ELN@FLI

Stage 0 - From market survey to performance specification

processes. This is why improvements, adjustments, and new ideas concerning the framework for GSP at FLI are constantly discussed. With the development of this GSP Concept, further improvements of GSP measures at FLI can be developed and ascertained also in the future.

GSP Measures

- ▼ Ombudsperson
- ▼ Weekly Institute Seminar
- ▼ Leibniz Graduate School on Aging - LGSA
- ▼ GSP Regulations in Work Contracts
- ▼ Compulsory GSP Trainings
- ▼ Electronic Lab Notebook - ELN
- ▼ Primary Data Management
- ▼ Guideline Quality Control
- ▼ Rules for Archiving of Publication-Relevant Data

Good Scientific Practice ▼

GSP Measures at FLI

Ombudsperson

Publications

Initial feature list

- questionnaire - lack of experience

Experiences

- someELN@jena
  - 10 years
  - regular user training
  - **≈ 30%**
- bring the people using it! (Y. R., MPIIB)
- portfolio of functionalities? (I. P., CSB)

Essence: Usability

- ➔ **Ease of use most important!**
- ➔ follow-up costs (BAuA)

Initial feature list

- questionnaire - lack of experience

Experiences

- someELN@jena
  - 10 years
  - regular user training
  - **≈ 30%**
- bring the people using it! (Y. R., MPIIB)
- portfolio of functionalities? (I. P., CSB)

Essence: Usability

- ➔ **Ease of use most important!**
- ➔ follow-up costs (BAuA)

Initial feature list

- questionnaire - lack of experience

Experiences

- someELN@jena
  - 10 years
  - regular user training
  - ≈ 30%
- bring the people using it! (Y. R., MPIIB)
- portfolio of functionalities? (I. P., CSB)

Essence: Usability

- ➔ **Ease of use most important!**
- ➔ follow-up costs (BAuA)

Budget

Core requirements

- **on premises** (no cloud!)
- **Web based** (only HTML5 browser)
- **Comprehensive Export**
- **Links to external files**
- Links to internal resources
- Search
- Templates, Reuse
- Collaboration
- GSP (Users+Groups • Rights • Versions) ➜ FDA 21 CFR part 11
- writable API
- Inline annotation of images • Freehand drawing
- **Usability**

Market

- LiMSwiki.org: 87 • The Gurdon Institute: 28 • LabsExplorer: 27 • Harvard Biomedical Data Management: 26 • Labii: 24

- + 10%

- OSS – sustainability? usability?

Market

- LiMSwiki.org: 87 • The Gurdon Institute: 28 • LabsExplorer: 27 • Harvard Biomedical Data Management: 26 • Labii: 24

- + 10%

- OSS – sustainability? usability?
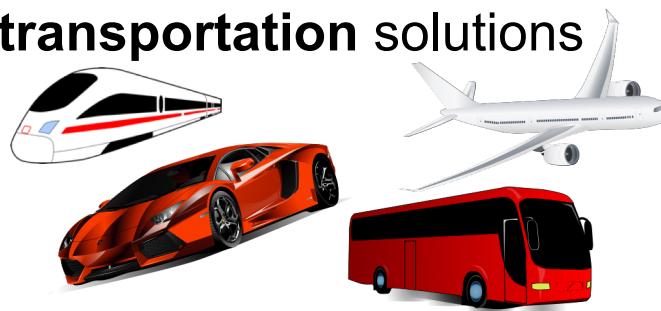
Aquire ELN

lab software solutions
- Abbott Informatics • Agilent • AgiLab
- BIOVIA • Dassault • Dotmatics • IDBS
- DeltaSoft • LabVantage • LabWare
- PerkinElmer • ThermoFisher • Waters
- …

Aquire transportation tool

transportation solutions

Type not targeted
- database centered
- structured data

- comprehensive

- specific
  ➔ needs adaption

- expensive?
- GUI on DB?

Aimed type
- document centered
  "scribbling" software
  + collaboration + GSP + RDM

- lightweight – standalone

- generic
  ➔ straightaway to use

- budget
- easy to use

Shortlisting

➔ **not** appropriate: webinars, demos, tutorials, …

➔ Test in practice!

(Easy access - cloud, VMs)

Shortlisting

     ➔ **not** appropriate: webinars, demos, tutorials, …

     ➔ Test in practice!

        (Easy access - cloud, VMs)

- ranked list $\lesssim$ dozen

User trial

- 3 ELNs   (7 weeks)   : 15 testers
- countable feedback from   :   6

Shortlisting

➜ **not** appropriate: webinars, demos, tutorials, …

➜ Test in practice!

(Easy access - cloud, VMs)

- ranked list $\lesssim$ dozen

User trial

- 3 ELNs   (7 weeks)   : 15 testers
- countable feedback from   :   6

ELN	all testers: 6 (points out from 10)	ranking	"power testers": 2 (time > 10h / ELN)	ranking
ELN 1	7.3	1	7.6	1
ELN 2	5.9	3	6.6	2
ELN 3	6.0	2	5.6	3

1. **Evaluating trial**

- Test environment ≠ real life (cloud ≠ on premises)
  - editing imported office files
  - data publishing (DOI)
  - Inventory
- initial judgements changed ≳ 1h
- "power testers" ➔ Achilles' heels (≫ sum of rush judgments)

2. **ELN suitability**

- all tested ELNs suitable (glitches ELN 3)
- ELN 4 – suspected also to fit

3. **Requirement specification**

- final drop of "must haves"
- phrasing ➔ practical needs
  - unambigous • exhaustive (§ 7 Abs. 1 VOL/A)

1. Evaluating trial

- Test environment ≠ real life (cloud ≠ on premises)
  - edit imported office files
  - data publishing (DOI)
  - LabInventory
- initial judgements changed ≳ 1h
- "power testers" ➔ Achilles' heels (≫ sum of rush judgments)

2. ELN suitability

- all tested ELNs suitable (glitches ELN 3)
- ELN 4 – suspected also to fit

3. Requirement specification

- final drop of "must haves"
- phrasing ➔ practical needs
  - unambigous • exhaustive (§ 7 Abs. 1 VOL/A)

