Information System Support for Analytical Methods

Martin Doerr

Center for Cultural Informatics,
Institute of Computer Science
Foundation for Research and Technology - Hellas

Heraklion, October 18, 2017

Analytics & Information

Requirements

Empirical science is based on observation.

Observation data cannot be understood without knowledge about the ways and circumstances of their creation.

Scientific analytical examinations are observations

Data Evaluation is based on observation records and hypotheses

Analytical examinations occur in:

medicine, environmental control, manufacturing control, geology, biodiversity, paleontology, art conservation, archaeology, forensics….where not?

Analytics & Information

Requirements

Generally, scientific data cannot be understood without knowledge about the meaning of the data and the ways and circumstances of their creation

We need metadata ("paradata") to assess

- meaning (view, experimental setup, instrument settings),
- relevance (measured things, their status, their conditions),
- quality (calibration, tolerances, **errors**, "artifacts"),
- possibilities of Improvement and Reprocessing.

From generation to use, permanent storage, reuse (life-cycle)

➔ No standards yet!

Analytics & Information

Requirements

Data Acquisition: *Reliable registration of the process and context conditions*

- The experimental setup and environment (tools, light/activation sources, geometry, sources of noise/reflections etc.)
- Capture device type, identity (individual behavior!)
- Need a hierarchical model: Inherit metadata common to series of “shots”
- The identity of the measured or depicted ***object***
  - *import identifiers, metadata of the object.*
  - *identity of location/ spot*

Analytics & Information

Requirements

Data Processing: *Reliable registration of parameters*

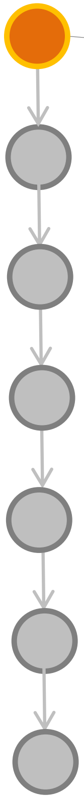
- Workflow logs, reliable identification of outputs with inputs
  - *input files (URIs!)*
  - *output files (URIs!), formats, warning and error reports.*
  - *S/W identifiers and parameters, manual adjustments!*
  - *process types for reasoning*
- Reliable linking with measured data

Data Use and Reuse*: parts, wholes and annotation:*

- Composition of final products from multiple trials, client/archival information packages
- Migration to other formats (compatibility and obsoletion)
- Authenticity, rights

LIBS-RAMAN-MULTISPECTRAL

FLOW DIAGRAM



*

Object Reception Report

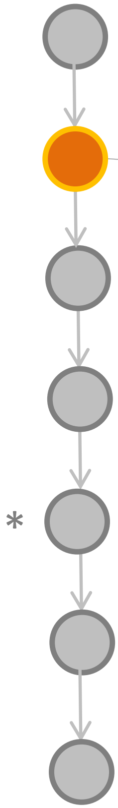
Description of object, client order and receipt

BRIEFLY

IESL-FORTH registration number: 20120717_Leontaraki_1

Condition Report

Object:	Oil painting
Theme:	Mythological
Artist:	Unknown-no signature
Painting Technique:	Oil painting on canvas with preparation
Dimensions:	Width: 64 cm Height: 76cm
Date:	-
Provenance:	Private collection (*Ms Danae Leontaraki*)
IESL Reference No.:	20120717_Leontaraki_1
Date of analysis:	17th and 18th of July, 2012



Optical Examination of Object

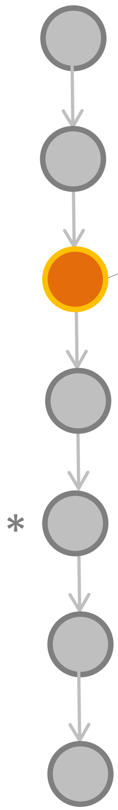
Find examination spots, create a map of spots and areas

