



# How to describe data: Metadata Schemas

---

Rossella Aversa

KARLSRUHE INSTITUTE OF TECHNOLOGY

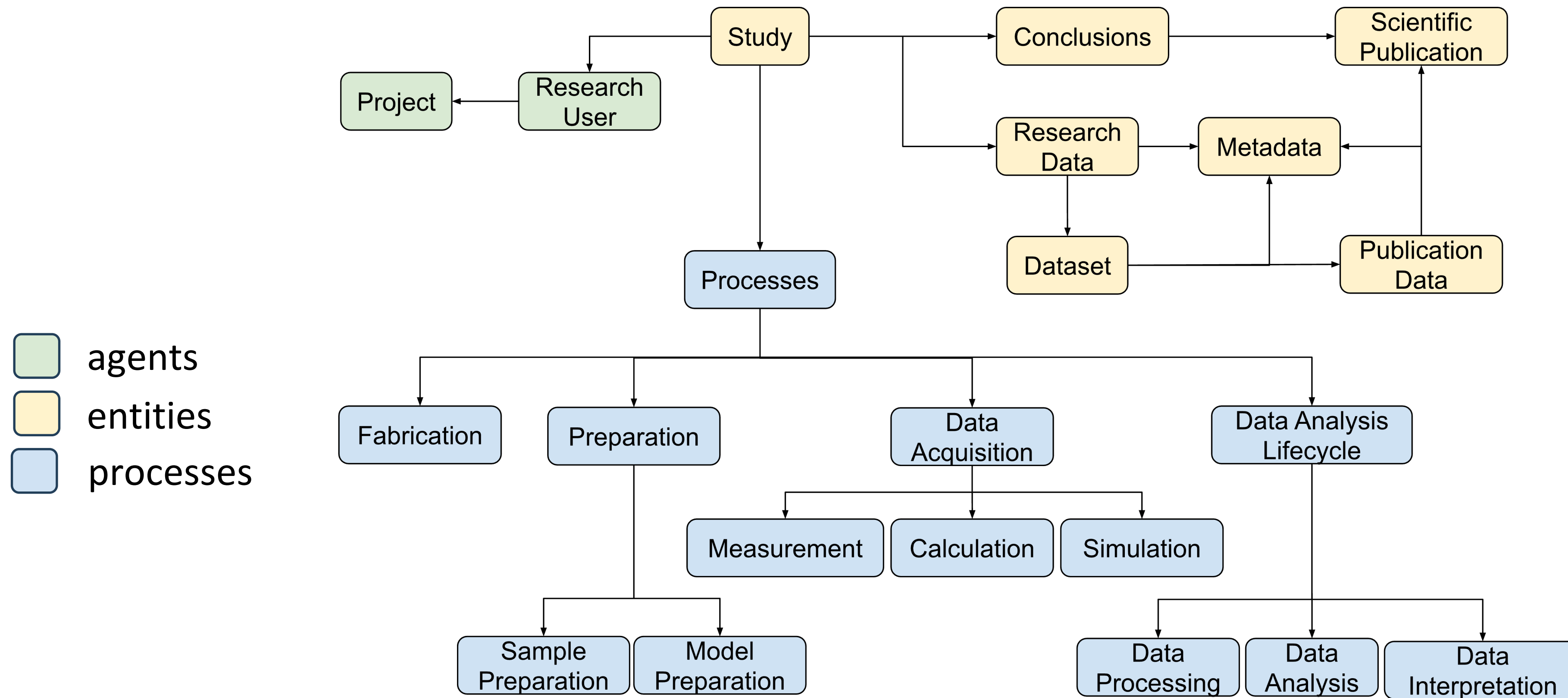
# Recap of what you have learnt

- What are the FAIR Principles and the advantages of following them
- The disadvantages of not following FAIR principles
- How the FAIR Principles are applied in NEP
- The difference between FAIR and Open
- How to plan your data activities using a Data Management Plan
- How to manage your data

# What you will learn today

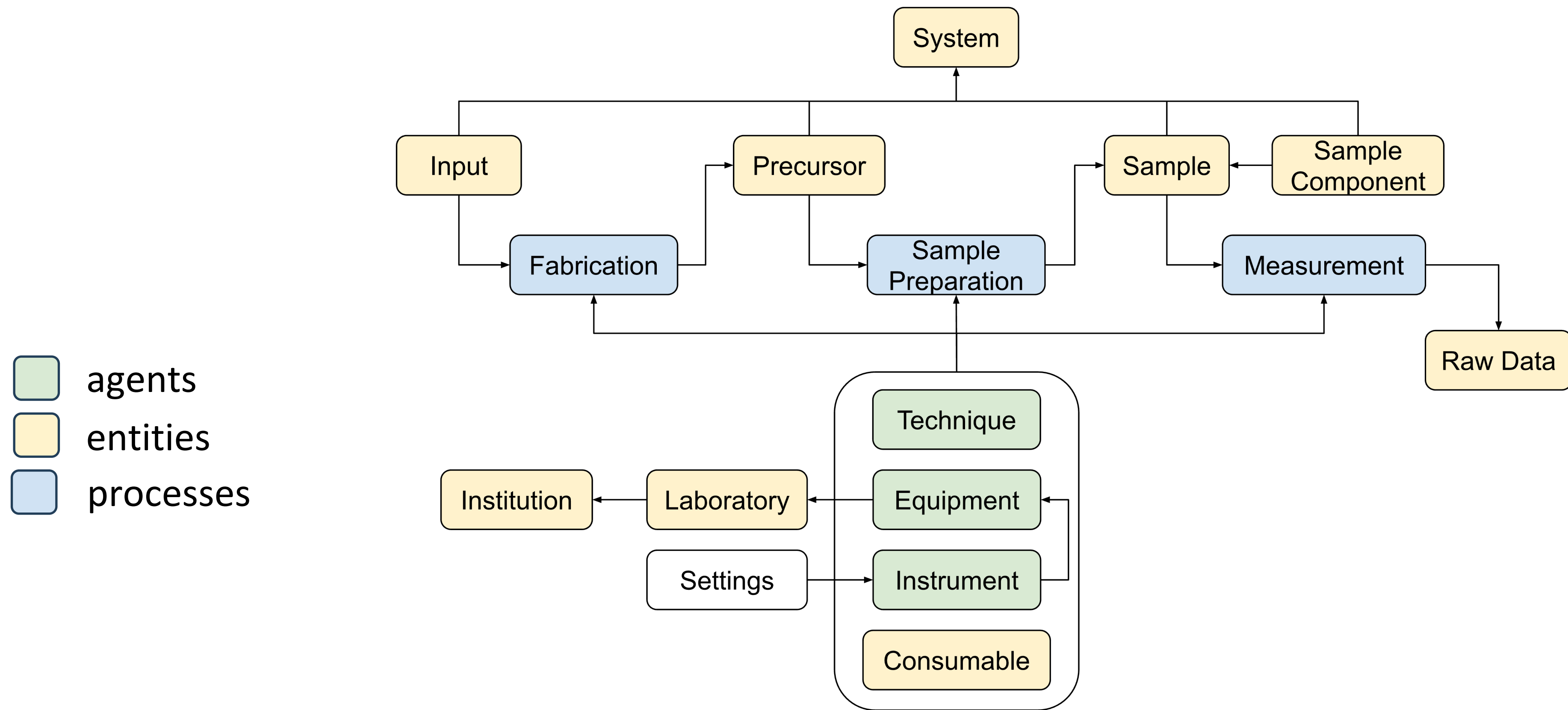
- (Almost) everything about metadata management:
  - How to write metadata to properly describe your data?
  - Where to store metadata?
  - How to find metadata and the data they describe?
  - Which tools can you use?