1. **Evaluating trial**

- Test environment ≠ real life (cloud ≠ on premises)
  - edit imported office files
  - data publishing (DOI)
  - LabInventory
- initial judgements changed ≳ 1h
- "power testers" ➜ Achilles' heels (≫ sum of rush judgments)

2. **ELN suitability**

- all tested ELNs suitable (glitches ELN 3)
- ELN 4 – suspected also to fit

3. **Requirement specification**

- final drop of "must haves"
- phrasing ➜ practical needs
  - unambigous • exhaustive (§ 7 Abs. 1 VOL/A)

	old prize (subscr. per seat and year)	new prize (subscr. per seat and year)	raise	comment
ELN X	$ 120	$ 600 (2016)	400 %	sic!
ELN 3 „unlimited server"	€ 1390 /institution/year (26th July 2015)	not available anymore	-	vanished from archive.org
ELN 3 (industry)	258 € (June 2017)	540 € (July 2017)	109 %	huge framework contracts 10/2016 04/2017 (no tender)
ELN 3 (academic)	120 € (June 2017)	180 € (July 2017)	50 %	

Commercial ELNs

- No open data format for ELN

- No data transfer between different ELNs

       ➔   high switching „costs"
       ➔   + obligation for continuous documentation

Commercial ELNs

- No open data format for ELN

- No data transfer between different ELNs

      ➔   high switching „costs"

      ➔   + obligation for continuous documentation

= vendor lock-in effect!

Commercial ELNs

- No open data format for ELN

- No data transfer between different ELNs

➔ high switching „costs"

➔ + obligation for continuous documentation

= vendor lock-in effect!

unavoidable, but

avoid <u>exploitation</u>

License & contracting

➔ keep on eye level with vendor

Get the data out!

➔ Export capabilities

Licensing and contracting

1. end of subscription period ➔ **unlimited (read-only) access!**

2. purchase of licenses. no budget? ➔ **hire** purchase

3. tiered prices

4. long period of notice ➔ time to switch

5. framework contract - fixed <u>prices</u>

   - long period

   - auto-prolonguation

6. "exit" ➔ "prompt adaptions to future changes"

   - legislation

   - requirements of funding organizations

   - technical requirements of GSP

Licensing and contracting

1. end of subscription period ➜ **unlimited (read-only) access!**
2. purchase of licenses. no budget? ➜ **hire** purchase
3. tiered prices
4. long period of notice ➜ time to switch
5. framework contract - fixed <u>prices</u>
   - long period
   - auto-prolonguation
6. "exit" ➜ "prompt adaptions to future changes"
   - legislation
   - requirements of funding organizations
   - technical requirements of GSP

Licensing and contracting

1. end of subscription period ➜ **unlimited (read-only) access!**
2. purchase of licenses. no budget? ➜ **hire** purchase
3. tiered prices
4. long period of notice ➜ time to switch
5. framework contract - fixed <u>prices</u>
   - long period
   - auto-prolonguation
6. "exit" ➜ "prompt adaptions to future changes"
   - legislation
   - requirements of funding organizations
   - technical requirements of GSP

Licensing and contracting

1. end of subscription period ➜ **unlimited (read-only) access!**
2. purchase of licenses. no budget? ➜ **hire** purchase
3. tiered prices
4. long period of notice ➜ time to switch
5. framework contract - fixed <u>prices</u>

   - long period
   - auto-prolonguation

6. "exit" ➜ "prompt adaptions to future changes"

   - legislation
   - requirements of funding organizations
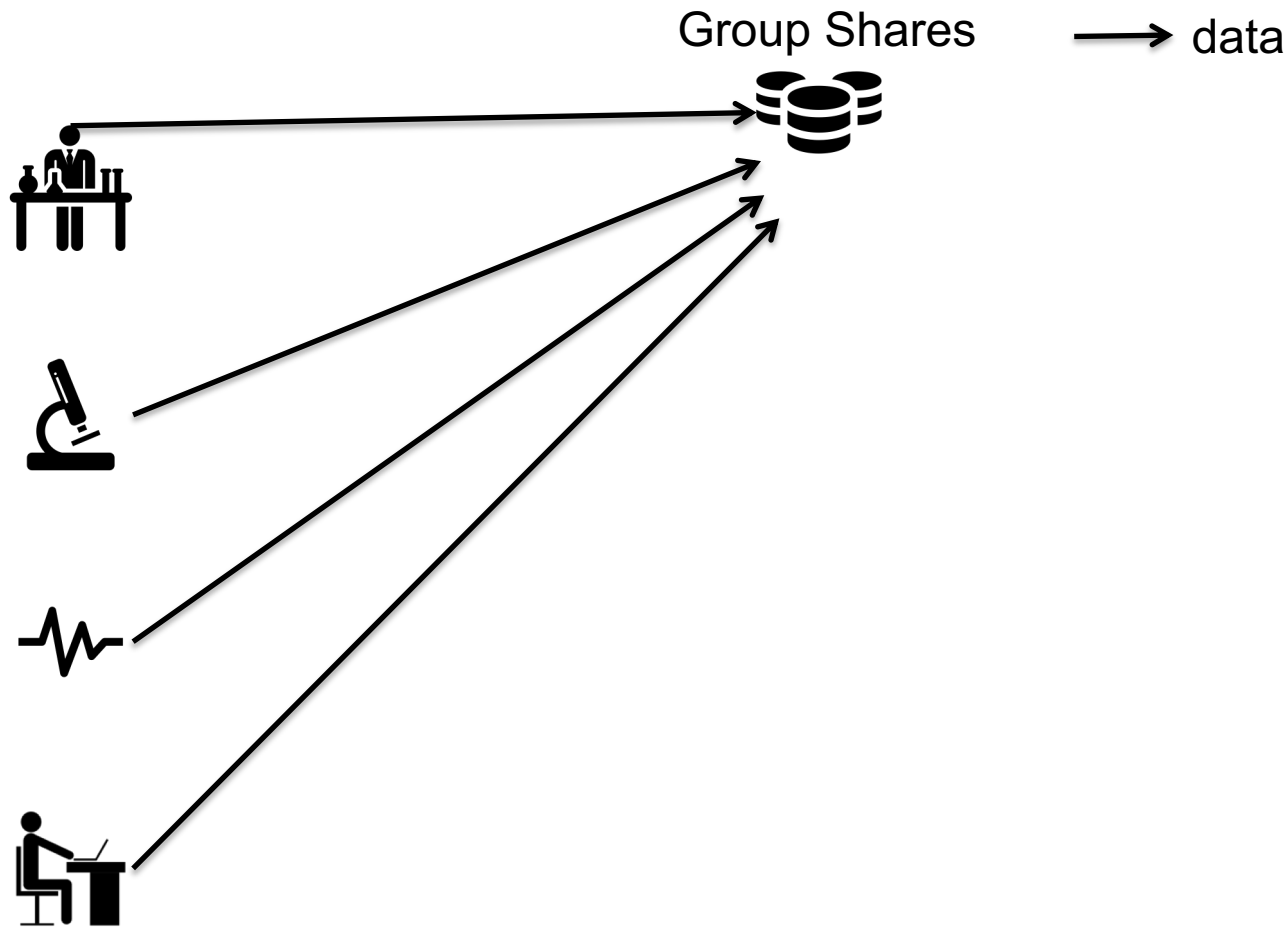   - technical requirements of GSP

