

# Dataflow from the field to a data repository: small-scale distribution of plant galls as use case

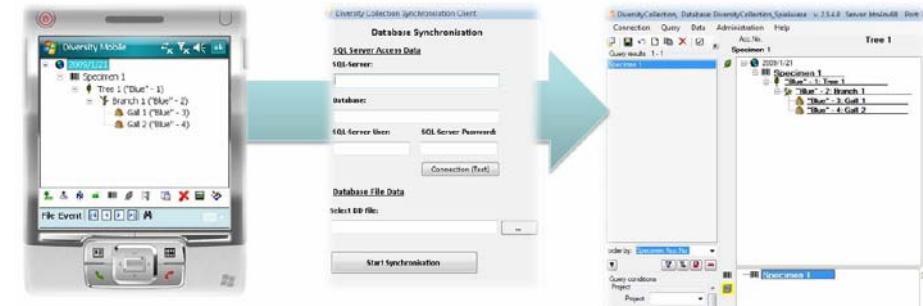


Dataflow from the field to a data repository:  
small-scale distribution of plant galls  
as use case

Alexandra Kehl  
Zentrales Labor für  
DNA-Analytik und Ökoinformatik

# Dataflow from the field to a data repository: small-scale distribution of plant galls as use case

- Introduction to the IBF-Project
- Plant galls  
as use case (IBF Ecology)
- Benefits of the aspired  
dataflow



Dataflow from the field to a data repository:  
small-scale distribution of plant galls  
as use case

Alexandra Kehl  
Zentrales Labor für  
DNA-Analytik und Ökoinformatik



# The IBF-Project

Setting up an **I**nformation **N**etwork on **B**iological  
Research Data gained in the **F**ield (I-B-F)

