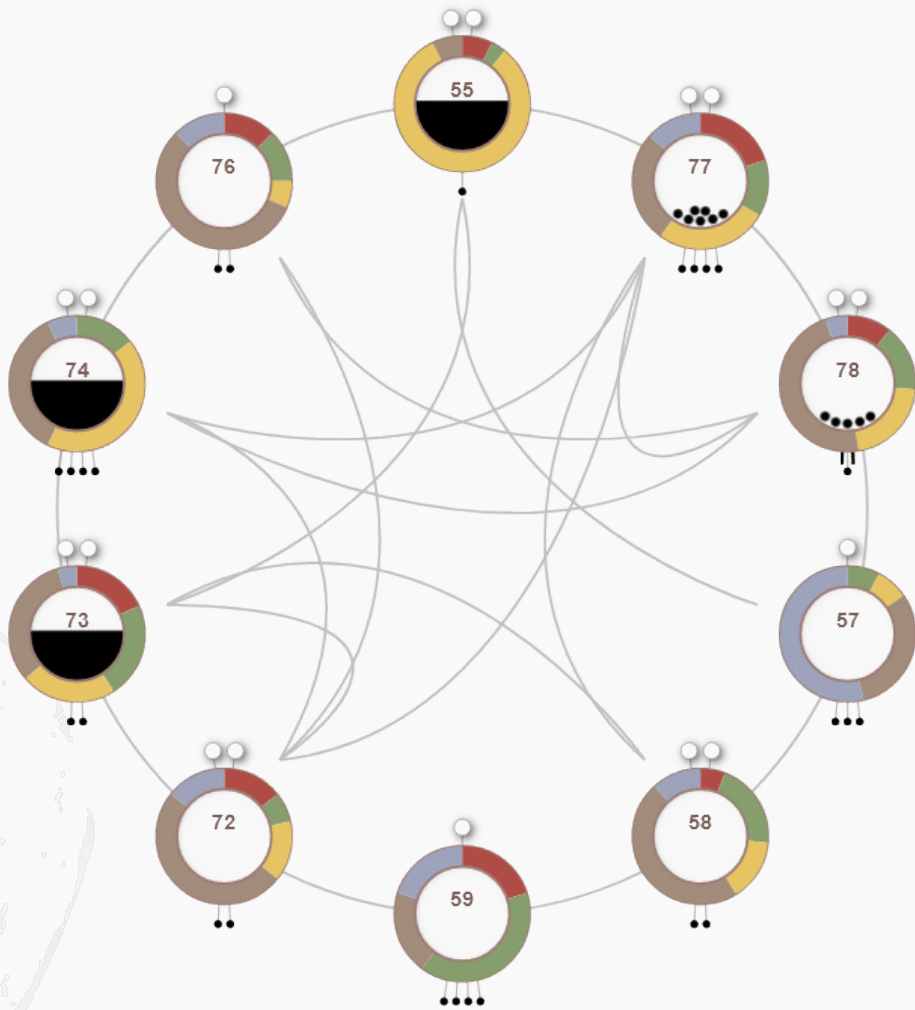
software:

- numerical analysis packages > MATLAB
- audio editing > REAPER
- visual effects, motion graphics and compositing > Insta 360 App