Licensing and contracting

1. end of subscription period ➜ **unlimited (read-only) access!**
2. purchase of licenses. no budget? ➜ **hire** purchase
3. tiered prices
4. long period of notice ➜ time to switch
5. framework contract - fixed <u>prices</u>
   - long period
   - auto-prolonguation
6. "exit" ➜ "prompt adaptions to future changes"
   - legislation
   - requirements of funding organizations
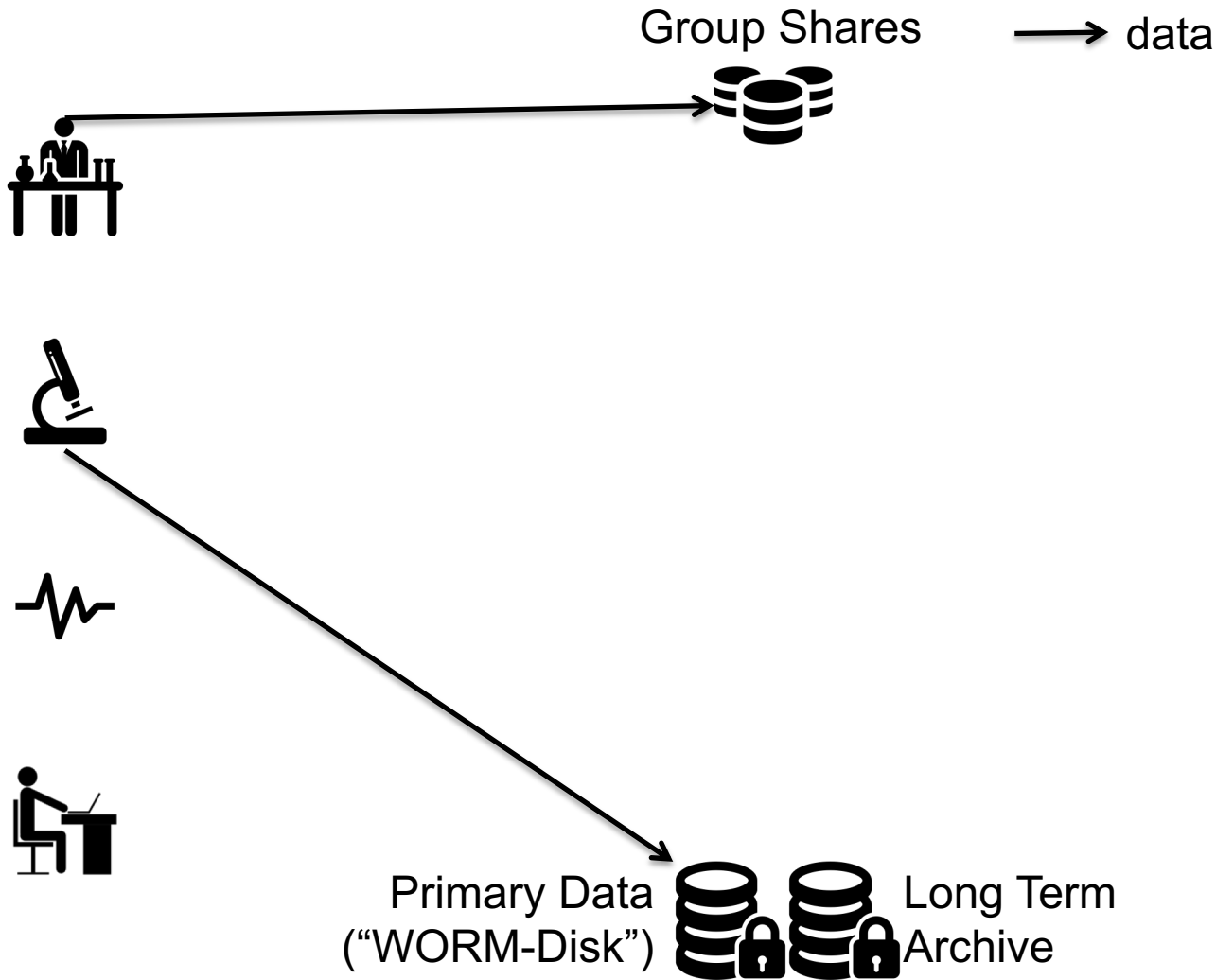   - technical requirements of GSP

Licensing and contracting

1. end of subscription period ➔ **unlimited (read-only) access!**
2. purchase of licenses. no budget? ➔ **hire** purchase
3. tiered prices
4. long period of notice ➔ time to switch
5. framework contract - fixed <u>prices</u>
   - long period
   - auto-prolonguation
6. "exit" ➔ "prompt adaptions to future changes"
   - legislation
   - requirements of funding organizations
   - technical requirements of GSP

Current situation (simplified)

Group Shares

→ data

Possible future scenario (RDM currently under development)