# What can be described?



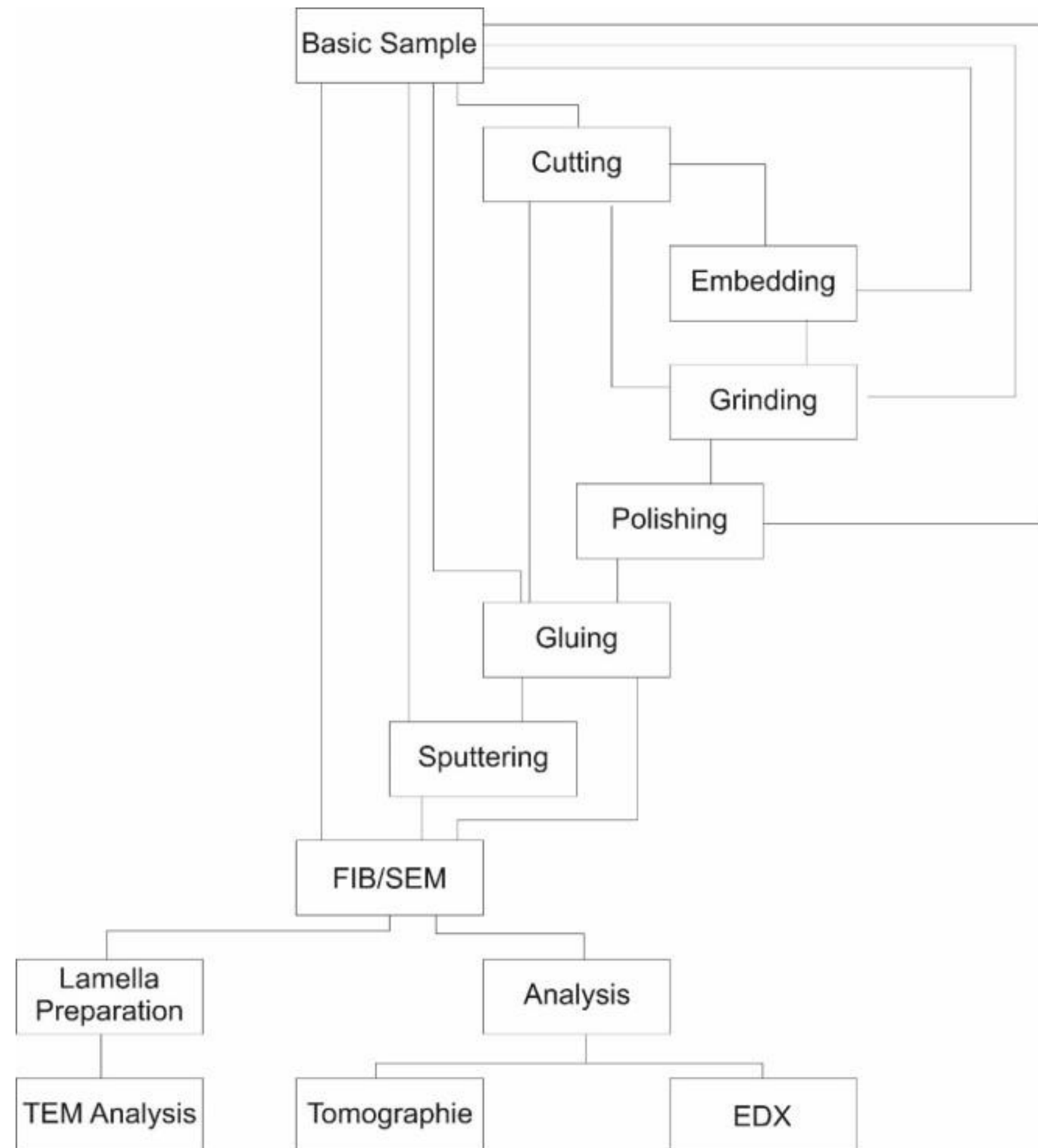
MDMC-NEP Glossary of Terms. DOI: [10.5281/zenodo.10663833](https://doi.org/10.5281/zenodo.10663833)

# What can be described?



MDMC-NEP Glossary of Terms. DOI: [10.5281/zenodo.10663833](https://doi.org/10.5281/zenodo.10663833)

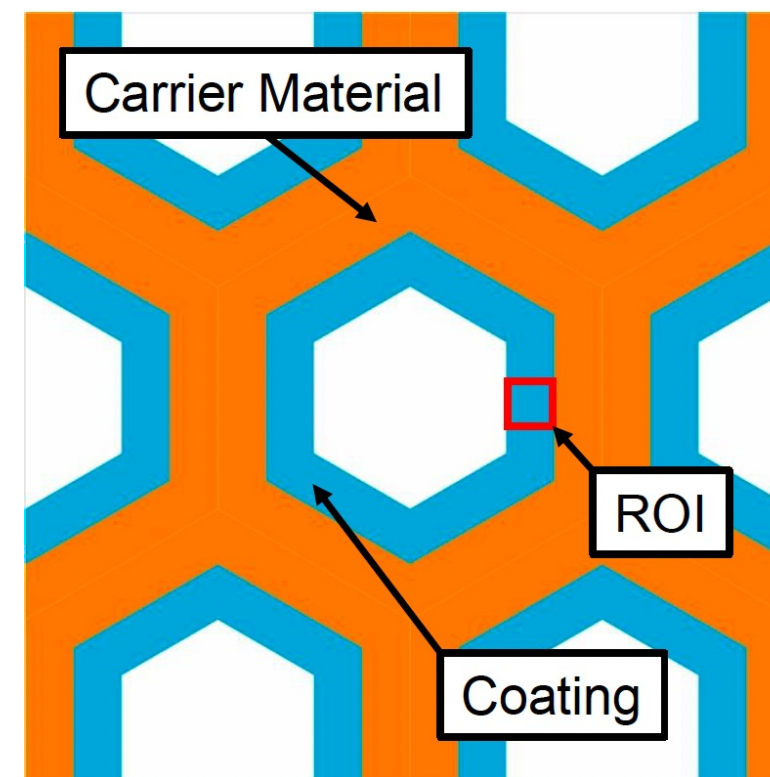
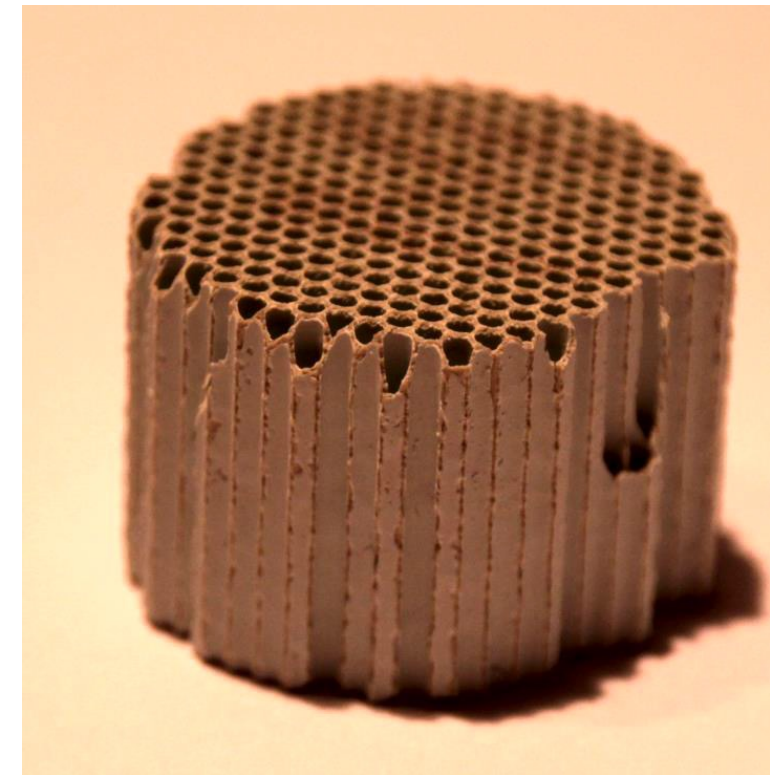
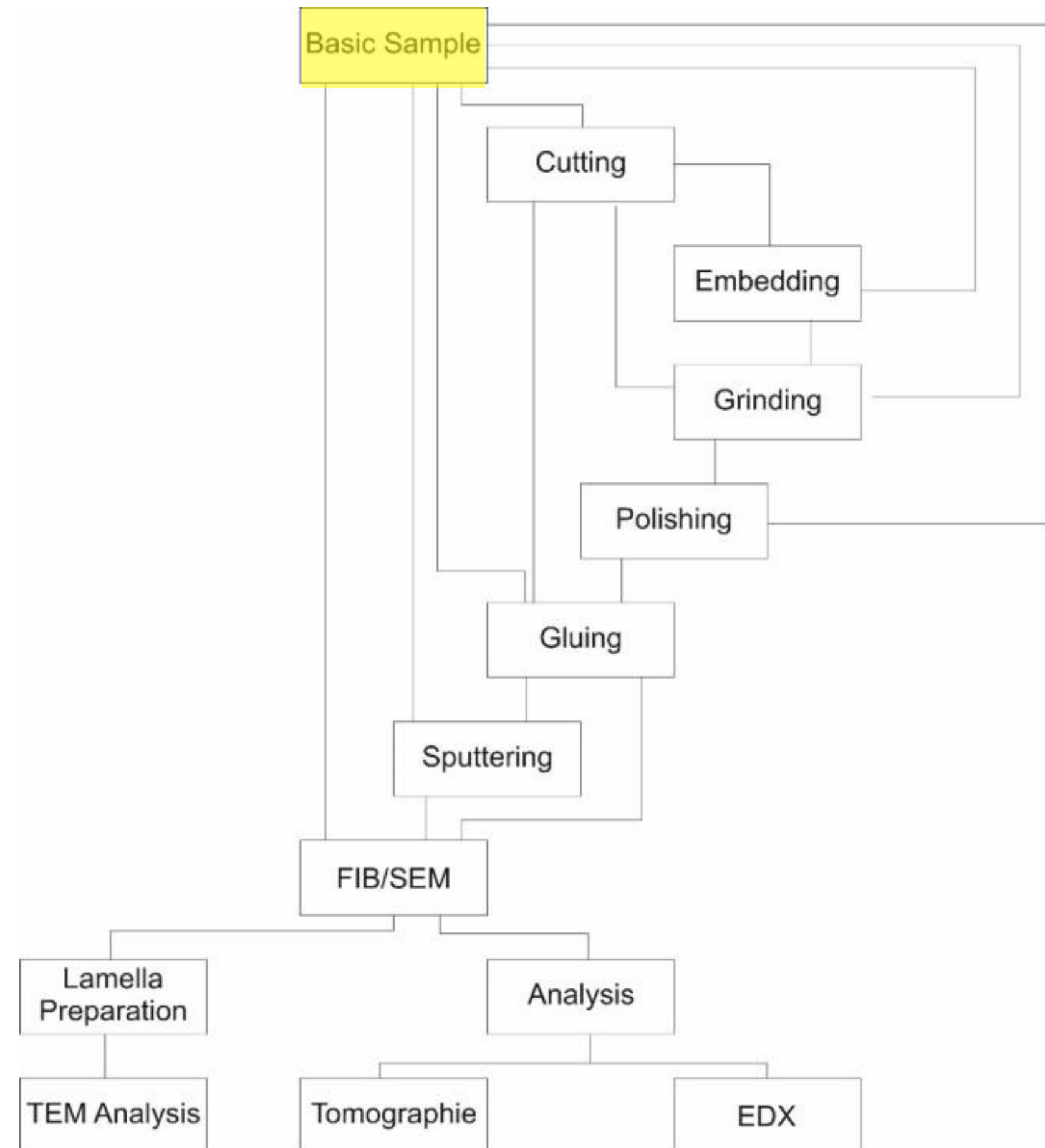
# What can be described?