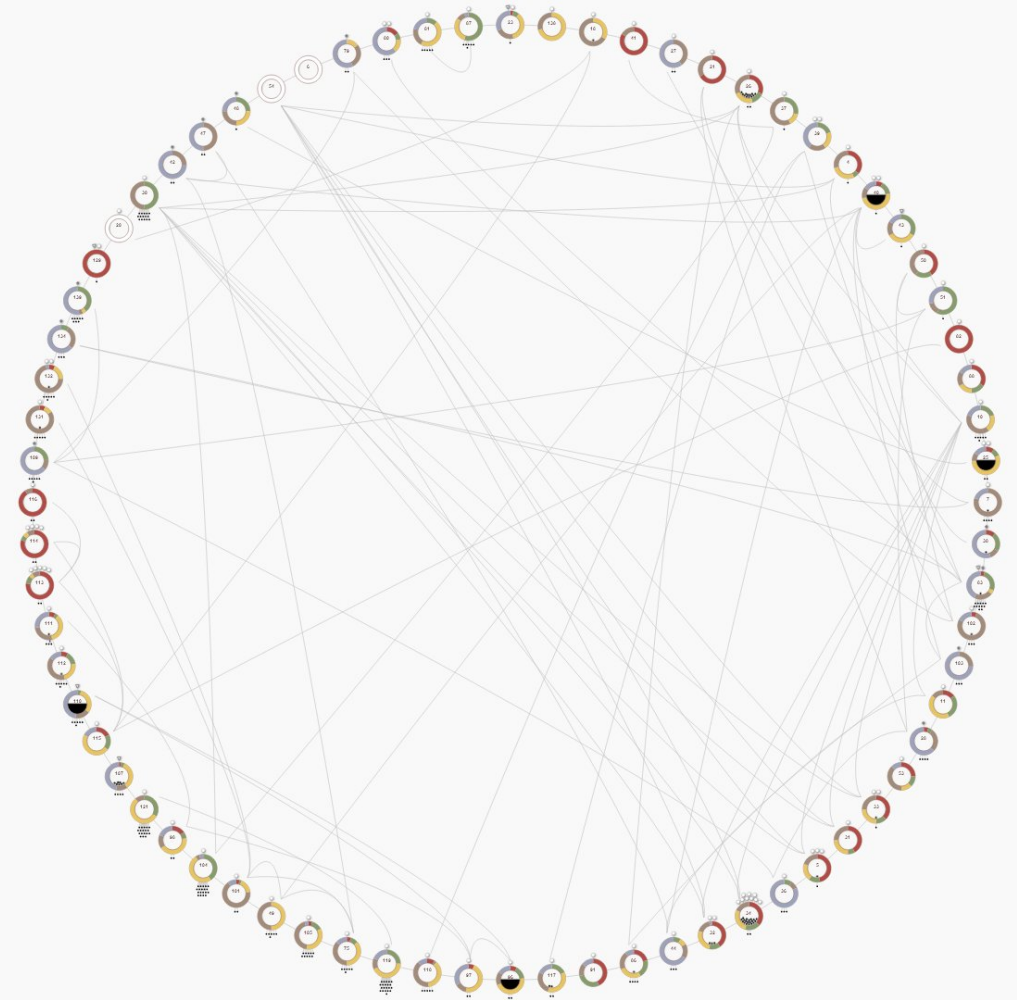


Information Visualisation

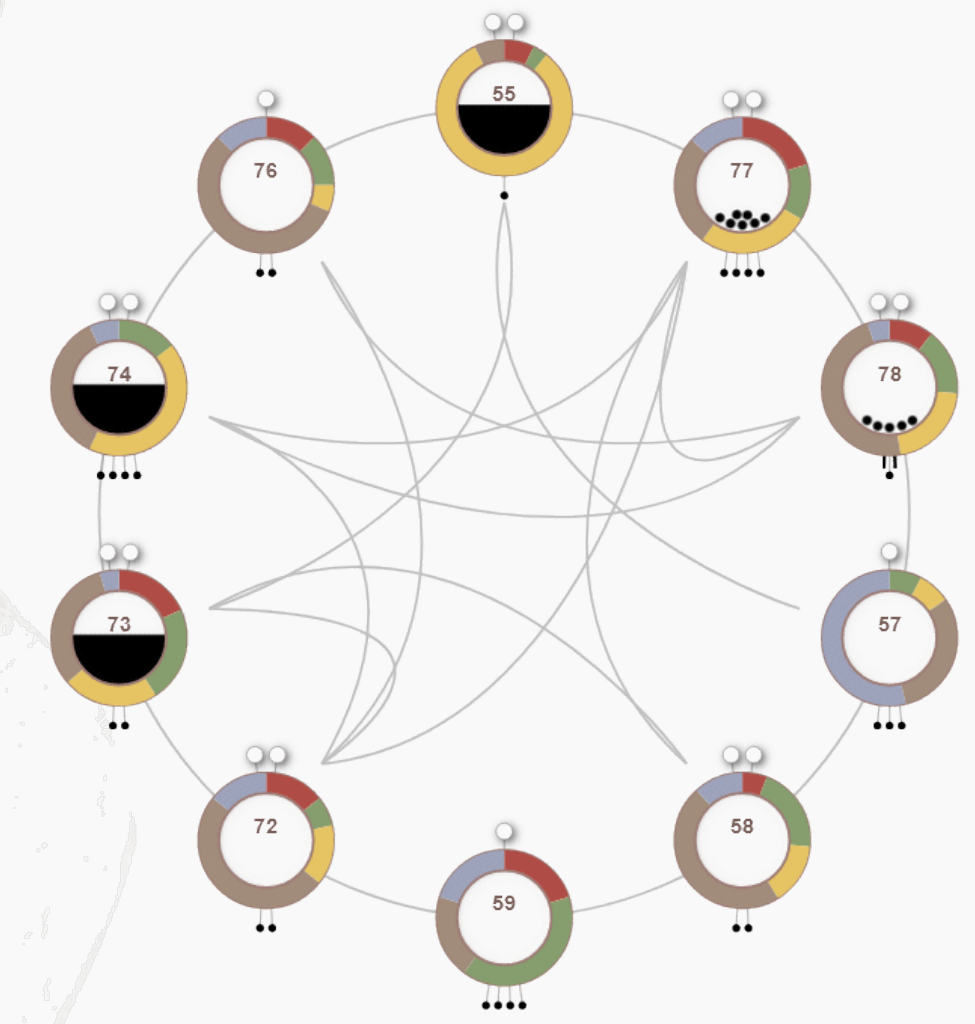
process variability chord diagram



project A



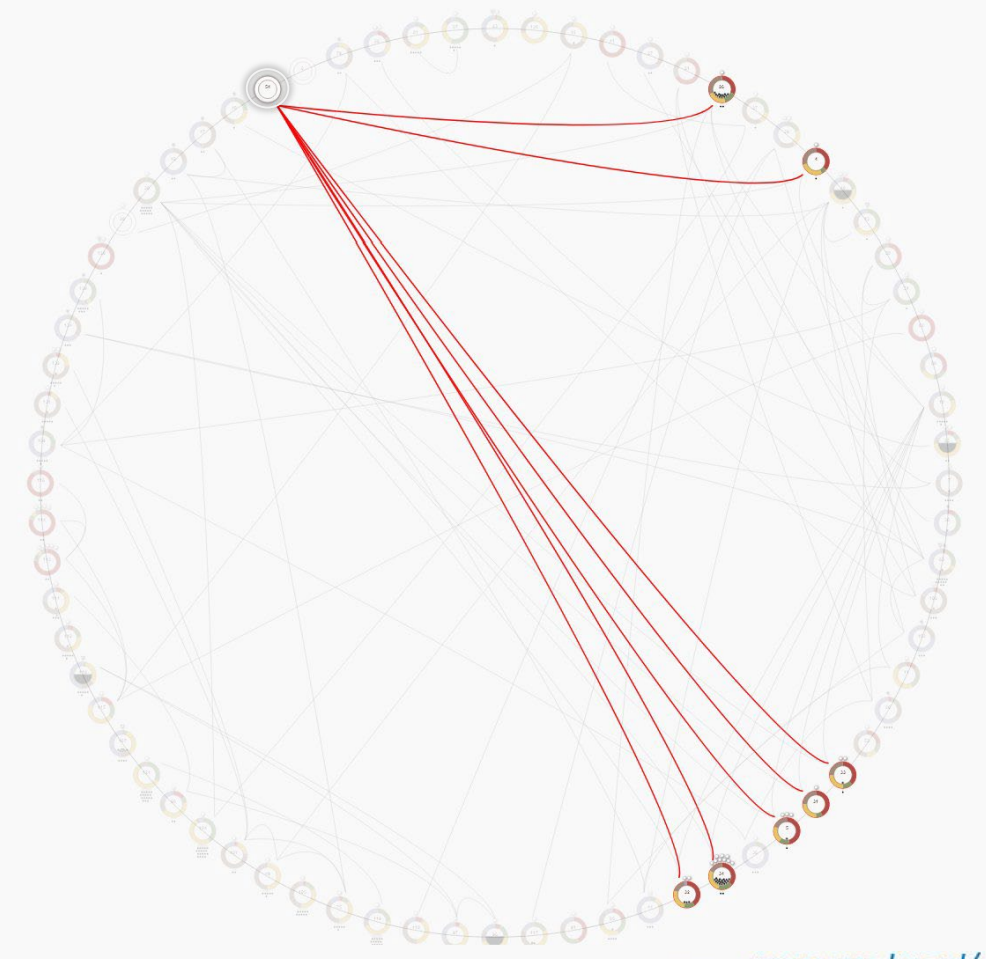
project B



project A

Information Visualisation

process variability chord diagram



project B

- processes based/depending on the selected process
- 'parent' or preceding process



Potential of the system

MEMORIA is an exploratory web-based prototype, an attempt to provide a methodological framework to meet the requirements of scientific integrity.

- documentation framework enables : tractability of results, verifiability of research protocols, comparability and analysis of different protocols, ... replicability and reproducibility of results
- support of the planning of interdisciplinary research protocols based on past experiences
- theoretical/tacit knowledge-sharing environment
- innovative visual analytical framework
- new documentation framework
by requiring more reflection, it can help us to work more effectively, more consciously and less routinely (*human inertia*)





Main constraints and limits of the system:

MEMORIA remains a human-centred system

Acceptability constrains:

- quality and reliability of the visualisations depend strongly on the data entered
- time-consuming documentation process, steep learning curve for some people
- unreadiness to openly share information about our mistakes, errors, shortcomings





Main constraints and limits of the system:

MEMORIA remains a human-centred system

Technical constrains:

- sustainability of web-based information systems
- costly and complex development steps

the constant evolution of the technology requires continuous mobilisation to ensure functionality of such systems - maintenance, development, user support and funding





future works

we'll be happy to discuss with you later today

Thanks for your attention!

Memoria Information System

for UMR 3495 CNRS/MC MAP
(open consultation)

Memoria sandbox

familiarization platform
(user account available without restrictions)

MEMORIA - classification of activities

online PDF , 334 pages
(in English)

Iwona DUDEK

UPR 2002 CNRS MAP

iwona.dudek@map.cnrs.fr



<http://memoria-dev.gamsau.archi.fr/projet/index.php>