Group Shares → data

Primary Data ("WORM-Disk")   Long Term Archive

ELN data capture

Group Shares

data outside ELN

data capture

ELN

Primary Data
("WORM-Disk")

Long Term
Archive
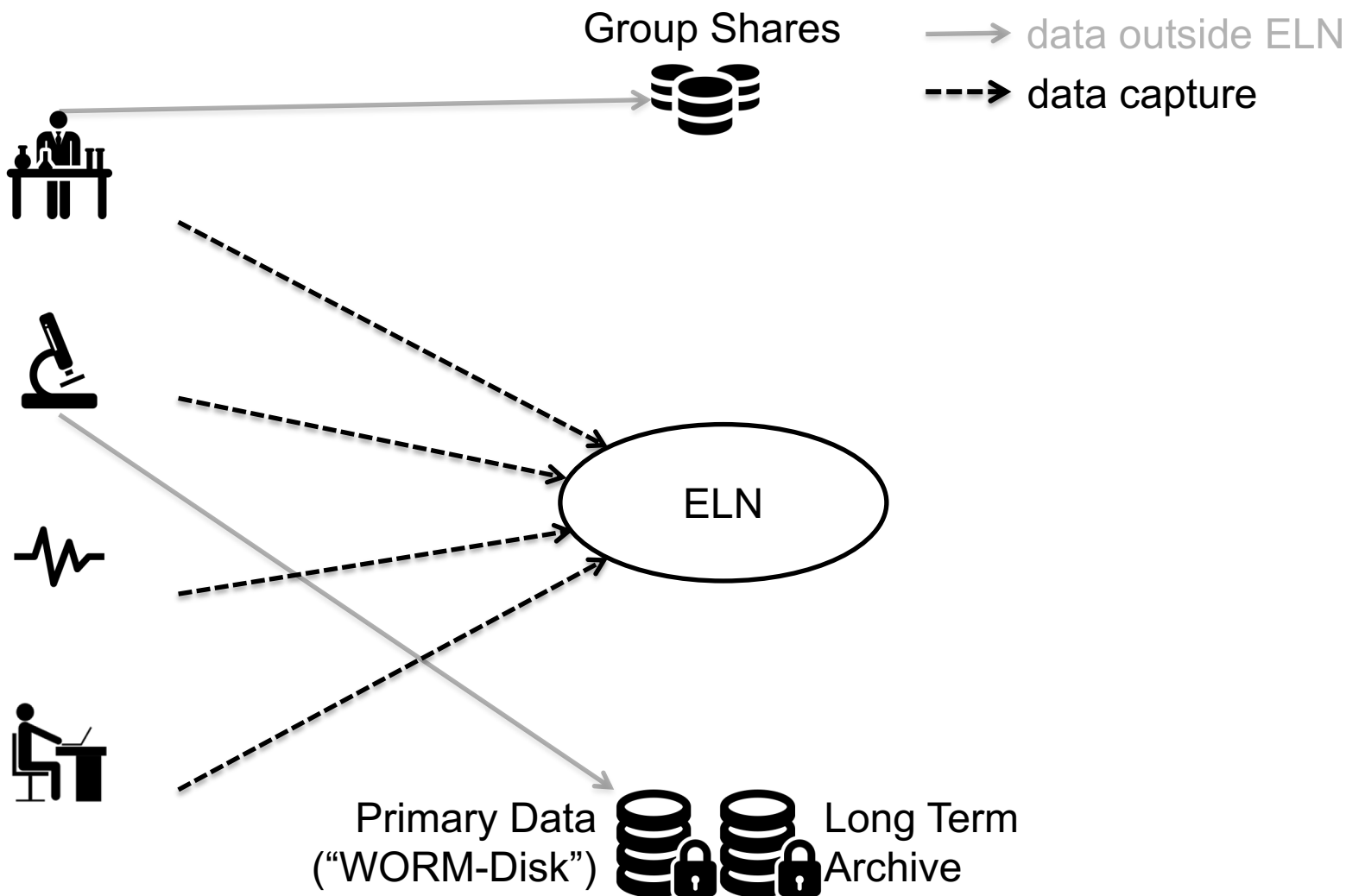
Masks: collect metadata as soon as data captured, provide structure

Group Shares

data outside ELN

data & metadata capture

masks provided by ELN

Masks

ELN

Primary Data ("WORM-Disk")

Long Term Archive

ELN – another data silo?

Group Shares

data outside ELN

data & metadata capture

masks provided by ELN

Masks

ELN

Primary Data ("WORM-Disk")

