

Dr. Matthias Lange

 [0000-0002-4316-078X](https://orcid.org/0000-0002-4316-078X)

 [@langemger](https://twitter.com/langemger)

# Research Data Management at IPK Gatersleben

## Data Life Cycle, Profiles, Roles and Governance



### Conservation, Activation and Digitalisation of Plant Genetic Resources



M. Lange, — Licensed under CC BY 4.0: <https://creativecommons.org/licenses/by/4.0>

# Research Data Management – Bridging IT service and science

## Data Capture

## Research Data Management

## Immersive Data Visualisation

Imaging



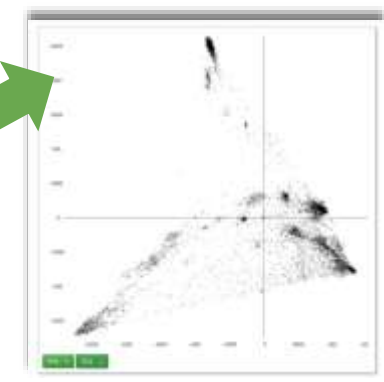
Trait Scoring



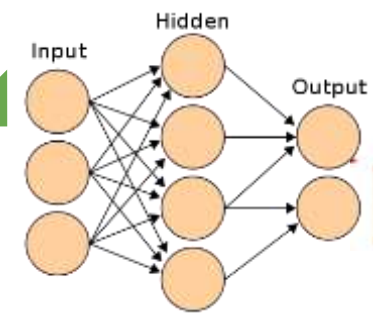
Genotyping



Environment



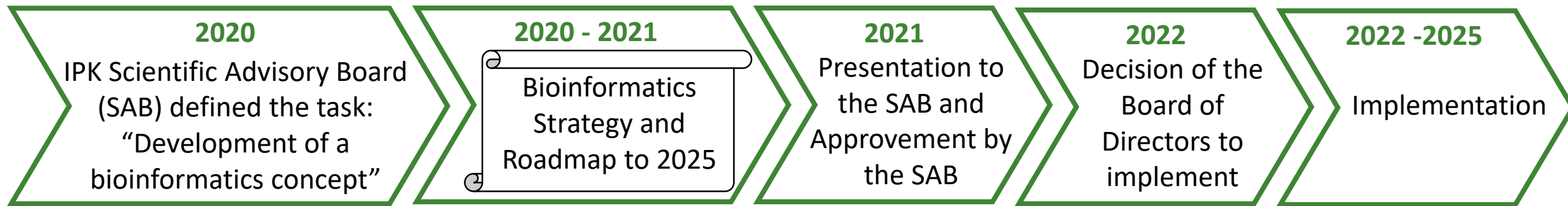
## Deep Learning & Model Building



M. Lange, – Licensed under CC BY 4.0: <https://creativecommons.org/licenses/by/4.0/>

es/by/4.0

# IPK Bioinformatics Concept



## Implementation Process ( RDM focal points: 2+1 of 5 )

<b>Service Infrastructure</b>	RISE-DE self assessment and implementation of a unique material management (digital twin)
<b>Search Portal</b>	Central entry point to search and access IPK data
<b>Overarching Challenge</b>	Organizational Adaption: Integration of the scientific library and a decentralized data stewardship team