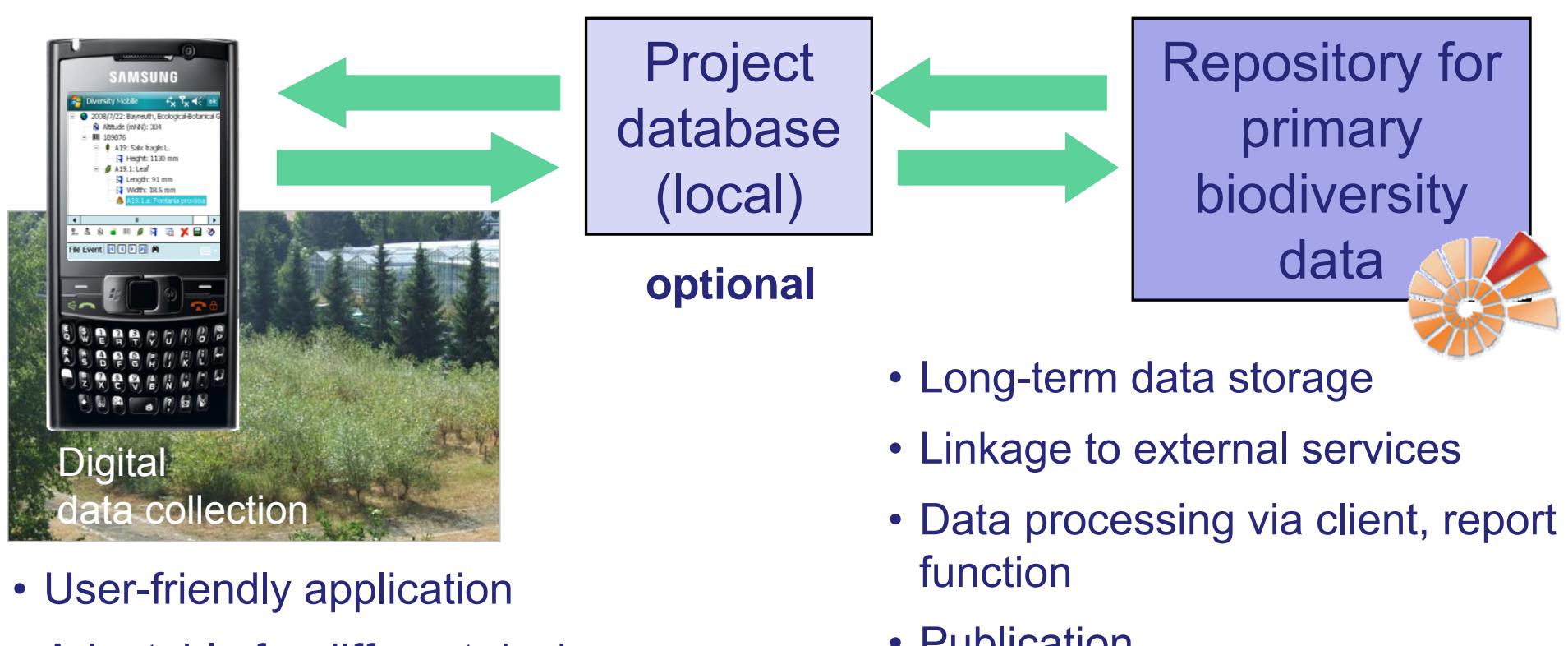


Aufbau eines **I**nformationsnetzes für **B**iologische  
**F**orschungsdaten von der Erhebung im Feld bis  
zur nachhaltigen Sicherung in einem  
**P**rimärdatenrepository (IBF)

# The IBF-Project



Establishment of the dataflow from the field to a data repository (and back):





# The IBF-Project

- Start of project: February 2009
- Participating institutions:
  - IT-Centre SNSB (München)
  - Institute of Botany (University of Regensburg)
  - Applied Informatics IV (University of Bayreuth)
  - DNA-Analytics and Ecoinformatics (University of Bayreuth)

Dataflow from the field to a data repository:  
small-scale distribution of plant galls  
as use case

Alexandra Kehl  
Zentrales Labor für  
DNA-Analytik und Ökoinformatik



# The IBF-Project

Establishment of the dataflow from the field to a data repository (and back):

- Data gathering via mobile device
- Special user interface for data entry and for access taxonomic names, ecological descriptors, and general scientific term presets
- Transfer of data contents to a data repository (here: *DiversityCollection*)
- Redistribution of data to end-users, data exchange with PDAs, wrappers (ABCD schema) and external applications for data analysis and presentation

# The IBF-Project

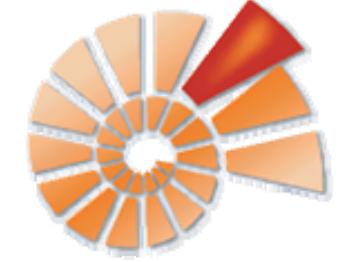


- Data gathering: IBF Monitoring,  
IBF Ecology
- User interface: IBF Mobile
- Data transfer: IBF Mobile,  
IBF Data repository
- Data exchange: IBF Data repository



Dataflow from the field to a data repository:  
small-scale distribution of plant galls  
as use case

Alexandra Kehl  
Zentrales Labor für  
DNA-Analytik und Ökoinformatik



# The IBF-Project

- Data collection: IBF Monitoring,  
**IBF Ecology**
- User interface: IBF Mobile
- Data transfer: IBF Mobile,  
IBF Data repository
- Data exchange: IBF Data repository

Dataflow from the field to a data repository:  
small-scale distribution of plant galls  
as use case

Alexandra Kehl  
Zentrales Labor für  
DNA-Analytik und Ökoinformatik

# IBF Ecology

## Small-scale distribution of gall-inducing sawflies (Hymenoptera, Tenthredinidae) on *Salix fragilis*



*Pontania proxima*



*Phyllocolpa oblita*



*Euura testceipes*

Dataflow from the field to a data repository:  
small-scale distribution of plant galls  
as use case

Alexandra Kehl  
Zentrales Labor für  
DNA-Analytik und Ökoinformatik

# IBF Ecology: scientific aims

- Differences in temporal and spatial distribution of three gall-inducing sawfly species on the same host plant species
- Relevance of the factors „shoot length“ and „shoot growth“ for oviposition (gall induction)