Macroscopic examination / Description of the preservation condition

Canvas, stretcher and lining:	The original broadly woven, plain weave linen canvas has been composition-glue lined onto a similar weave linen canvas. The canvas is attached with tacks to wooden stretcher which is in a good condition. The canvas is lined overall with what appears to be an aqueous	
Ground layer	The paint has been applied over a fine ground that appears to be in good cohesion and adhesion to the lining canvas. In certain areas the exposed ground layer has a rich red earth color.	
Painting layer	The paint layer is characteristic of an oil binding medium and appears in fairly good condition. The painting suffered from extensive damages and paint losses and shows irregular structure. Previous restorers have re-lined the canvas and over-painted the losses to a rather poor standard. As a result several features of the painting have been changed, as for example the face of "Venus" at the middle of the composition. Quite heavy age-cracking has occurred in the original paint; some of which is visually disturbing. In these areas a reddish ground layer has been exposed.	 Cracks and looses  Irregular structure  Retouching

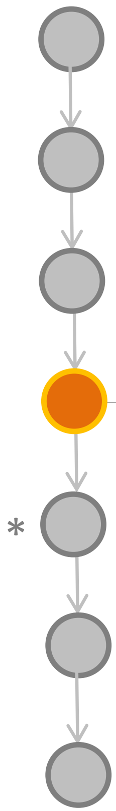
MAP

Points and areas on the painting, where LIBS and Raman analysis was conducted



Preparation

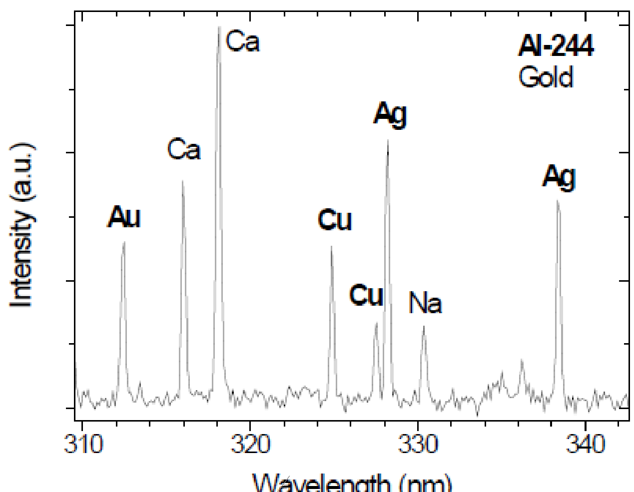
Calibration, surface cleaning, spot description, focussing and adjusting instruments



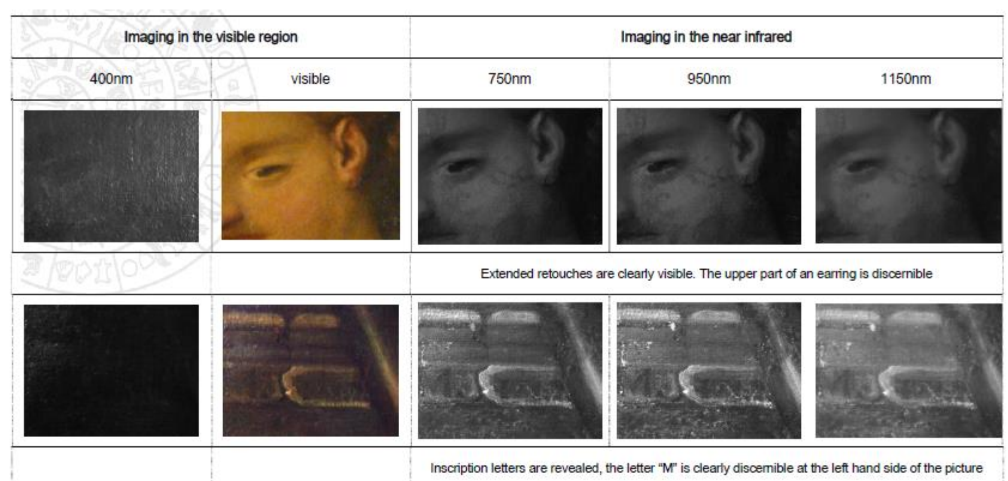
Scanning

LIBS-RAMAN: Εμφάνιση γράφου, εντοπισμός σημαντικών κορυφών και καταγραφή χαρακτηριστικών δεδομένων (π.χ. nm κορυφών)