Long Term Archive
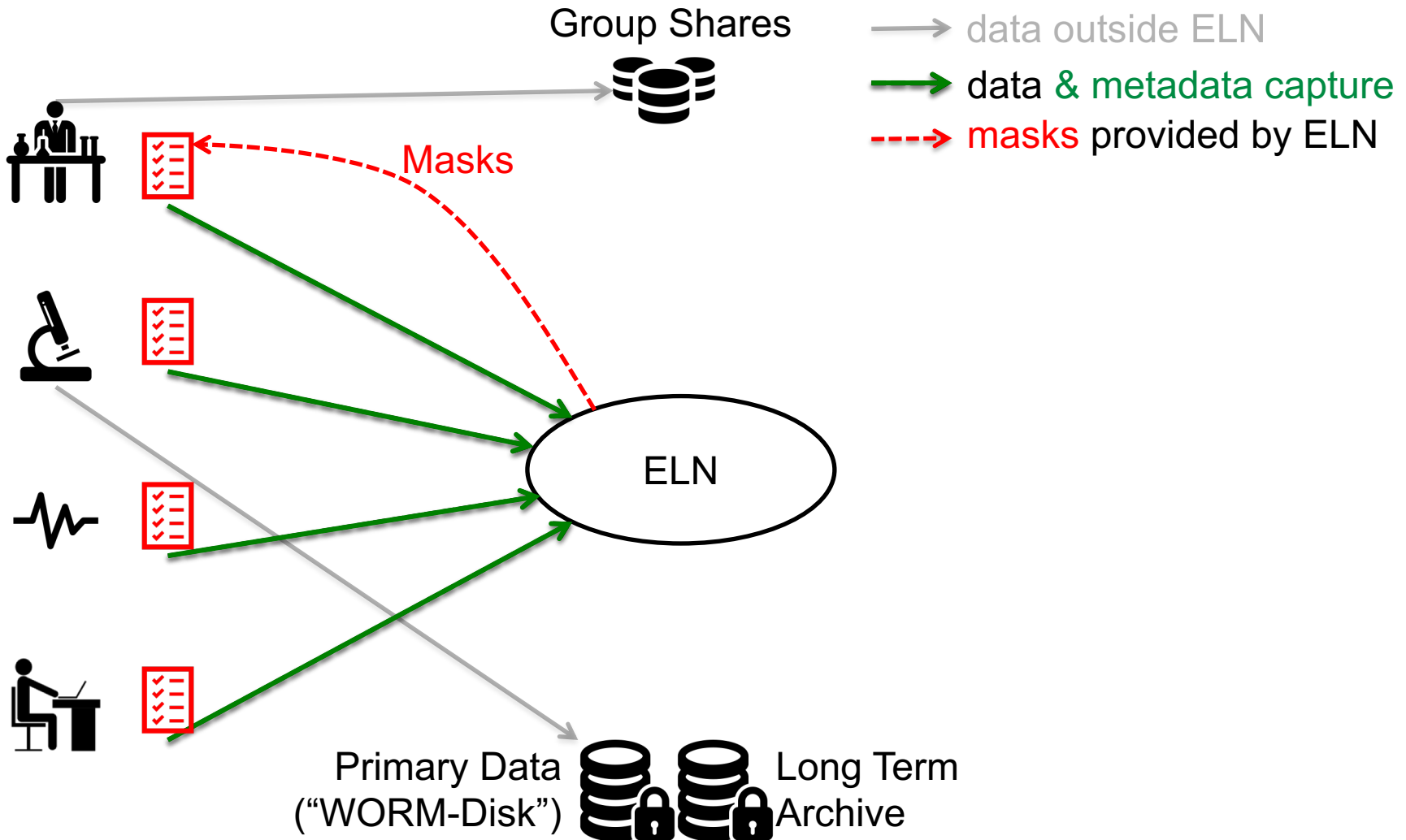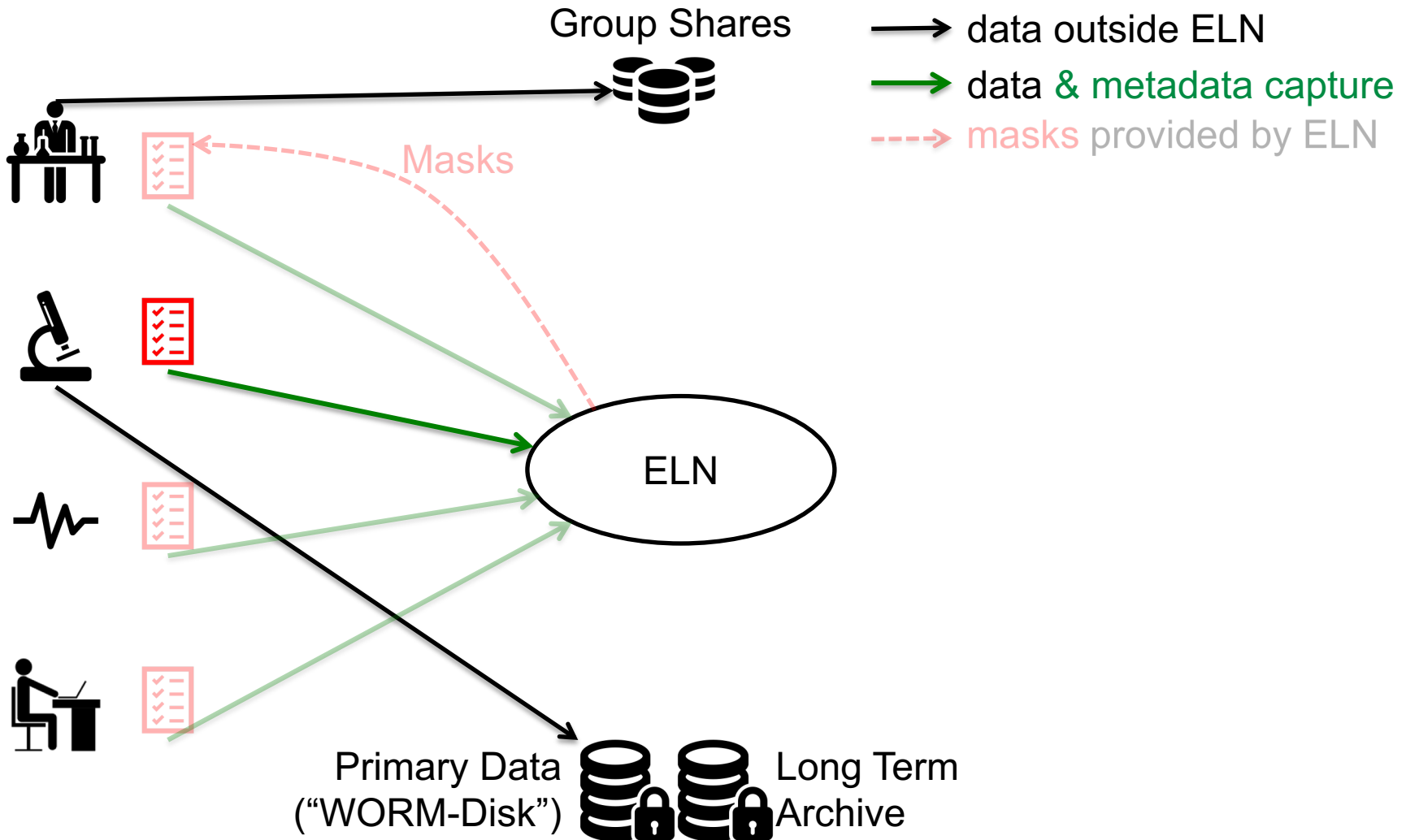
Links tells it: where is the file I created in this experiment?

Group Shares

data outside ELN

data & metadata capture

masks provided by ELN

external files linked

Masks

Link (break.)