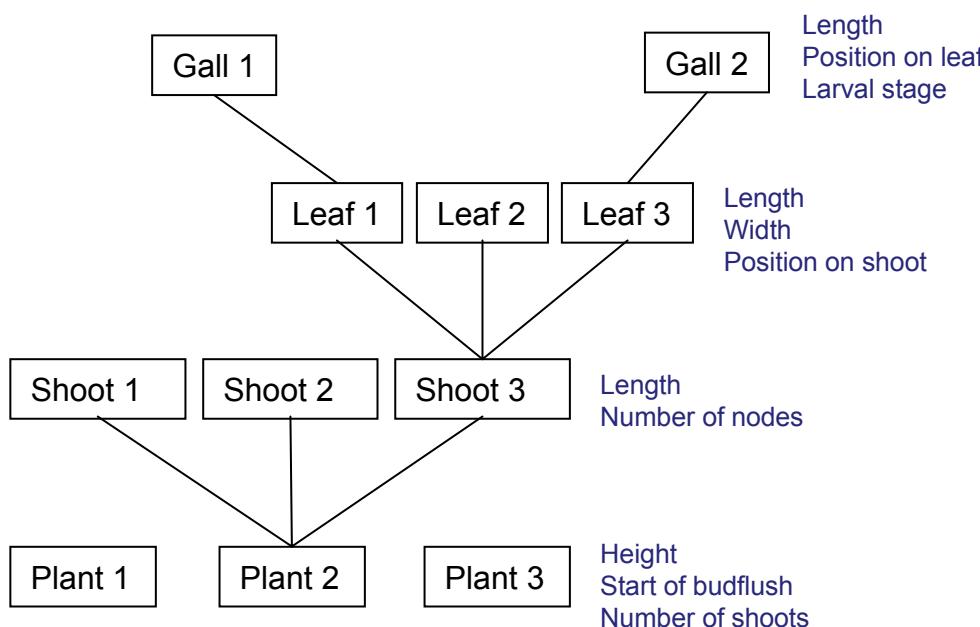


Dataflow from the field to a data repository:  
small-scale distribution of plant galls  
as use case

Alexandra Kehl  
Zentrales Labor für  
DNA-Analytik und Ökoinformatik

# IBF Ecology: technical aims

→ To test and further develop the user interface by collecting “multidimensional” data at regular time intervals



Dataflow from the field to a data repository:  
small-scale distribution of plant galls  
as use case

Alexandra Kehl  
Zentrales Labor für  
DNA-Analytik und Ökoinformatik

# IBF Ecology: approach

Long-term investigation area since 2002 (Ecological-Botanical Garden, University of Bayreuth):

→ Host plant selectivity of gall-inducing sawflies on 240 individuals derived from 19 clones (*S. alba/S. fragilis*-aggregate)



# IBF Ecology: approach

- 40 plants (4 male clones)
- Regular monitoring of plant growth and phenology
- Regular monitoring of selected shoots on these plants (growth and gall presence)
- Regular monitoring of gall presence and location on the plants
- Manipulation experiment (in 2010)

# Dataflow from the field to a data repository: small-scale distribution of plant galls as use case

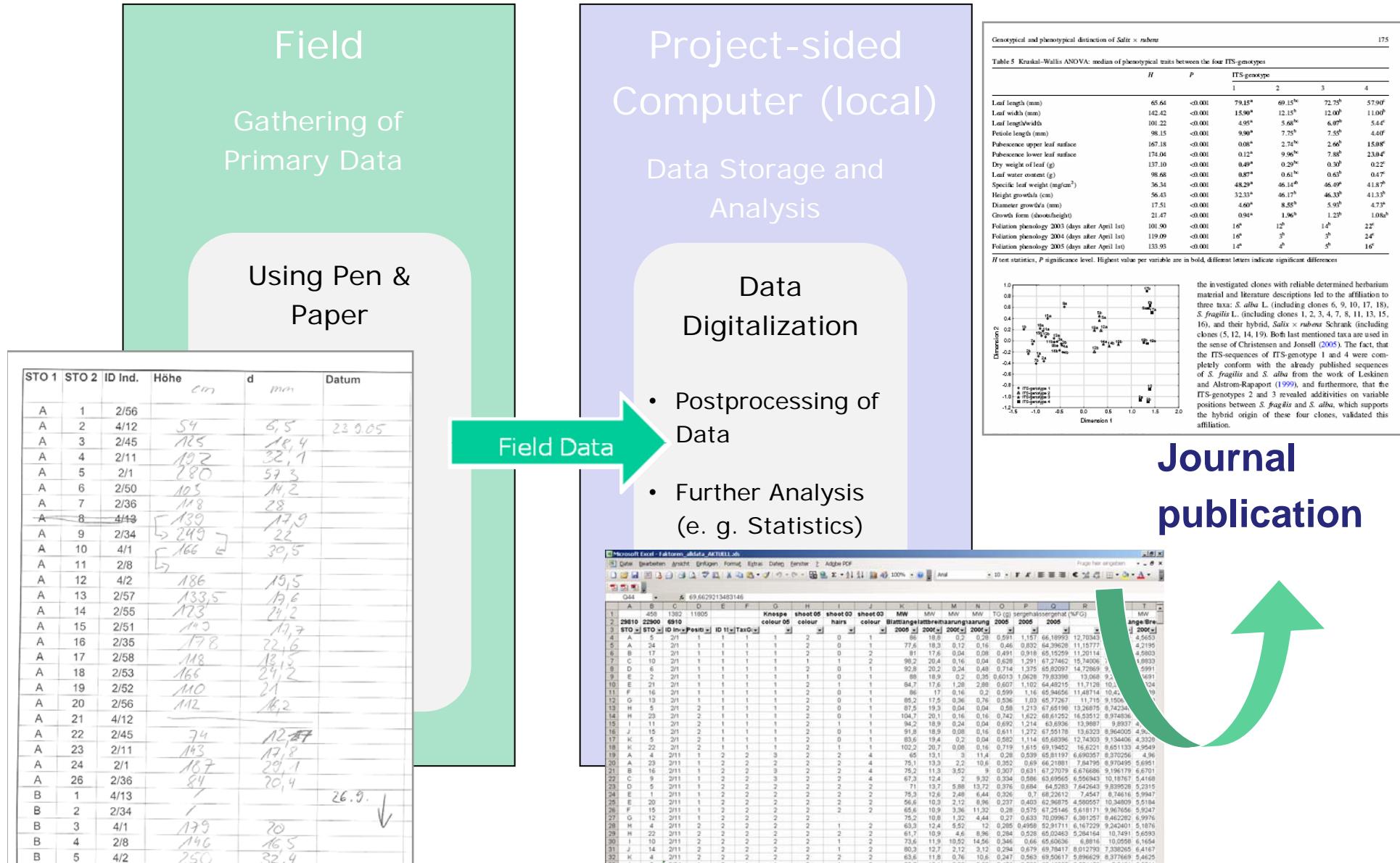
- Introduction to the IBF-Project
- Plant galls  
as use case (IBF Ecology)
- Benefits of the aspired  
dataflow