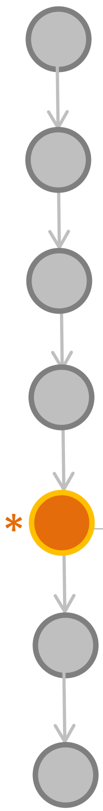
MULTISPECTRAL: Φωτογραφικές λήψεις σε διάφορα μήκη κύματος



LIBS-RAMAN



MULTISPECTRAL



Post-processing

MULTISPECTRAL: Image flattening, Registration, Spectral response

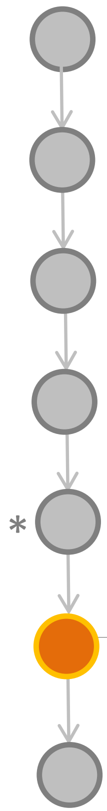
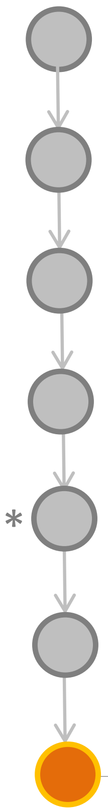


Table 4 Summary of results of blue paint analysis

Sample	LIBS	Raman	XRD	SEM-EDX (X-ray mapping)	Comment
KN01	See Table 6	See Table 6	Calcite, talc, riebeckite, muscovite, quartz, clinochlore	Calcite, talc, Mg-riebeckite, muscovite, quartz, clinochlore	Macroscopically the color is black but in polished section grey-blue
KN02	Al, Ba, Ca, Fe, Mg, Na, Si, Sr	Intense fluorescence at 890 nm, Egyptian blue	Riebeckite, calcite, talc, muscovite, quartz	Mg-riebeckite, calcite, talc, muscovite, quartz	~30–80 μm red layer over a ~200 μm layer of Mg-riebeckite and Egyptian blue (see Fig. 2)
KN03	Al, Ba, Ca, *Cu*, Fe, Mg, Na, Si, Sr	Intense fluorescence at 890 nm, Egyptian blue	Cuprorivaite, calcite, quartz, kaolinite	–	
KN04	Al, Ba, Ca, *Cu*, Fe, Mg, Na, Si, Sr	Intense fluorescence at 890 nm, Egyptian blue	Cuprorivaite, calcite, cuprite, talc, riebeckite, illite	–	
KN05	Al, Ba, Ca, *Cu*, Fe, Mg, Na, Si, Sr	Intense fluorescence at 890 nm, Egyptian blue	Cuprorivaite, calcite	–	
KN06	Al, Ba, Ca, *Cu*, Fe, Mg, Na, Si, Sr	Intense fluorescence at 890 nm, Egyptian blue	Cuprorivaite, calcite, talc, riebeckite, muscovite, chlorite	Cuprorivaite, calcitetalc Mg-riebeckite	~200 μm layer similar to KN01
	Mg, Na, Si, Sr	890 nm, Egyptian blue			
VD01	Al, Ba, Ca, *Cu*, Fe, Mg, Na, *Pb*, Si, Sr	Intense fluorescence at 890 nm, Egyptian blue	–	–	
VD04	Al, Ba, Ca, *Cu*, Fe, Mg, Na, *Pb*, Si, Sr	Intense fluorescence at 890 nm, Egyptian blue	Cuprorivaite, calcite, quartz, cristobalite	–	Cristobalite is relict or residue from the Egyptian blue synthesis
VD08	Al, Ba, Ca, *Cu*, Fe, Mg, Na, Si, Sr	Intense fluorescence at 890 nm, Egyptian blue	Cuprorivaite, calcite, quartz, hematite, dioptase		
VD09	Al, Ba, Ca, *Cu*, Fe, Mg, Na, Si, Sr	Intense fluorescence at 890 nm, Egyptian blue	Calcite, quartz, hematite, talc, aragonite	–	
G01	Al, Ba, Ca, *Cu*, Fe, Mg, Na, *Pb*, Si, Sr	Calcite	–	–	
G04	Al, Ba, Ca, Fe, Mg, Na, Si, Sr	–	Calcite, quartz, kaolinite, hematite, glauconite	–	
G06	Al, Ba, Ca, *Cu*, Fe, Mg, Na, Si, Sr	–	Calcite, quartz	–	
G07	Al, Ba, Ca, Fe, Mg,	Calcite	Calcite, quartz	–	