# RDM Roles and Responsibilities – Job Profile



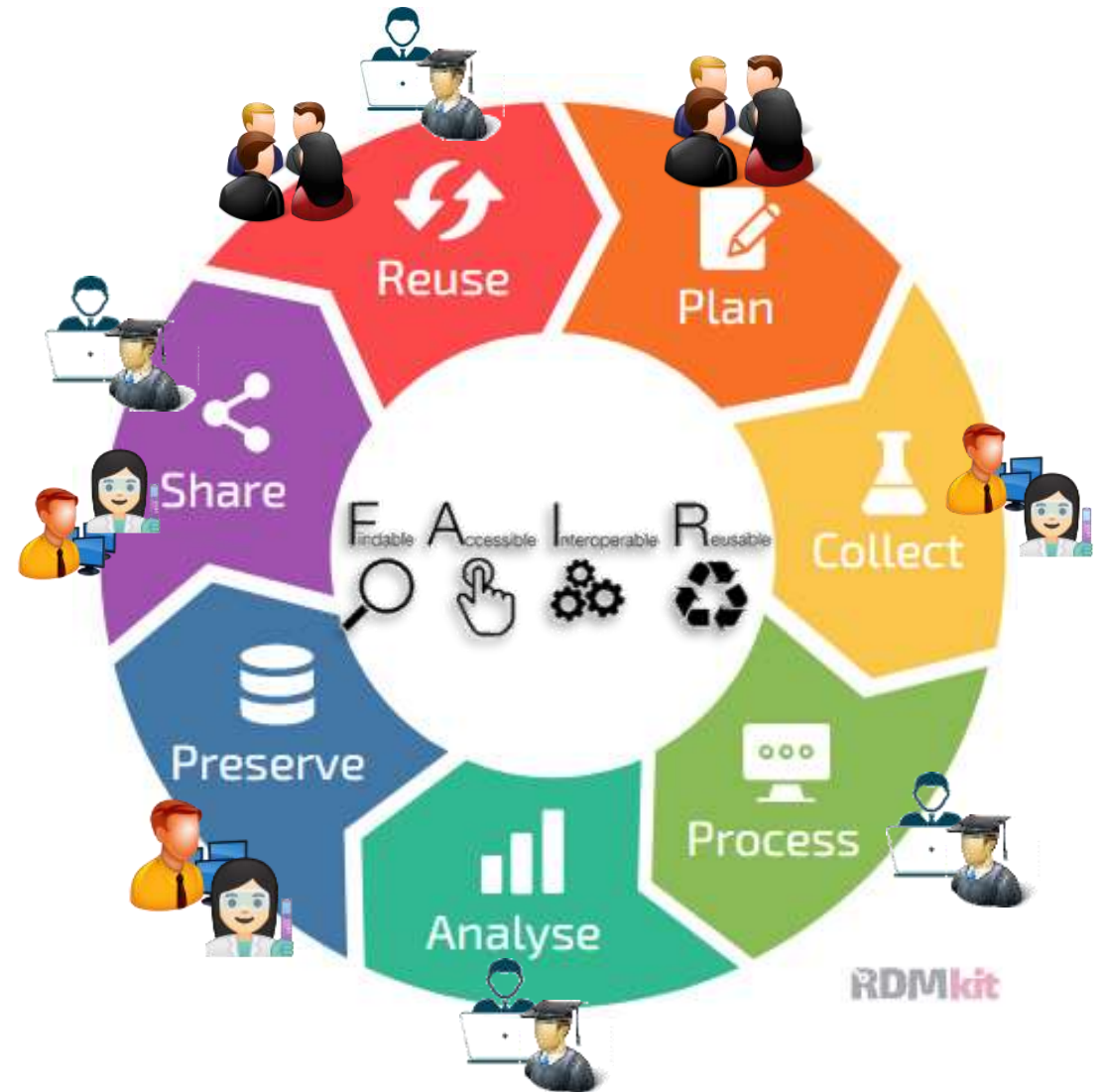
**Principal investigator**  
grant holder, research group leader