- TEM lamella preparation for further measurements
- Different treatments, depending on the samples and/or on the scientific questions
- What does it need to be documented?

*M. Mail, Workflow for TEM lamella preparation – and more (2021)*

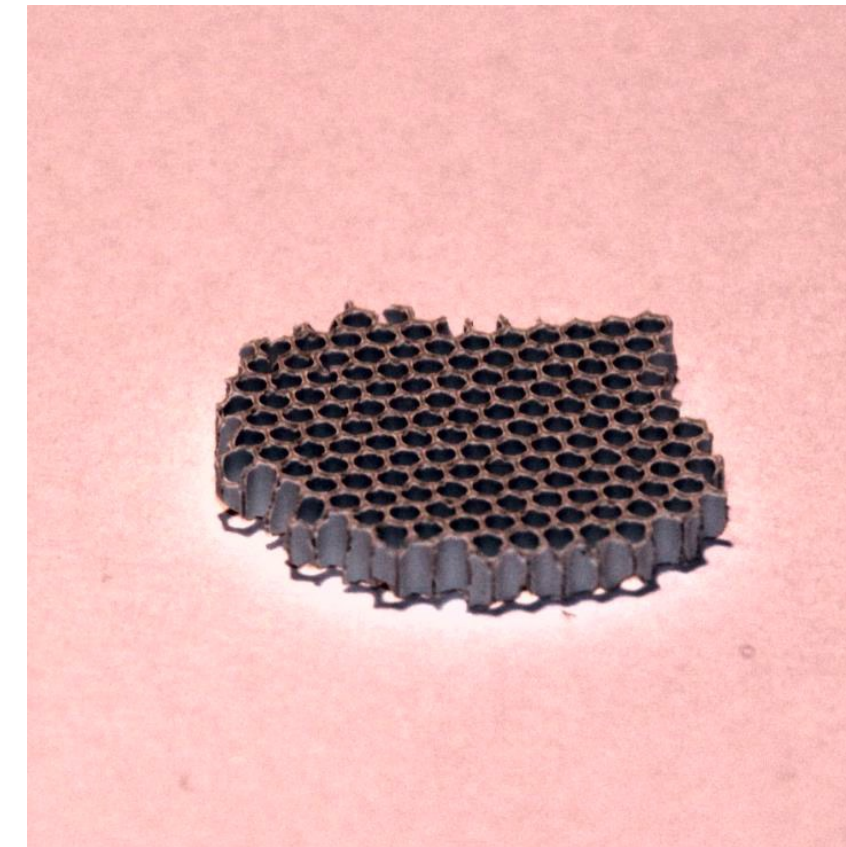
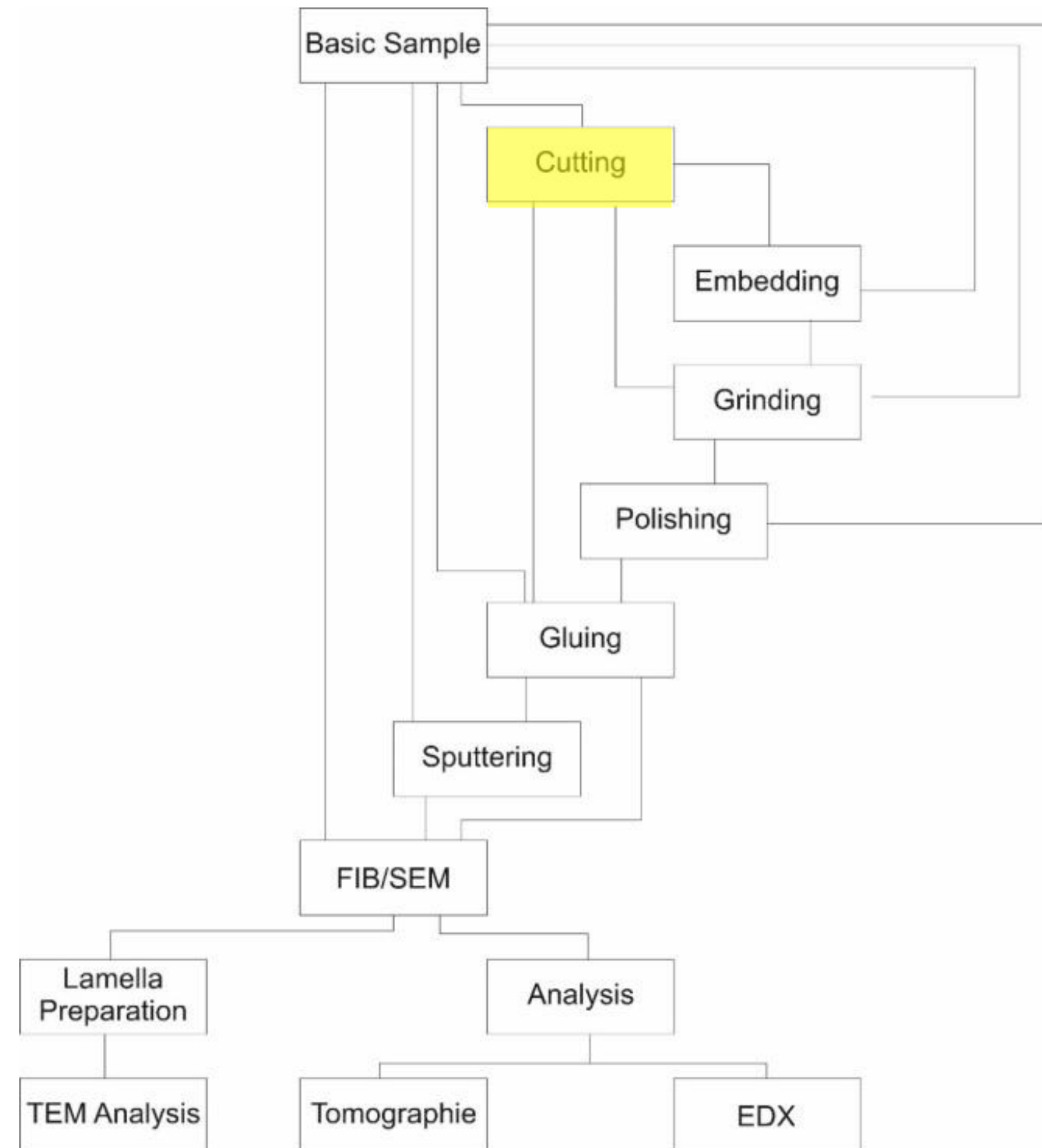
# What can be described?