Interpretation:

Optical observation, identification of spectral lines, comparison with material properties etc. , conclusions

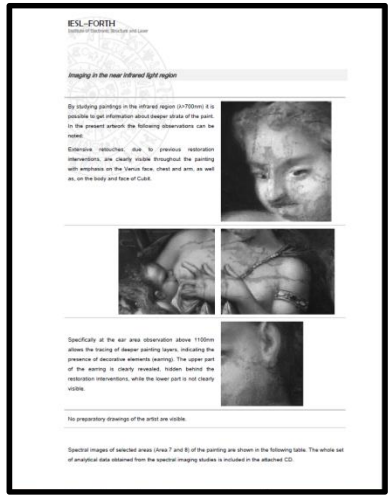


Report generation


IESL – FORTH

ANALYSIS REPORT

- Spectral Imaging
- Laser Induced Breakdown Spectroscopy (LIBS)
- Raman Spectroscopy


IESL–FORTH

Imaging in the near infrared light region

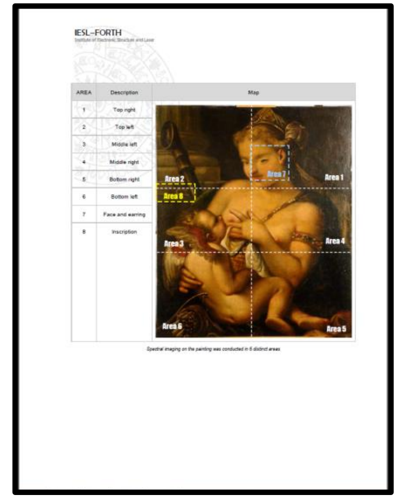
By studying paintings in the infrared region (λ>700nm) it is possible to get information about deeper strata of the paint. In the present artwork the following observations can be noted:

Extensive retouches, due to previous restoration interventions, are clearly visible throughout the painting with emphasis on the Venus face, chest and arm, as well as, on the body and face of Cubit.

Specifically at the ear area observation above 1100nm allows the tracing of deeper painting layers, indicating the presence of decorative elements (earring). The upper part of the earring is clearly revealed, hidden behind the restoration interventions, while the lower part is not clearly visible.

No preparatory drawings of the artist are visible.

Spectral images of selected areas (Area 7 and 8) of the painting are shown in the following table. The whole set of analytical data obtained from the spectral imaging studies is included in the attached CD.


IESL–FORTH

AREA	Description	Map
1	Top right	
2	Top left	
3	Middle left	
4	Middle right	
5	Bottom right	
6	Bottom left	
7	Face and earring	
8	Inscription	

Spectral imaging on the painting was conducted in 6 distinct areas

Riddle: 10.000 methods => 10.000 different systems?

What are the best abstractions to simplify system design?

Experimental science:

- Choose a phenomenon C_i,("colorant"), activation signal A
- Measure signal response S
- isolate ("clean"), vary, impact factors F_j, until signal becomes predictable.
- "predictable" means constant or formula.
- Formula requires parametrization of things & activation:
- Requires hypothesis of negligible factor N_k

=> $S = F(C_i, F_j, A, N_k)$

Analytics & Information

How to Make Information Systems

Analytical science:

- Given some object, unknown phenomenon of known kind (“which colorant?”)
- Select a measurement (and activation) method
- Measure signal response S
- Find the phenomena.