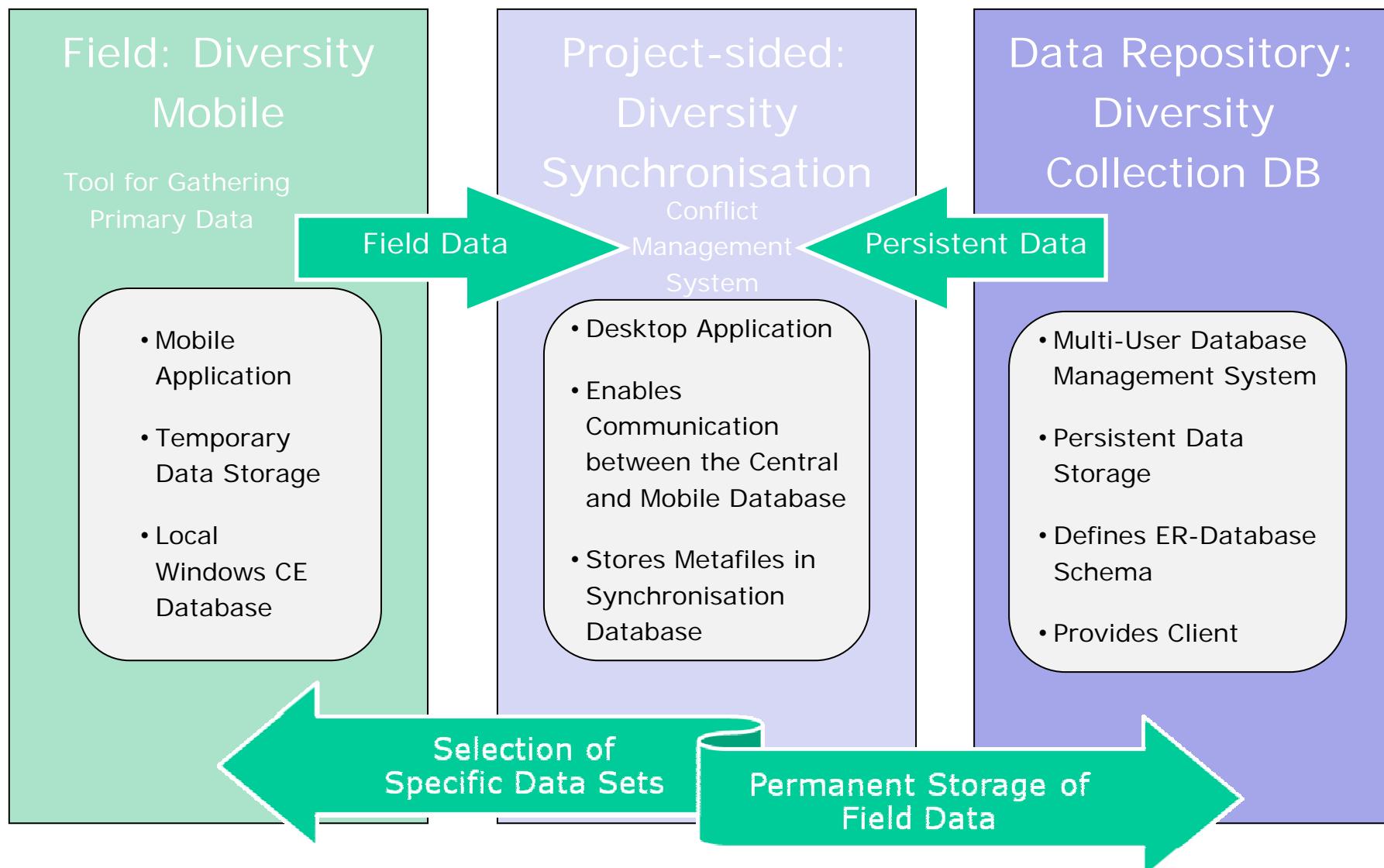
Dataflow from the field to a data repository:  
small-scale distribution of plant galls  
as use case