- Material of the initial sample (a.k.a. precursor)
- Type of coating
- Type of carrier
- ...

*M. Mail, Workflow for TEM lamella preparation – and more (2021)*

# What can be described?

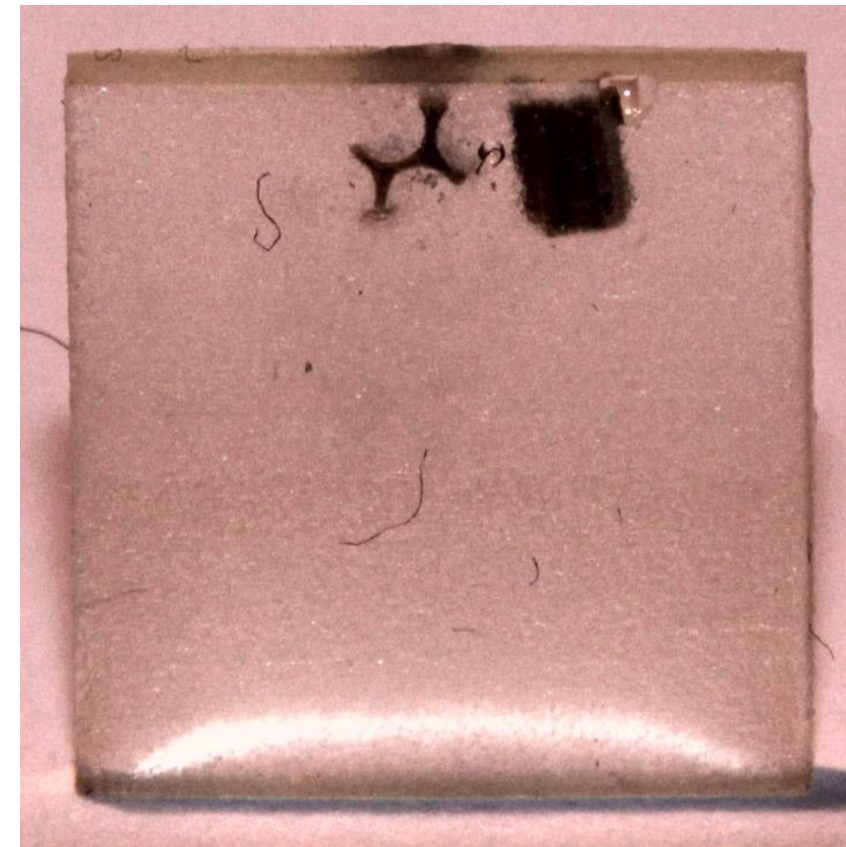
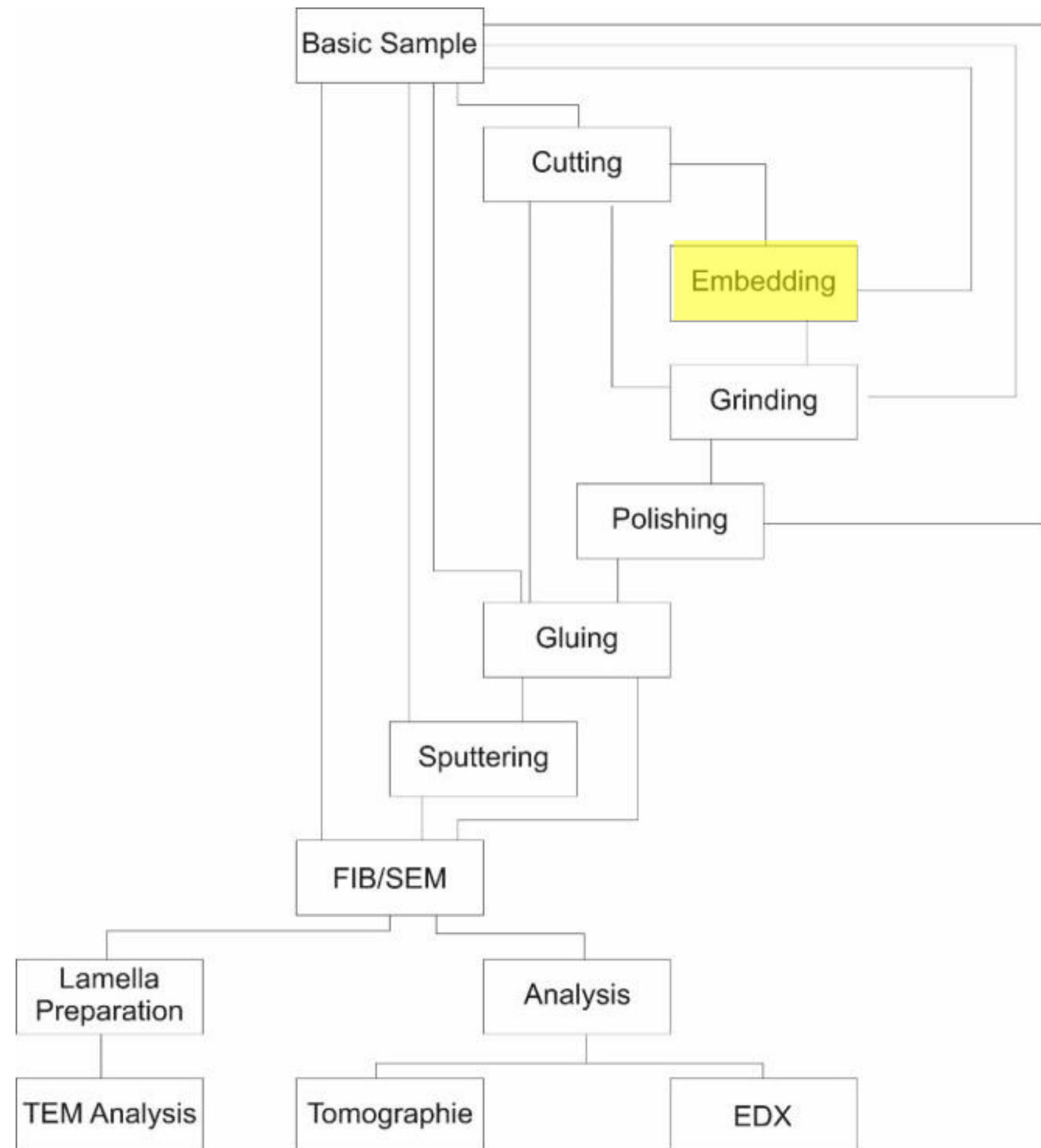


- Cutting parameters (e.g. sawing speed, etc.)
- Cooling
- If so: cooling agent, ...