=>

- Make hypothesis of negligible factors N_k
- Determine known impact factors F_j
- Invert signal function for C_i

$$\langle C_i \rangle = F^{-1}(S, F_j, A, N_k)\ldots\ldots$$

….in general impossible

*An **analytical method** exists, if F^{-1} can be found!*

Analytics & Information

How to Make Information Systems

Problem A, the reference knowledge:

- Document signal function $F(C_i, F_j, A, N_k)$
- Find inverse function
- Knowledge of F, impact factors F_j, negligible factors $N_{k,}$ change:

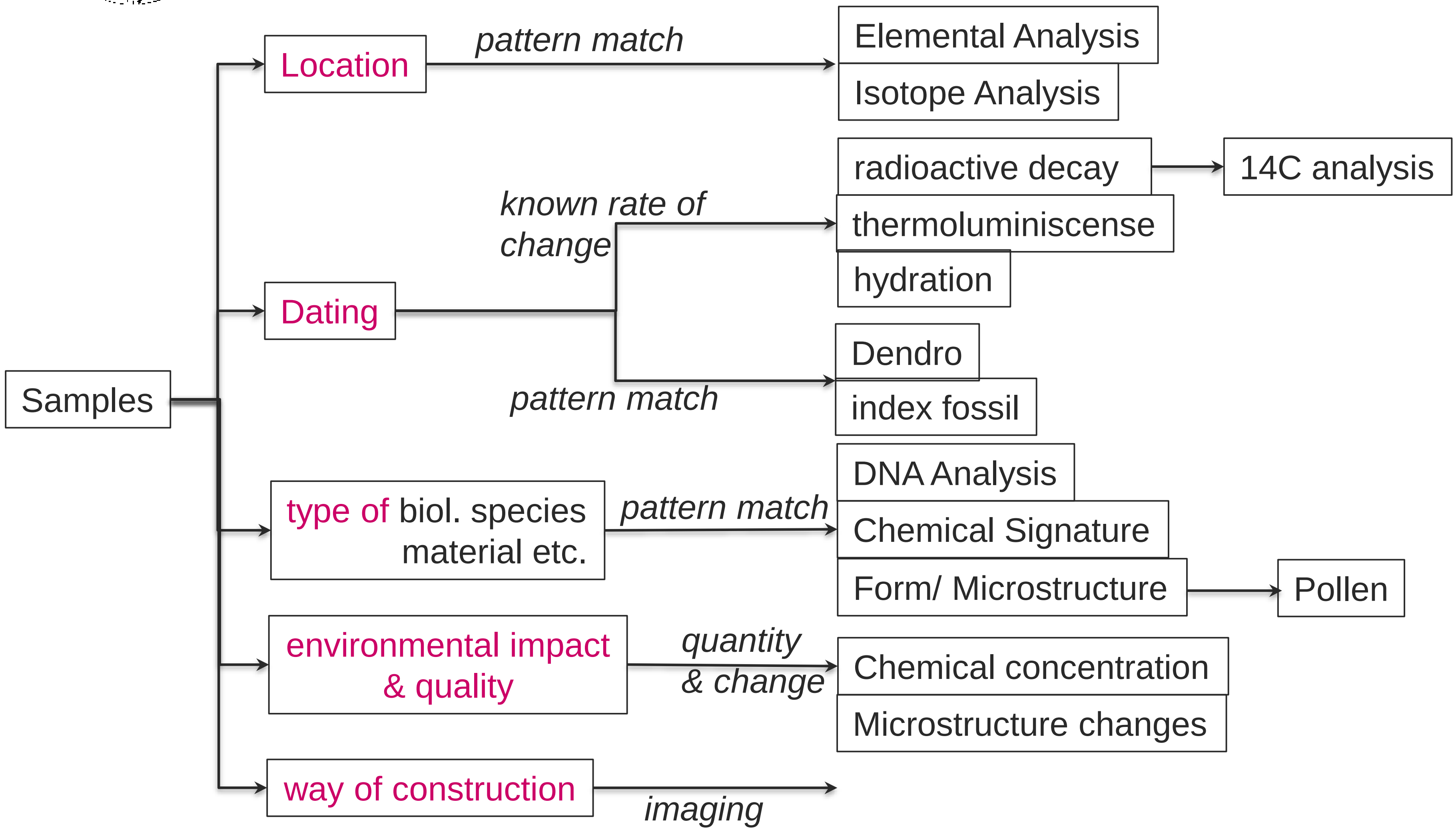
  We learn, we need to learn !
- ⇒ Create a ***Reference database***