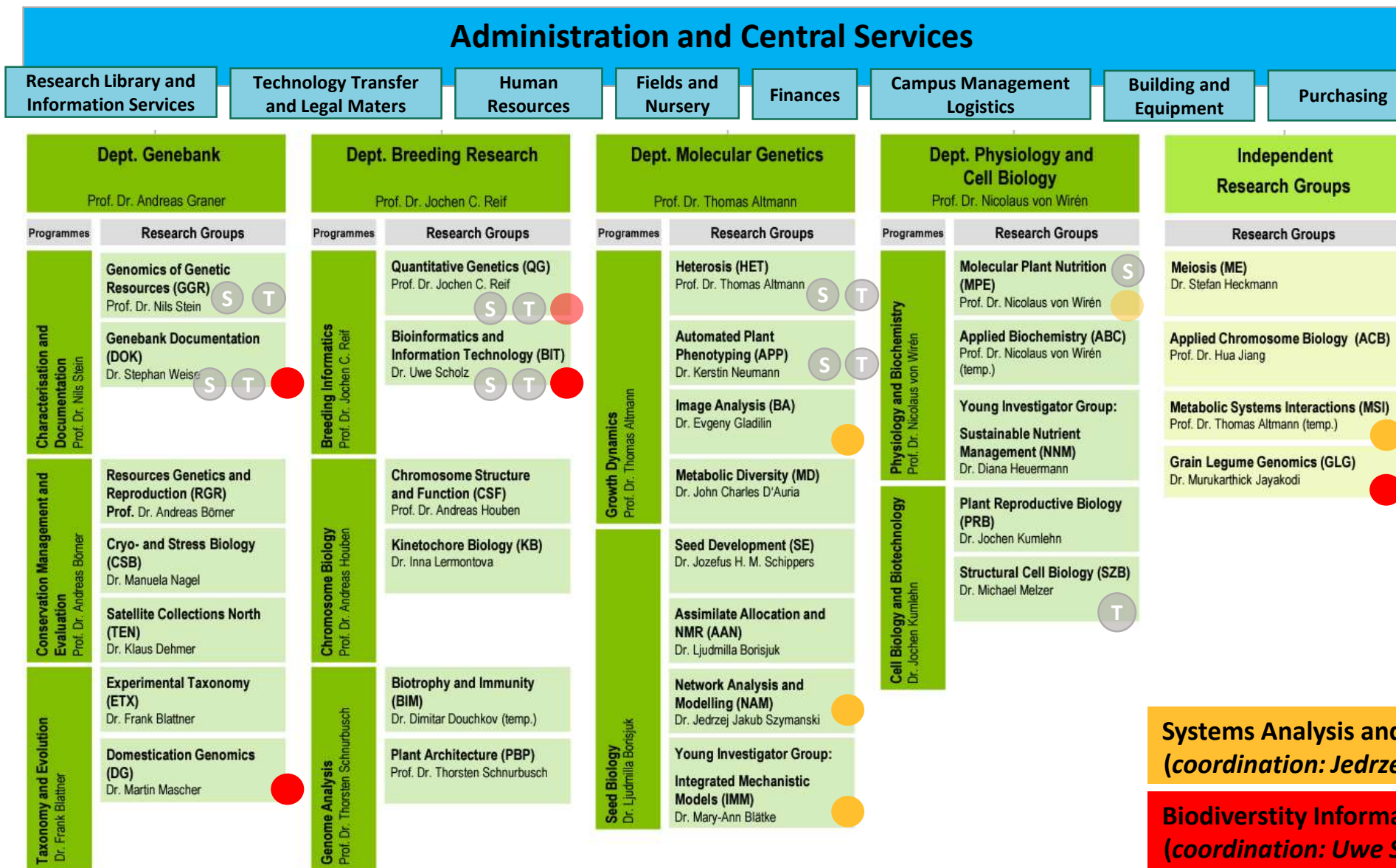
**Scientist**  
researcher work in a project or core facility



**Technician**  
worker in a field of technology



# Coordinated RDM Roles and Responsibilities



RDM pilot groups



S **Scientist**



T **Technician**

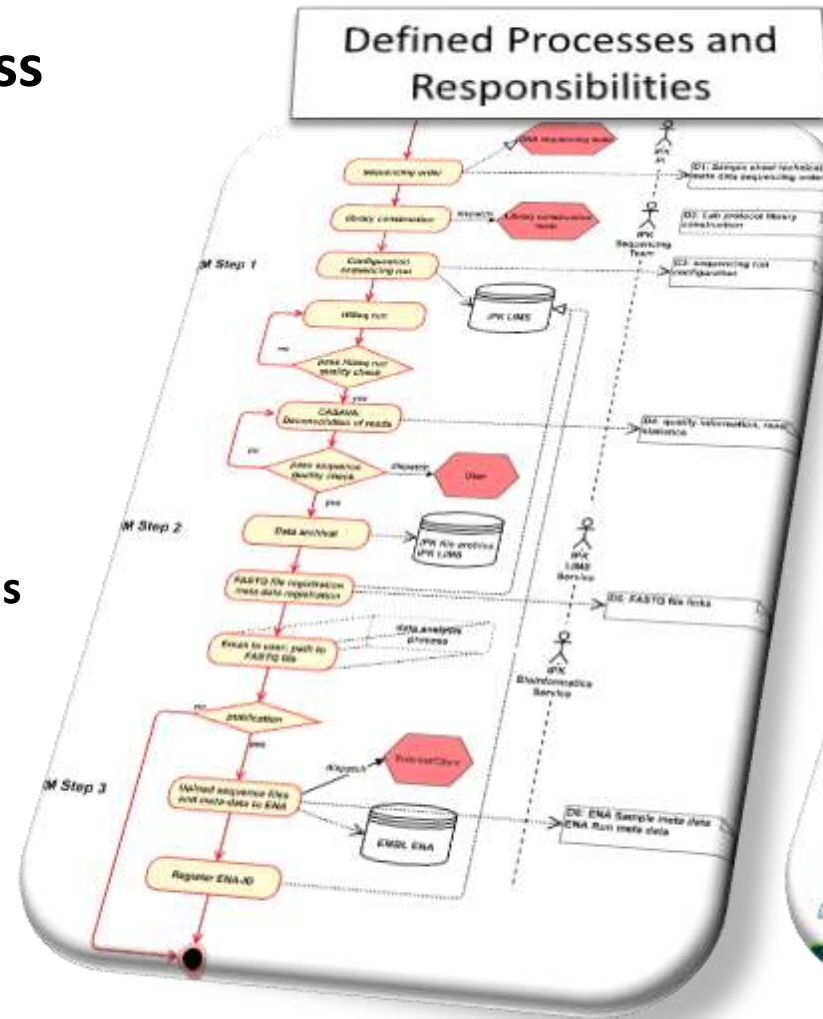
**Systems Analysis and Modelling**  
*(coordination: Jędrzej J. Szymanski)*