ELN

Link (unbr.)

Primary Data ("WORM-Disk")

Long Term Archive
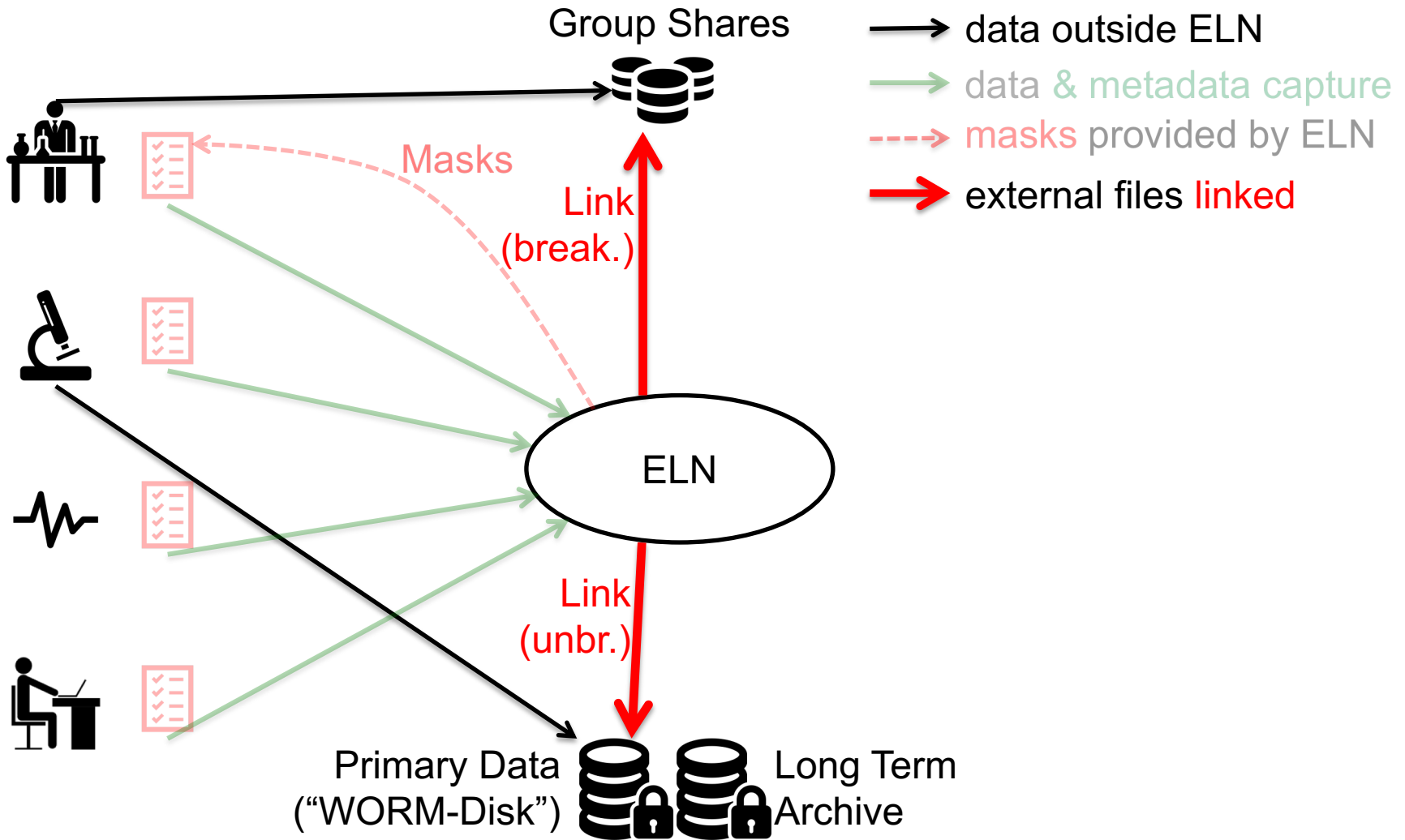
Linking to external files

- linking to local (institutional) network resources
  - SMB/CIFS
  - NFS
  - sftp
- several shares, simultaneous

- referring external files: "extremely important"

- avoid data duplication ➜ RDM

Get it out - Archive (GSP) - Resist vendor lock-in

Group Shares

data outside ELN
data & metadata capture
masks provided by ELN
external files linking
export data & metadata

Masks

Link
(break.)

Export

ELN

Link
(unbr.)

Export
(PDF/A, XML)