Problem B, the description of the observation ("provenance"):

- Document for each observation $F(C_i, F_j, A, N_k)$
- Completely, even for unknown factors??
- So that quality of (old) diagnoses can be judged
- So that old diagnoses can be improved
- So that the reference base **can learn**!

Analytics & Information

Questions of Analytics



Good News, analytical methods are very similar in a sense!

We have investigated 10 methods:

DNA,
Thermoluminescence
C14
Isotope Analysis
Dendrochronology
LIBS for clay provenance, for colorant analysis
Ground Penetrating RADAR
Multispectral Imaging
….

Processes, Parameters & Error sources ofScientific Analysis

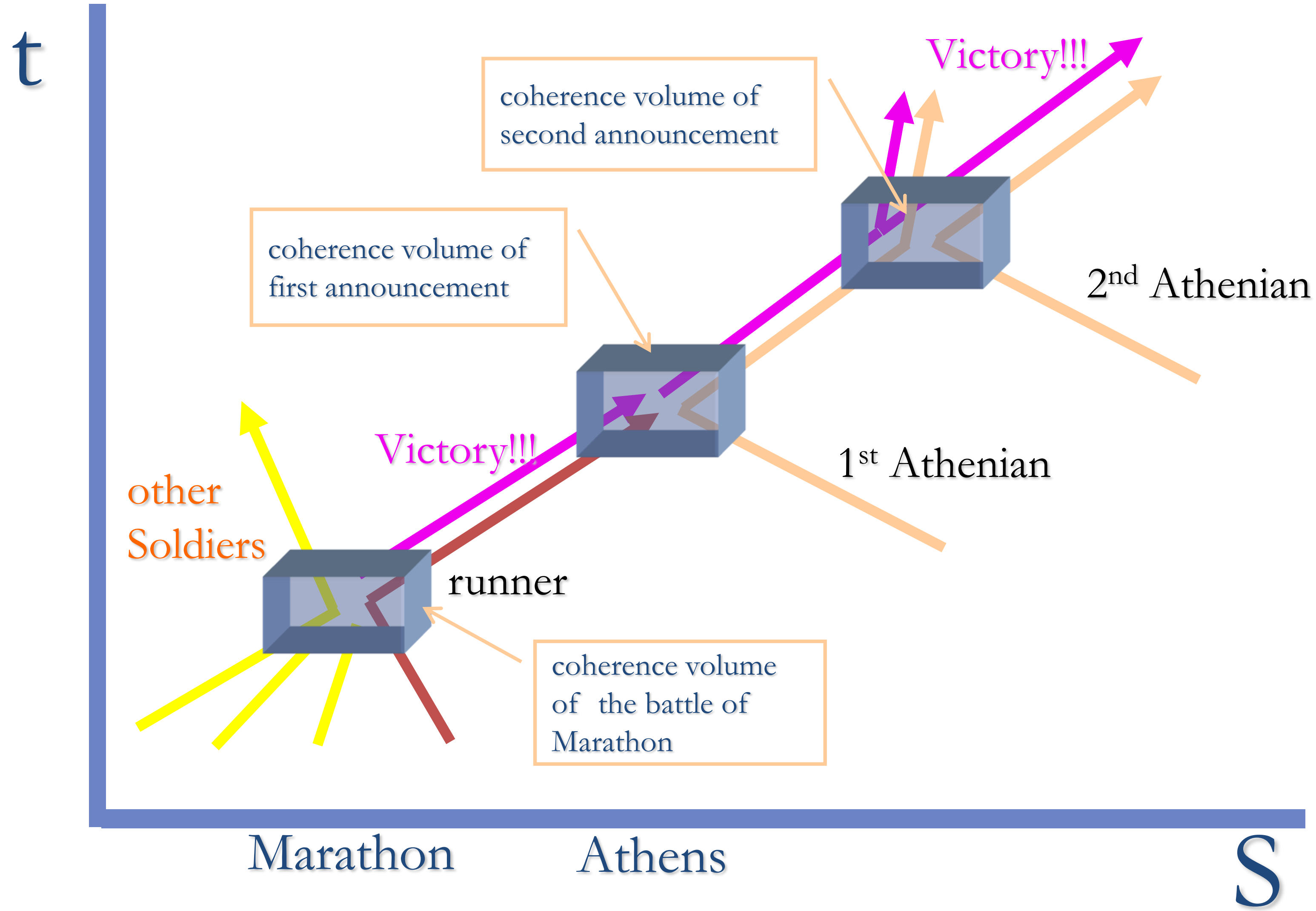
Proces	Parameters		Error Source
1. Environment/ Context	Position	Phenomenal Place	Wrong position in environment:
	Surrounding substance	Physical Feature	Poluted, not representative, not suited for analysis
	known history		not suited for analysis
2. Object	adequacy of the kind	Physical Object	
	state of preservation/condition		
3. Sample taking	Position within object - Kind of structural part (surface, tooth, bone, core,…)	Sample Taking	Wrong position in object,
	condition at place / position where sample is taken	Samples	Wrong accuracy, essential property not measured
	Purity		
4. Sample preparation for measurement	Kind of method, agents used, time passed since sample taking (age of sample)	Design or Procedure	bad quality/quantity of agents
5. Measurement	method, kind of device, instance of device,..	Measurement	
6. Postprocessing	algorithm/software release, parameters used, interactive intervention	Design or Procedure	
7. Pattern for matching	quality of measured signal format, resolution	Pattern for matching	
8. Compare with reference pattern / matching	method, degree of matching, estimation of errors from the previous steps	Design or Procedure	