*M. Mail, Workflow for TEM lamella preparation – and more (2021)*



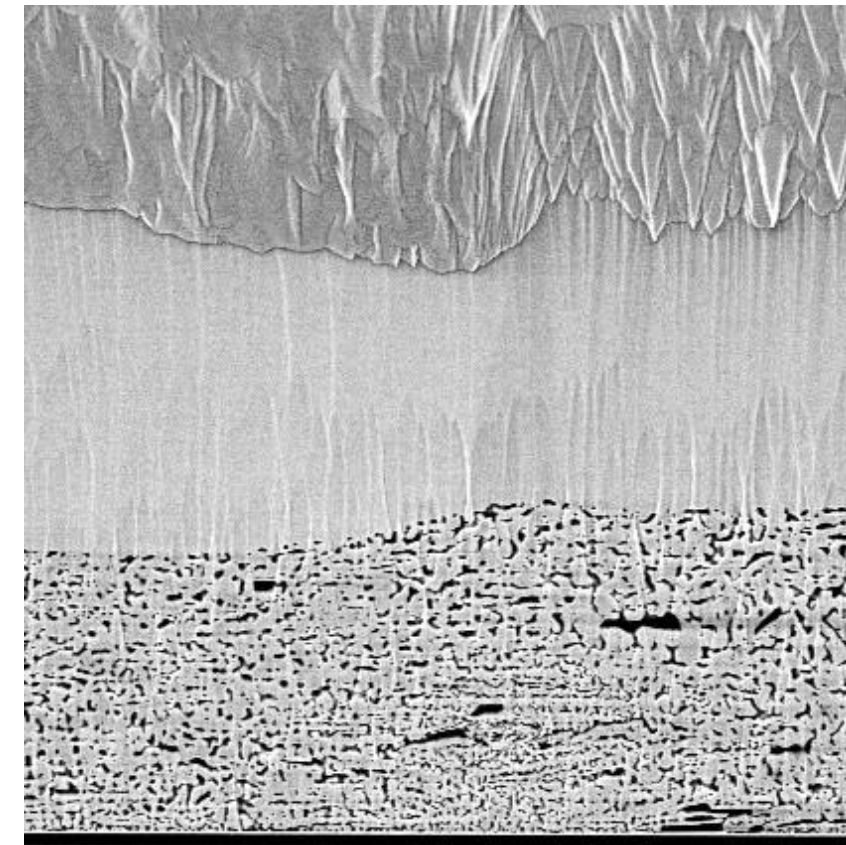
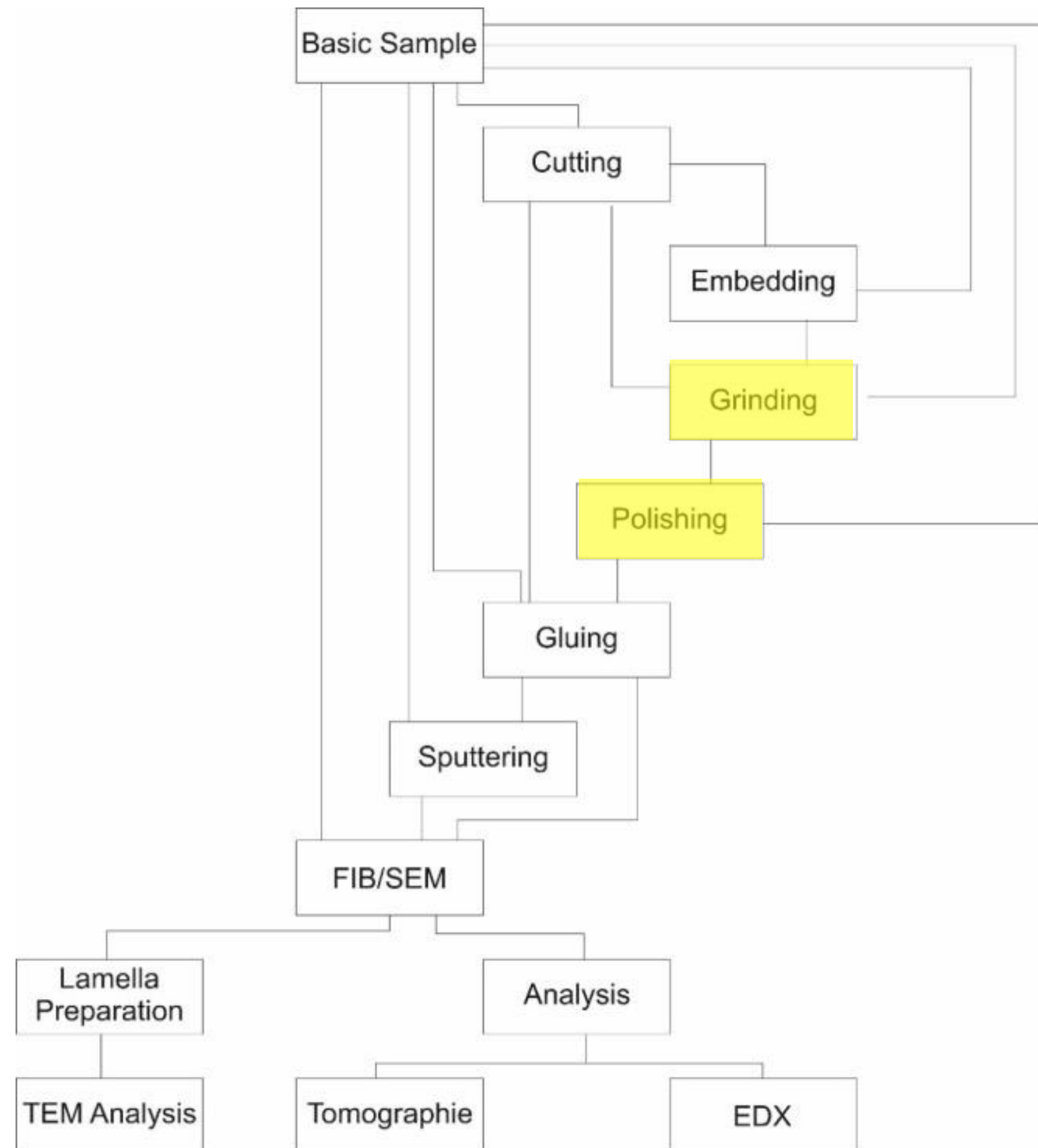
# What can be described?



- Embedding material
- Temperature
- Pressure
- ...