**Biodiversity Informatics**  
*(coordination: Uwe Scholz)*

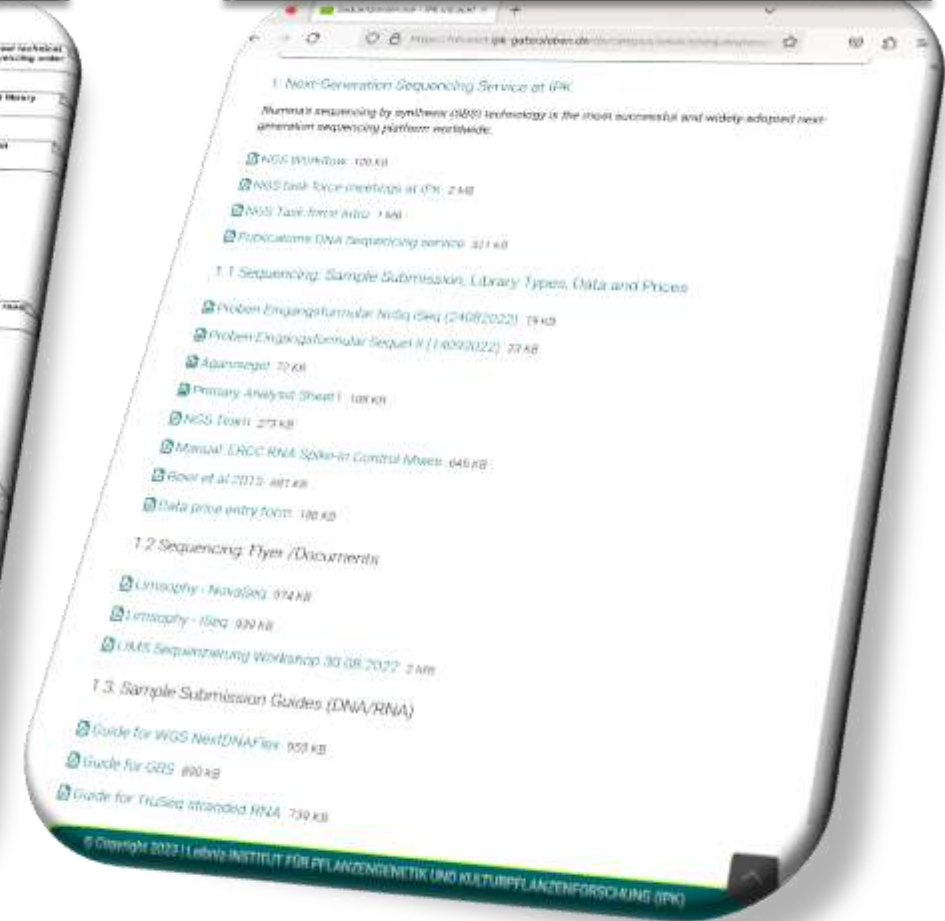
# IPK RDM Service Processes

## Example Sequencing Process

- Defined Data Submission System
- Published, Binding Process
- Defined Responsibilities across IPK
- Defined Standards and Fault Actions
- Defined Data Transition Points
- LIMS and API based Data Flow
- User Notification System