Export

Primary Data
("WORM-Disk")
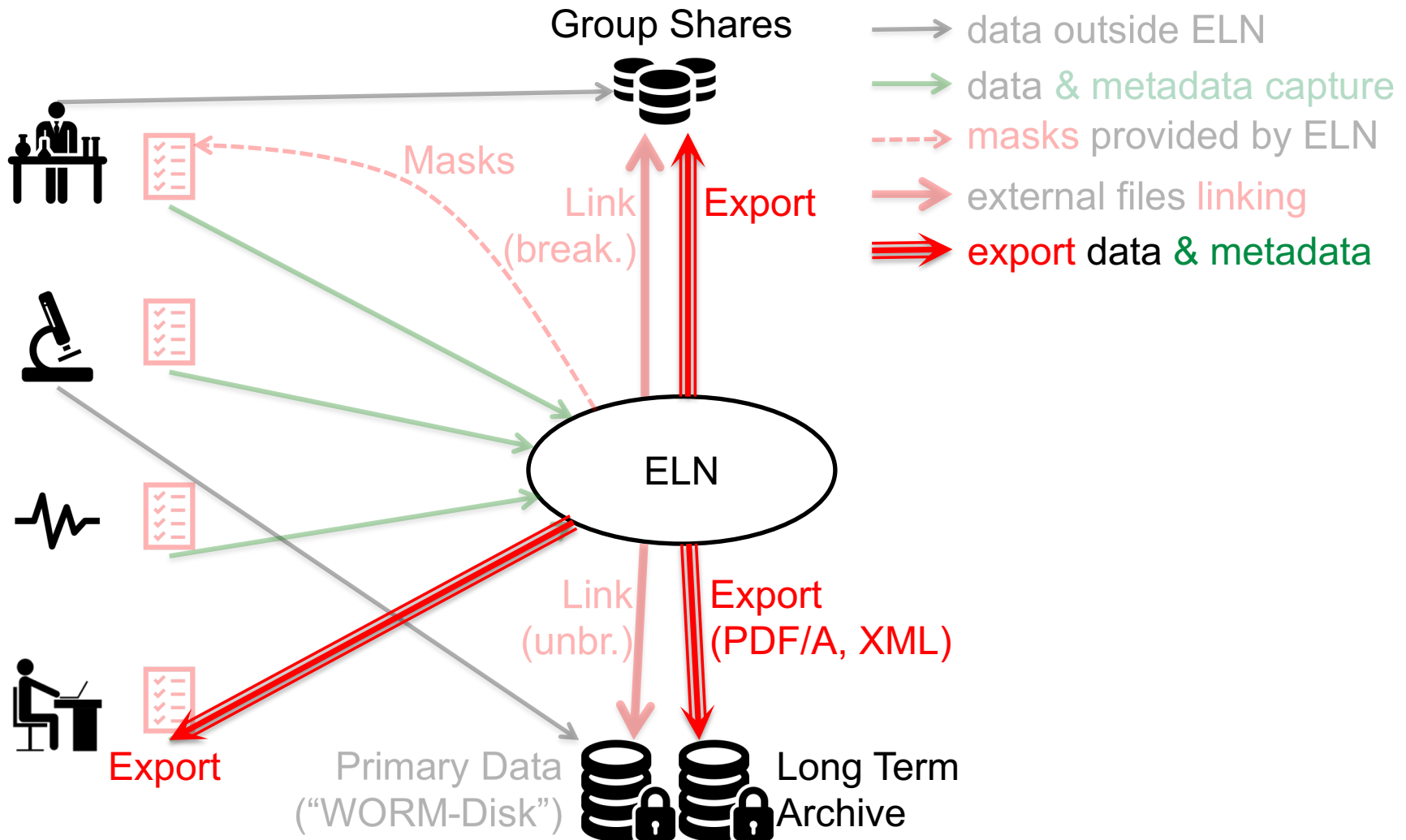
Long Term
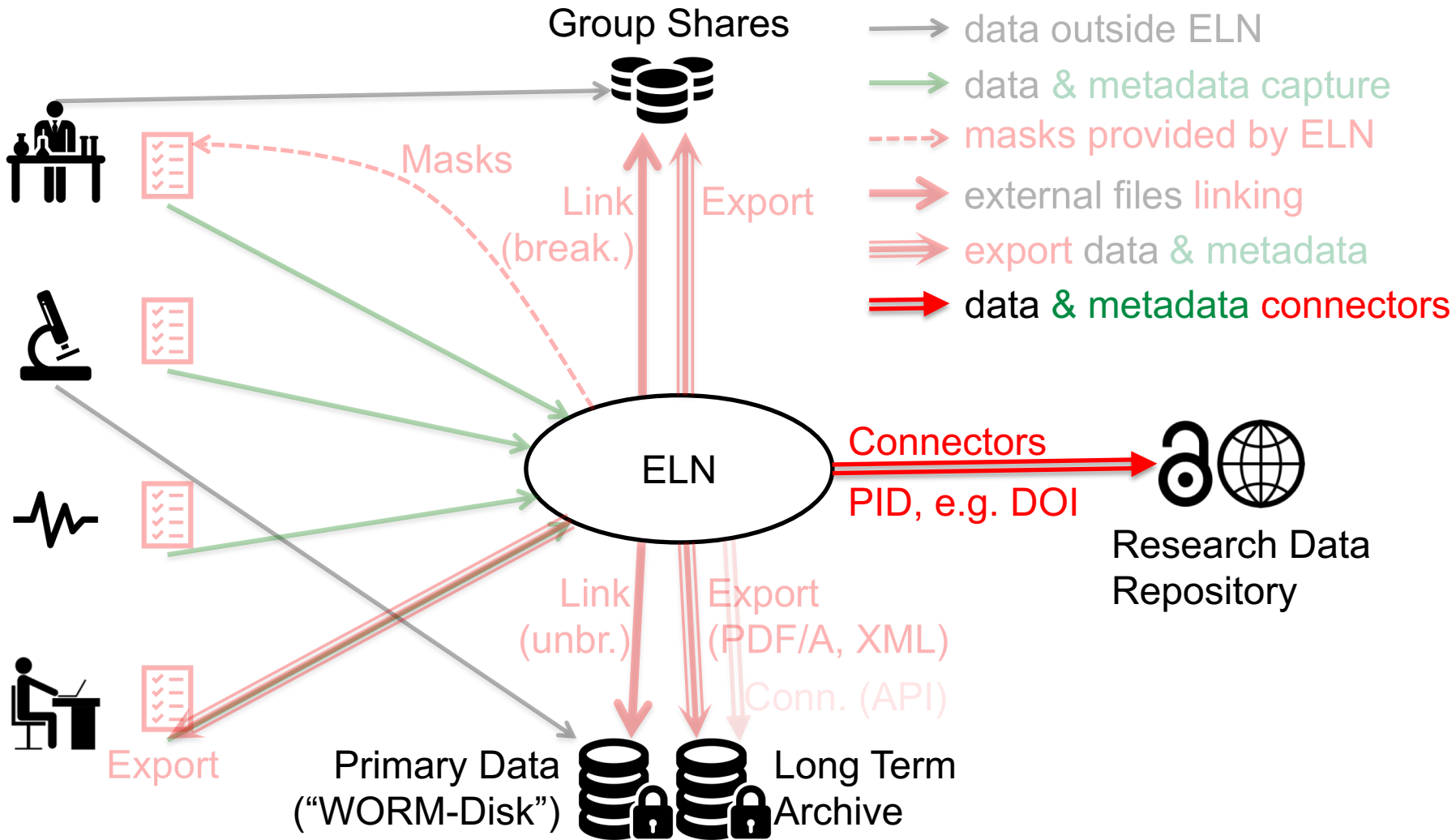Archive

Export functionality

1. Export formats ("how")
   - long term archiving: BSI (M 4.170)
     - humans ➔ PDF/A
     - machines ➔ XML + schema definition XSD (W3C)
   - "offline version"
     - HTML (<u>not</u> for archiving - BSI)
     - preservation of links, unlimited linking depth
   - format allowing for complete reimport
     - preserving the original ELN structure
     - take all my labbooks to my next research position
     - cloud ⇆ on premises
   - single documents ➔ common office formats

Export functionality (continued)

2. Export content selection ("what")
   - single + multiple selection
   - on every granularity level
     – from documents, folders, projects, notebooks, group, ...
     – ... whole ELN content

3. Export performers ("who")
   - Users
   - Group leaders
   - Admins
     – whole content
     – no need for vendor support

Deposit data & metadata: bench ➜ repository

Group Shares

data outside ELN

data & metadata capture

Masks

masks provided by ELN

Link
(break.)