*M. Mail, Workflow for TEM lamella preparation – and more (2021)*

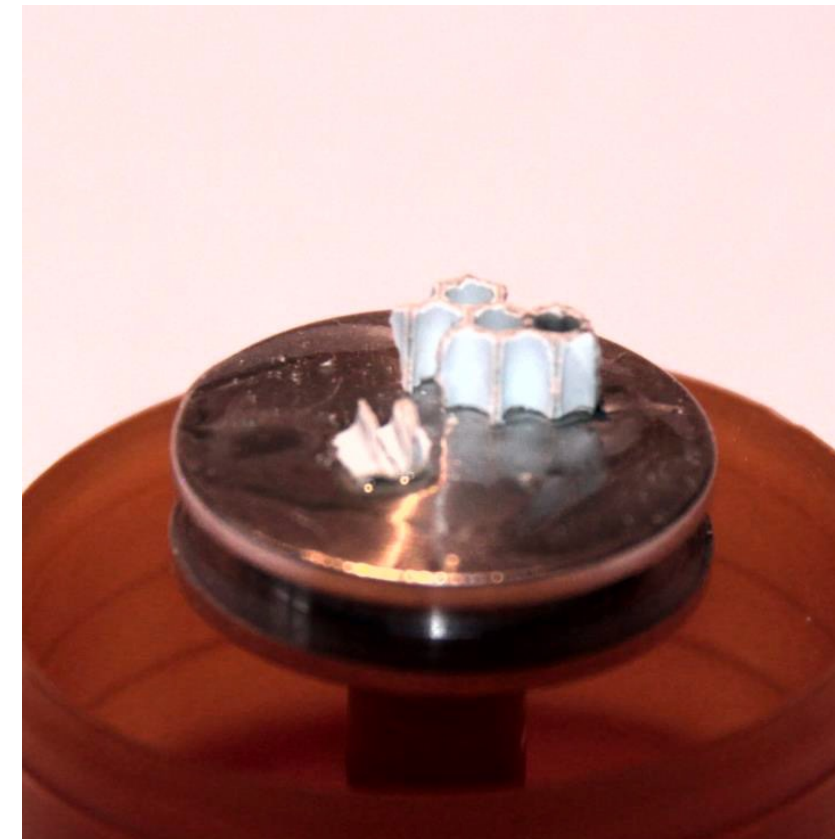
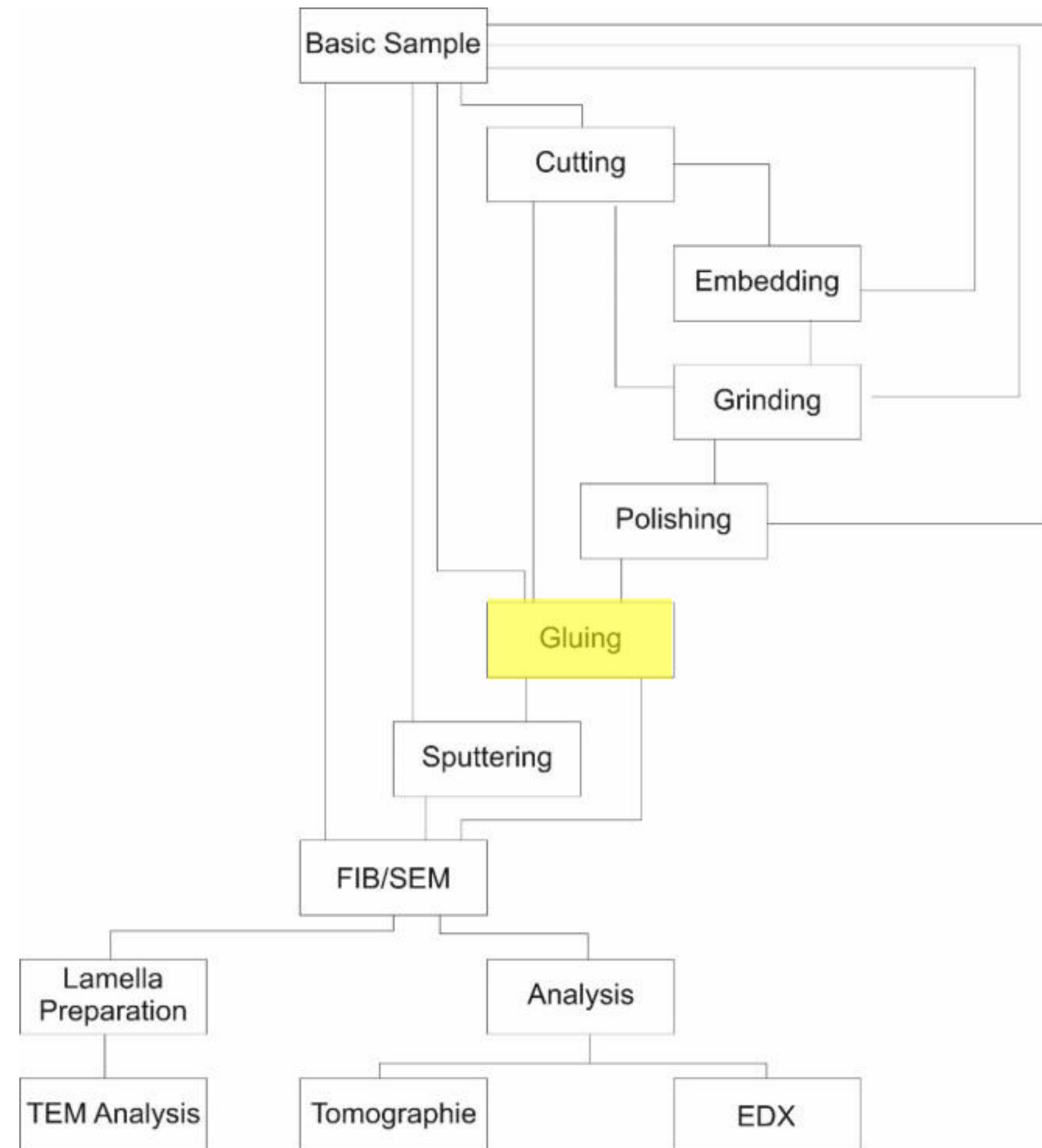
# What can be described?



- Type of sandpaper
- Rotation speed
- Water
- Polishing steps
- ...

*M. Mail, Workflow for TEM lamella preparation – and more (2021)*

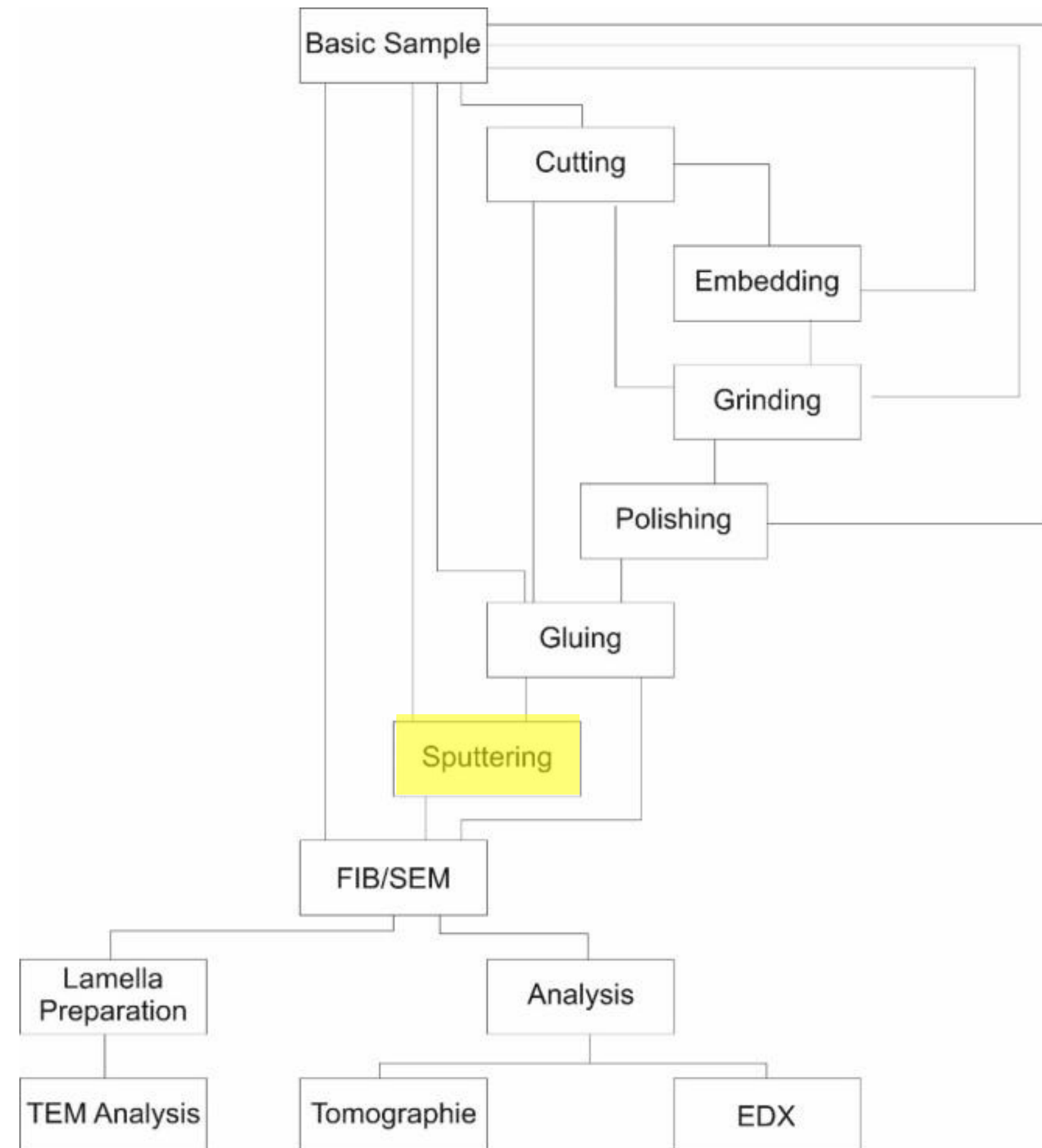
# What can be described?



- Type of glue used
- Position of the sample
- ...

*M. Mail, Workflow for TEM lamella preparation – and more (2021)*

# What can be described?



- Position on the holder
- Sputter material
- Sputter parameters (current, pressure, ...)
- ...

*M. Mail, Workflow for TEM lamella preparation – and more (2021)*

# How to describe data?

- **Metadata:** descriptive information about data (data describing data)

I told you, metadata  
is data!

It's so much more.