Alexandra Kehl  
Zentrales Labor für  
DNA-Analytik und Ökoinformatik

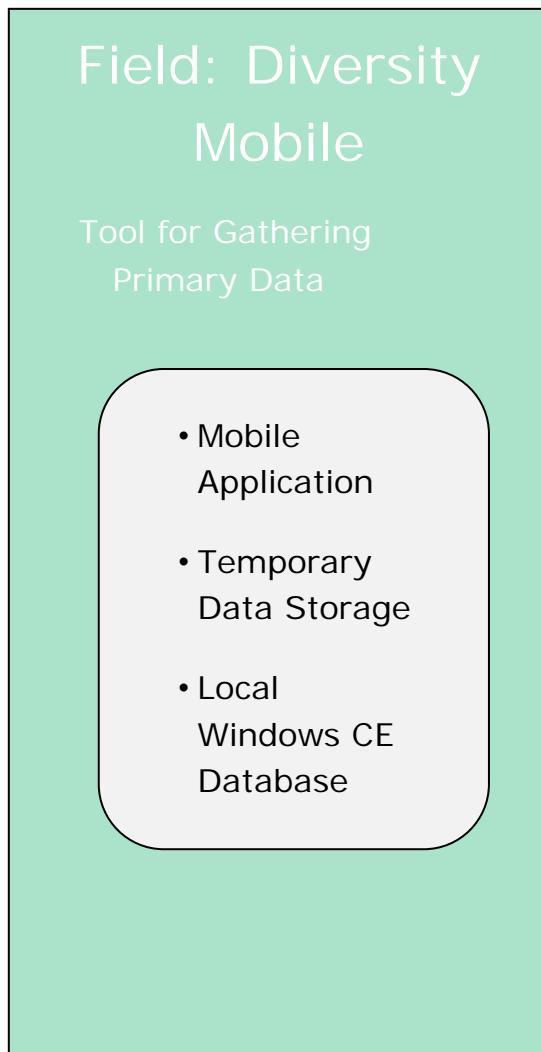
# Data flow: typical situation



# Data flow: aspired situation



# Data flow: aspired situation



- Previously recorded data accessible in the field
- Access to external data sources (e. g. standard list of taxon names)
- Use of GPS functionality
- Collection of multimedia data (images, video, sound)
- All entered data will be transferred to the data repository (usually via local project database)

# Data flow: aspired situation



- Long-term data storage
- Access to data history
- Individual observations become referable (GUIDs)
- Publication of (selected) data (→ WWW)
- Selective data export (formated by report function) to the project-sided computer for statistical processing
- Traditional publication of data

Data Repository:  
Diversity  
Collection DB

- Multi-User Database Management System
- Defines ER-Database Schema
- Provides Client

# Conclusion: Benefits of the aspired dataflow

- Facilitation by using a mobile device for data collection
- Long-term storage of primary data (data repository)
- Data processing possible at any time:
  1. via client on local project-database
  2. via client on remote repository-database
- Linkage to (external) web services (→ completion with standard data and improvement of data quality)
- Publication of primary data: access for others

# Thank you!



Zentrales Labor  
für DNA-Analytik  
und  
Ökoinformatik



Dataflow from the field to a data repository:  
small-scale distribution of plant galls  
as use case

Alexandra Kehl  
Zentrales Labor für  
DNA-Analytik und Ökoinformatik