Export

external files linking

export data & metadata

data & metadata connectors

ELN

Connectors

PID, e.g. DOI

Research Data
Repository

Link
(unbr.)

Export
(PDF/A, XML)

Conn. (API)

Export

Primary Data
("WORM-Disk")

Long Term
Archive

Connectors to general purpose repositories

- publish data <u>directly</u> from ELN

- metadata <u>together</u> with data as suppl. mat.

- create PID (DOI, Handle)

Which repo?

1. open source (one of)

    - CKAN

    - ePrints

    - DSpace

    - Dataverse

    - (also: Invenio)

2. institutional

    - Zenodo or Harvard Dataverse

3. commercial

    - Figshare (amongst)

Bilingual

- broadening the competition
- sticking point
- invitation / performance spec. / communication
- contract: German

Next steps

- invitation to bid – mid of June
- Setup: August 2018
- Stage 1: Preparation
  - metadata capture
  - commitments

Thanks to ELN testers

Nicolas Huber

Mara Sannai

Thomas Bates

Danny Schnerwitzki

Omid Omrani

Silke Förste

+ 9 others for valuable feedback

Thanks to many experts

for disclosure of their experiences and insights
by visitations, presentations, workshops, open discussions.

Graphic	Slide(s)	URL
Bus	8	https://openclipart.org/download/24476/Anonymous-Bus1.svg
Race Car	8	https://openclipart.org/download/190844/1392140719.svg
ICE	8	https://openclipart.org/image/2400px/svg_to_png/18947/mbs-ICE-Train.png
Airplane	8	https://openclipart.org/download/28921/sagar-ns-Plane.svg
Roller	9	https://upload.wikimedia.org/wikipedia/commons/8/81/St-4428695-s1000_%281%29.jpg
Storage	26-31,33,36	https://www.flaticon.com/free-icon/server_675480 Icon made by Freepik from www.flaticon.com licensed by Creative Commons BY 3.0
Microscope	26-31,33,36	https://upload.wikimedia.org/wikipedia/commons/6/64/Microscope_-_The_Noun_Project.svg
Scientist	26-31,33,36	https://www.flaticon.com/free-icon/scientist_488085 Icon made by Freepik from www.flaticon.com licensed by Creative Commons BY 3.0
Office User	26-31,33,36	https://upload.wikimedia.org/wikipedia/commons/thumb/7/7a/Office_room_icon.svg/600px-Office_room_icon.svg.png
RO-Storage	27-31,33,36	https://www.flaticon.com/free-icon/secure-database_83704 Icon made by Freepik from www.flaticon.com licensed by Creative Commons BY 3.0

1
- collection of all protocols from all research groups
  - start with most used experiments (e.g. DNA extraction, Western Blot)

2
- review, evaluation, comparison, structure

3
- extraction of a general version
  - get out all experimental steps
  - for each step: define to-be included metadata

4
- development of metadata scheme
  using DMPs and metadata standards (e.g. Dublin Core? MIBBI?)
- apply ontology based vocabulary

5
- provision of templates ("scientists unable/unwilling to develop")
- revision, several iterations

Phrasing of requirements

- product neutral
  ➔ no hidden product specific requirements

- evaluation of requirement: transparent for bidders (yes / no)
  ➔ no adjectives, e.g.:
  🚫 "~~leicht~~ verständliche u. ~~leicht~~ zu bedienende Nutzerschnittstelle"

  "Gebrauchstauglichkeit nach DIN EN ISO 9241 Teile 11 und 110"
  ➔ verifiable!

- requirement restricting competition ➔ justify reasons

Right order of trial – tender?

1. trial first – then tender
    –   early consultation of purchase dept. for instructions on
        ➔ prevention of distortion of market competition
    –   tender results with not anticipated ELN: additional test necessary
        ➔ **tender first**!

2. tender first – then trial
    –   requirement specification needs practical experience
    –   market exploration: <u>allowed</u>, often <u>wished</u> by VOL/A before tender
    –   tender for purpose of market survey forbidden (§ 2 Abs. 3 VOL/A)
        ➔ **trial first**!

EU-DSGVO / GDPR

- consultation of data protection officer

- additional requirements in performance specification

- service / support by vendor (e.g. teleservice)
  ➔ mandatory: contract for commisioned data processing

Works council – contact early

- ➔ agreement needed

ELN-requirements to be part of a RDM according to FAIR data princ.

1.	masks	to capture metadata <u>together</u> with data	**FR**
2.	linking	to external files	**FA**
3.	export	to archives	**A**
4.	connectors	to research data repositories (PID, DOI)	**FA**

F

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

A

- A1. (meta)data are retrievable by their identifier using a standardized communications protocol
- A1.1 the protocol is open, free, and universally implementable
- A1.2 the protocol allows for an authentication and authorization procedure, where necessary
- A2. metadata are accessible, even when the data are no longer available

I

- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (meta)data use vocabularies that follow FAIR principles
- I3. (meta)data include qualified references to other (meta)data

R

- R1. meta(data) are richly described with a plurality of accurate and relevant attributes
- R1.1. (meta)data are released with a clear and accessible data usage license
- R1.2. (meta)data are associated with detailed provenance
- R1.3. (meta)data meet domain-relevant community standards

F indable A ccessible I nteroperable R eusable

https://nl.wikipedia.org/wiki/FAIR_principes