## Mandatory Data Standards, Transition Points and Regulations

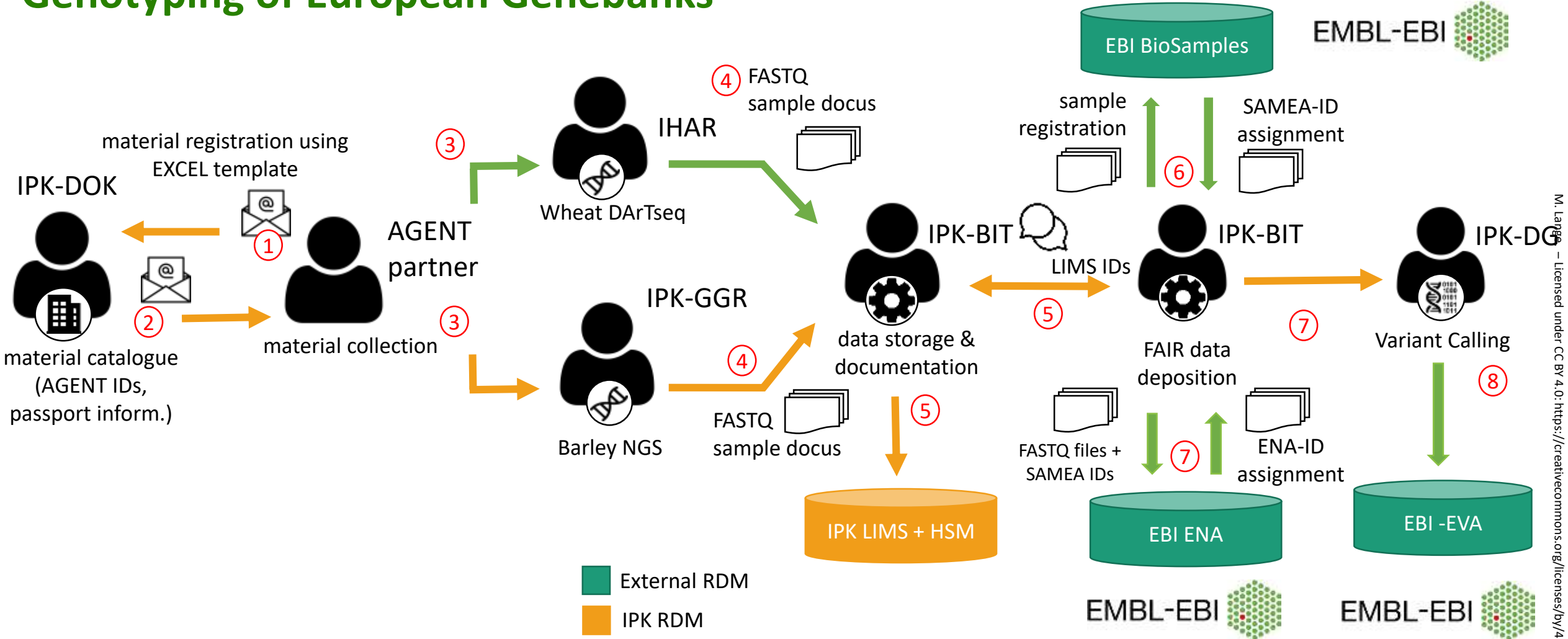




# Integration of external RDM Process



## Genotyping of European Genebanks



M. Lange – Licensed under CC BY 4.0: <https://creativecommons.org/licenses/by/4.0>

# Summary – Profiles and Roles

- **Location of RDM professionals**
  - positions are **spread across IPK** research groups, persons **share several roles** in one position
  - diverse and heterogenous quality of RDM
- **Job profiles and responsibilities**
  - responsibilities and task in RDM is in **responsibility within research groups**
  - **pilot groups** provide institute wide RDM services:  
e.g. sequencing, Gene Bank management, microscopy
  - common RDM aspects across groups moderated by a **federate Bioinformatics coordination unit**
  - IPK administration and **scientific library** is **not part** of **RDM** activities
  - **no binding RDM processes**, but recommendations linked to WGL guidelines
- **Integration in research themes**
  - central, **core facility processes**, e.g. Genebank, microscopy, sequencing service
  - RDM is in **majority project driven**
- **Resources and scaling**
  - **permanent** positions for **scientist and technician** for RDM in **pilot groups**
  - RDM responsibility persons are **group leader** as permanent position
  - RDM responsibility in projects are **project PI** with support of Bioinformatics coordination



# Acknowledgement

- IPK Bioinformatics Coordination

- Uwe Scholz
- Jędrzej Szymanski

- Research Group IPK-BIT

- Anne Fiebig
- Manuel Feser
- Patrick König
- Team IT
  - Heiko Mieke
  - Jens Bauerfeind
- Team Scientific Data Management
  - Daniel Arend
  - Danuta Schüler
  - Jorge Brizuela
  - Elena Rey-Mazon
  - Theodor Strauch

- IPK research groups

- AAN
- HET
- GGR
- ADP
- DOK
- DG
- QG

- all external co-operation partner

all funding agencies: BMBF, DFG, EC