# Data vs metadata

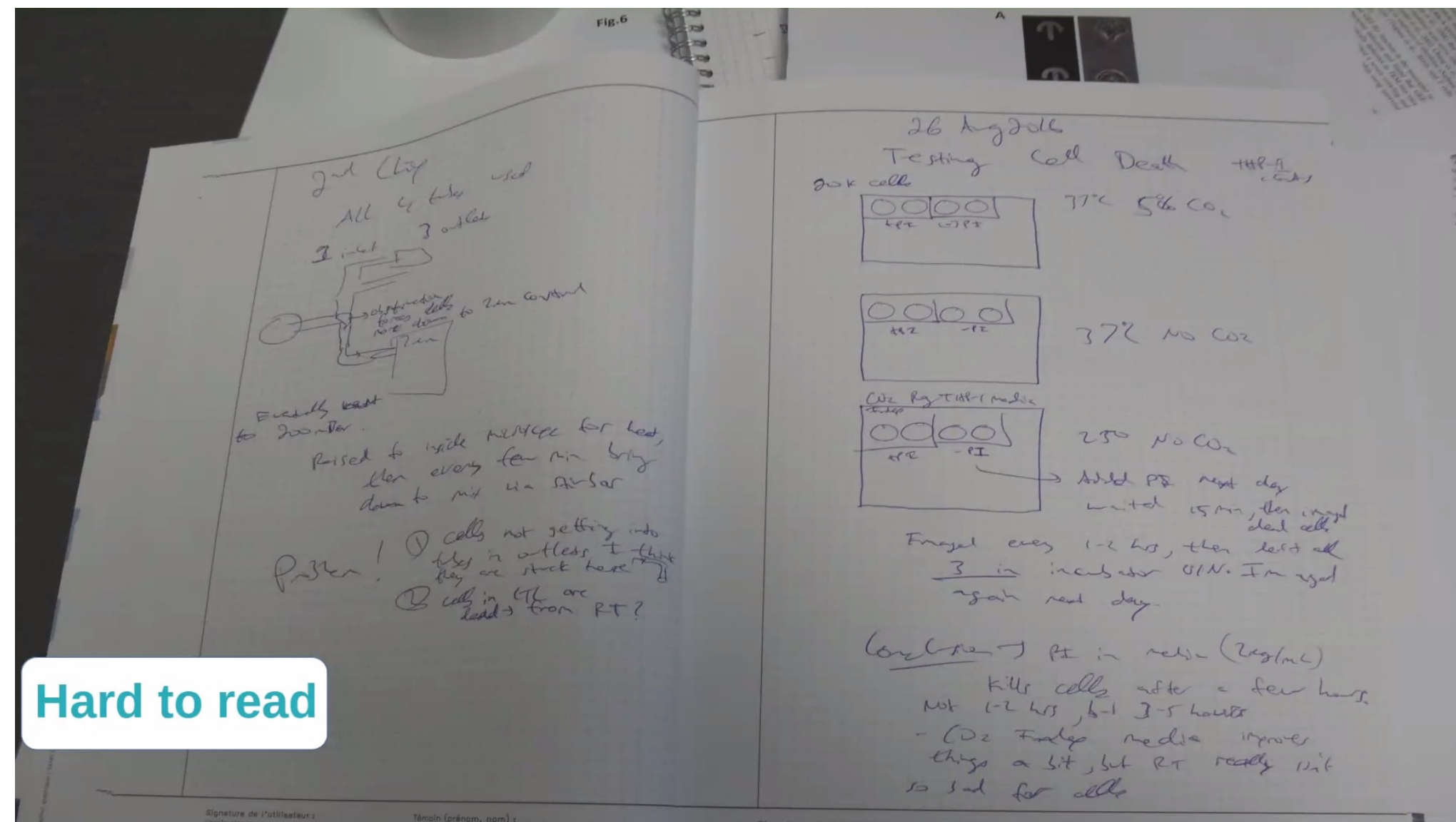
	Data	Metadata
<b>Nature and content</b>	<b>Raw facts</b> , measurements, observations	Information <b>providing context</b> and attributes
<b>Usage</b>	Analysis, decision-making, <b>research</b>	Data management, discovery, <b>interpretation</b>
<b>Representation</b>	Mostly <b>unstructured</b>	<b>Structured</b>
<b>Relationship</b>	Independent and <b>stand-alone</b>	<b>Linked</b> to data/other metadata
<b>Purpose and function</b>	Primary <b>source of information</b>	<b>Supporting framework</b> for organization, management, interpretation

# How to describe data?

- **Metadata:** descriptive information about data (data describing data)
  - **Administrative metadata:** facilitates the management of resources. It can include elements such as version, rights, and licence
  - **Scientific metadata:** enables discovery, identification, and reproducibility of resources. It usually includes the domain-specific attributes.