Idea for provenance documentation:

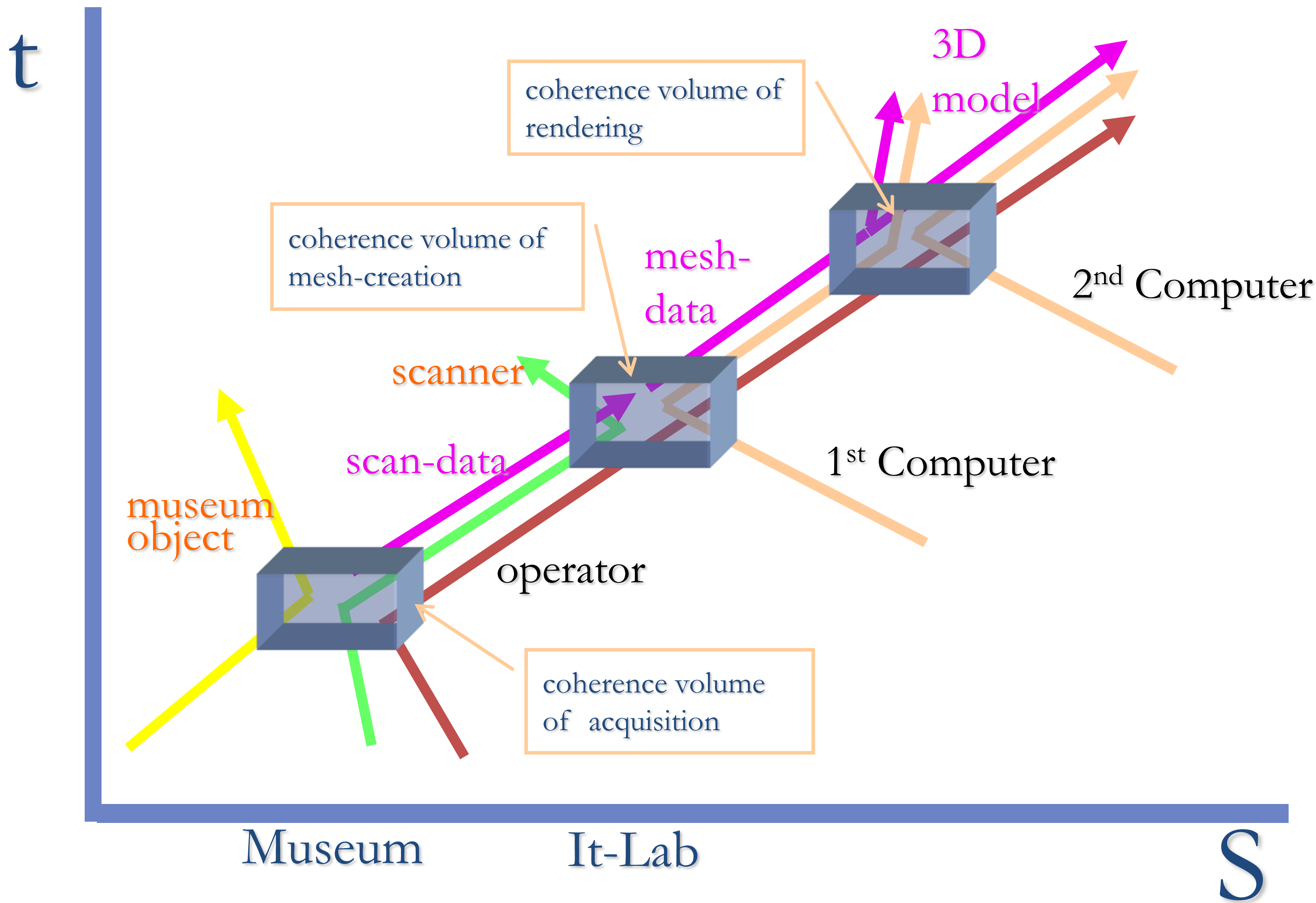
Meetings of scientists with objects, devices, data carriers…

“Feynman graphs…”

CRM:Information Exchange as Meetings



3D Model Creation as Meetings



Competitors

Competitors:

- *INSPIRE –earth science oriented*
- *OBOE – life science oriented*
- *SEEK – ecology oriented*
- *Darwin Core - biodiversity*

Problems of competitors:

- *Confuse observation process with observation record*
- *Confuse sample taking with observation*
- *Confuse finding with preparation*
- *No persistent sample identity*
- *Poor, inconsistent description of methods, environment and participants*
- *Poor identity of observed "thing"*

Analytics & Information

How to Make Information Systems

FORTH-ISL provenance metadata ("CRMSci) are currently the most powerful provenance metadata.

To be developed:

General Theory of Reference databases, an idealisation/ learning process:

- *Register **ALL** samples*
- *Find mean-value/deviation of individual phenomena ("one colorant"), decide protypicality of samples, individual variation on sample, variation between samples.*
- *Find parameters (time, provenance, heat, etc.) causing signal variation*
- *Find diagnostic power: distance of signals between phenomena*
- *Find combination behavior: Linear? Complex?*
- *Take each analysis as a new example, **REVISE** all above if necessary.*

Analytics & Information

Conclusion

Even though each analytical method is a universe of theory and experience of its own,

They follow a generic process pattern, and a few fundamental ways how data types, evaluation sequences and error sources connect.

This can be used for effective, highly parametrizable information systems for monitoring, reporting and experience building of analytical methods.