# Lab Notebooks

- Current efforts of storing metadata are... less than ideal



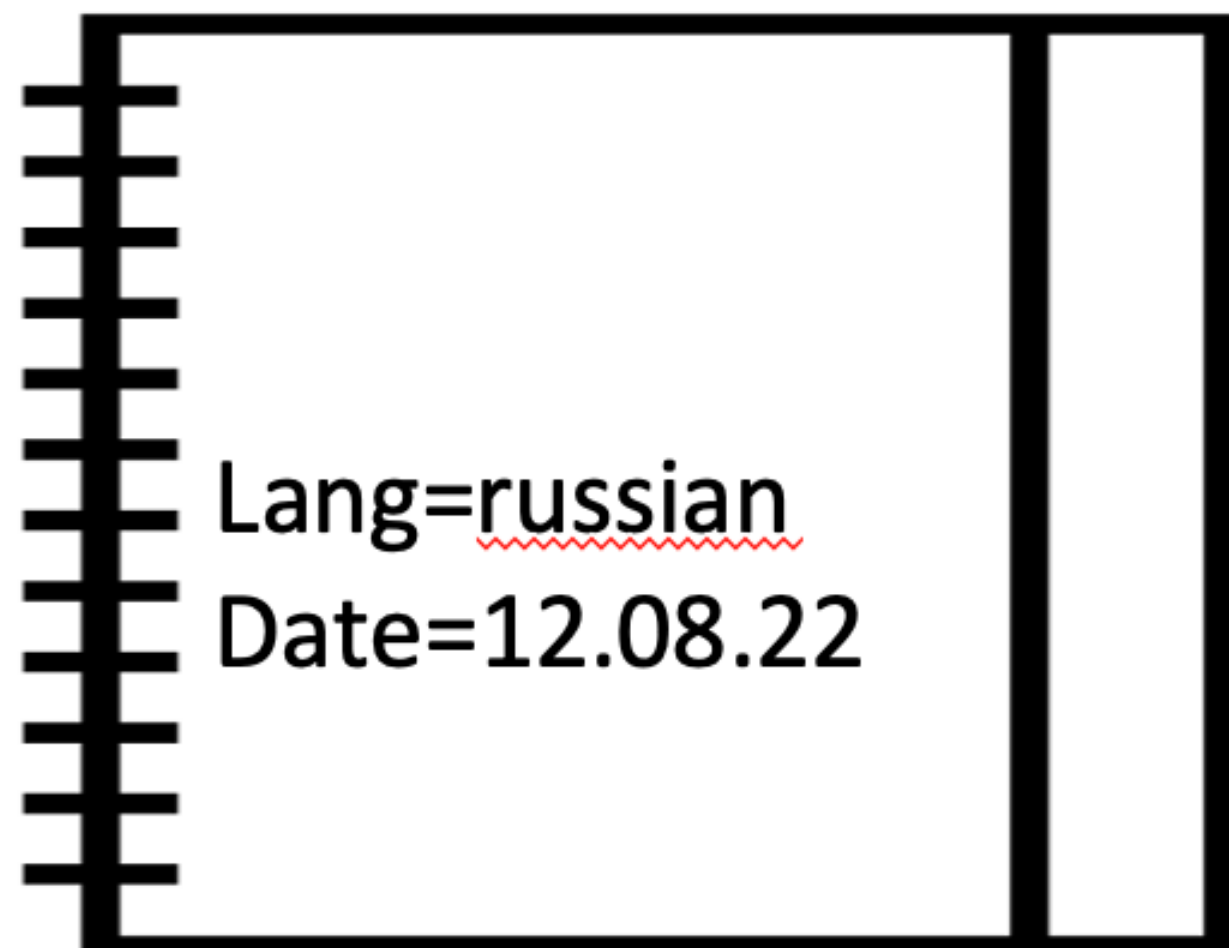
*N. Carpi, eLabFTW (2022)*



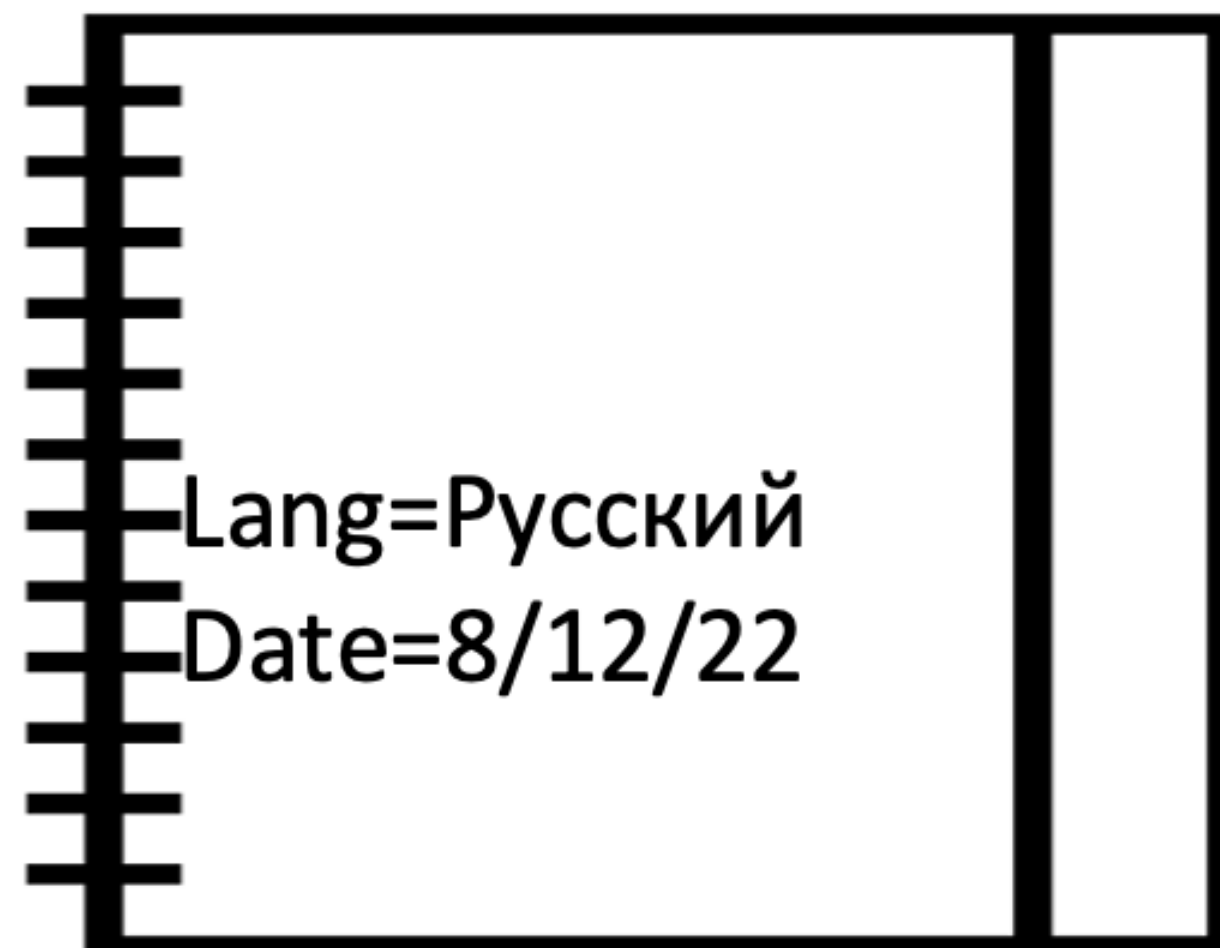
# Electronic Lab Notebooks (ELNs)

- Current efforts of storing metadata are... less than ideal

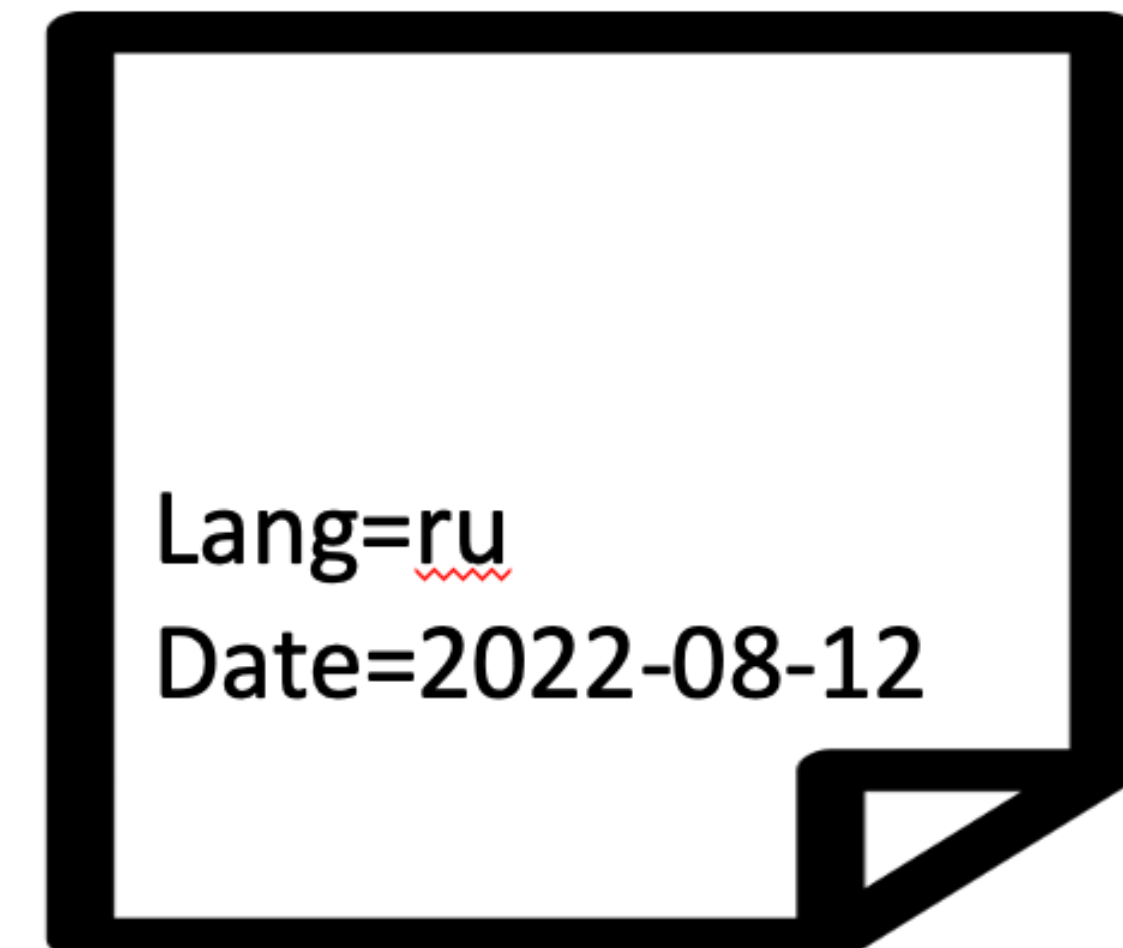
Researcher 1's notebook



Researcher 2's notebook



Researcher 3 using  
ISO standards



*E. Vitali, The MetaStore and its Use Case in NFDI-MatWerk (2022)*

# How to describe data?

- **Metadata document:** digital, structured description about a data resource

```
{
  "phaseOfMatter": "solid"
}
```

- **Metadata schema:** outline of the overall structure of the metadata (elements, value types, rules, ...)

```
"phaseOfMatter": {
  "description": "(Recommended) -",
  "type": "string",
  "enum": [
    "not applicable",
    "solid - definition: https://",
    "liquid - definition: https://",
    "gas - definition: https://e",
    "plasma - definition: https://",
    "mixture - definition: https://"
  ]
},
```

# Advantages of using metadata schemas

- Same parameters for all data, “standardized” description
- Structured metadata can be interpreted (also by machines)
- Data can be found based on their properties
- Data can be compared and/or exchanged
- Results can be reproduced



# Contact us

Rossella Aversa – [rossella.aversa@kit.edu](mailto:rossella.aversa@kit.edu)

[www.nffa.eu](http://www.nffa.eu)  
[secretariat@nffa.eu](mailto:secretariat@nffa.eu)

# Acknowledgements

- The Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under the National Research Data Infrastructure – NFDI 38/1 – project number 460247524
- The Joint Laboratory Model and Data driven Materials Characterization (JL MDMC), a cross-centre platform of the Helmholtz Association
- NFFA-Europe-Pilot (EU H2020 – n. 101007417)
- The research program “Engineering Digital Futures” of the Helmholtz Association of German Research Centers
- The Helmholtz Metadata Collaboration Platform