



Universität Bayreuth
Lehrstuhl für Angewandte Informatik IV
Datenbanken und Informationssysteme

Prof. Dr.-Ing. Stefan Jablonski

DiversityMobile und Mechanismen der Datensynchronisation

Tobias Schneider, Stefan Jablonski, Bernhard Volz

Universität Bayreuth, Lehrstuhl Angewandte Informatik IV

Gerhard Rambold, Alexandra Kehl

Universität Bayreuth, DNA-Analytik und Ökoinformatik

Dagmar Triebel, Markus Weiss, Dieter Neubacher

SNSB IT Zentrum

Peter Poschlod

Universität Regensburg

Ausgangspunkt: papierbasierte Datenerfassung

Feld

Aufnahme von Primärdaten

papierbasierte Datenerfassung

Primärdaten

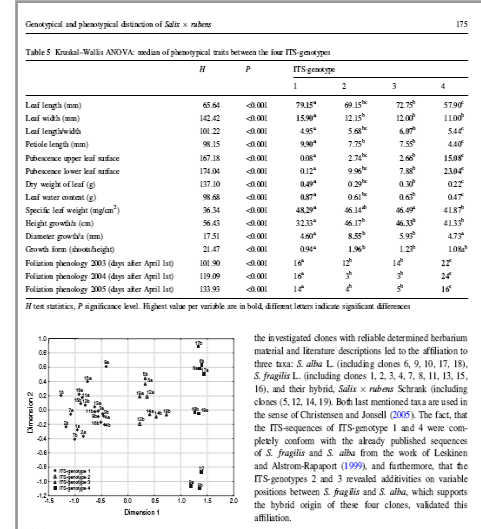
Projekt

Temporäre Datenspeicherung

Digitalisierung

- Nachbearbeitung der Daten
- Auswertung (Statistiken)

STO 1	STO 2	ID Ind.	Höhe	d	Datum
			cm	mm	
A	1	2/56			
A	2	4/12	54	6,5	23.9.05
A	3	2/45	105	18,4	
A	4	2/11	102	22,1	
A	5	2/1	280	59,3	
A	6	2/50	105	14,2	
A	7	2/36	118	28	
A	8	4/14	139	17,9	
A	9	2/34	249	22	
A	10	4/1	166	30,5	
A	11	2/8			
A	12	4/2	186	19,5	
A	13	2/57	133,5	19,6	
A	14	2/55	173	24,2	
A	15	2/51	143	19,7	
A	16	2/35	178	22,6	
A	17	2/58	118	14,3	
A	18	2/53	166	24,2	
A	19	2/52	110	21	
A	20	2/56	112	16,2	
A	21	4/12			
A	22	2/45	74	12,7	
A	23	2/11	183	17,8	
A	24	2/1	187	29,1	
A	26	2/36	84	20,4	
B	1	4/13			26.9.
B	2	2/34			
B	3	4/1	179	20	



Veröffentlichung

Anstieg der Anforderungen in der Primärdatenspeicherung

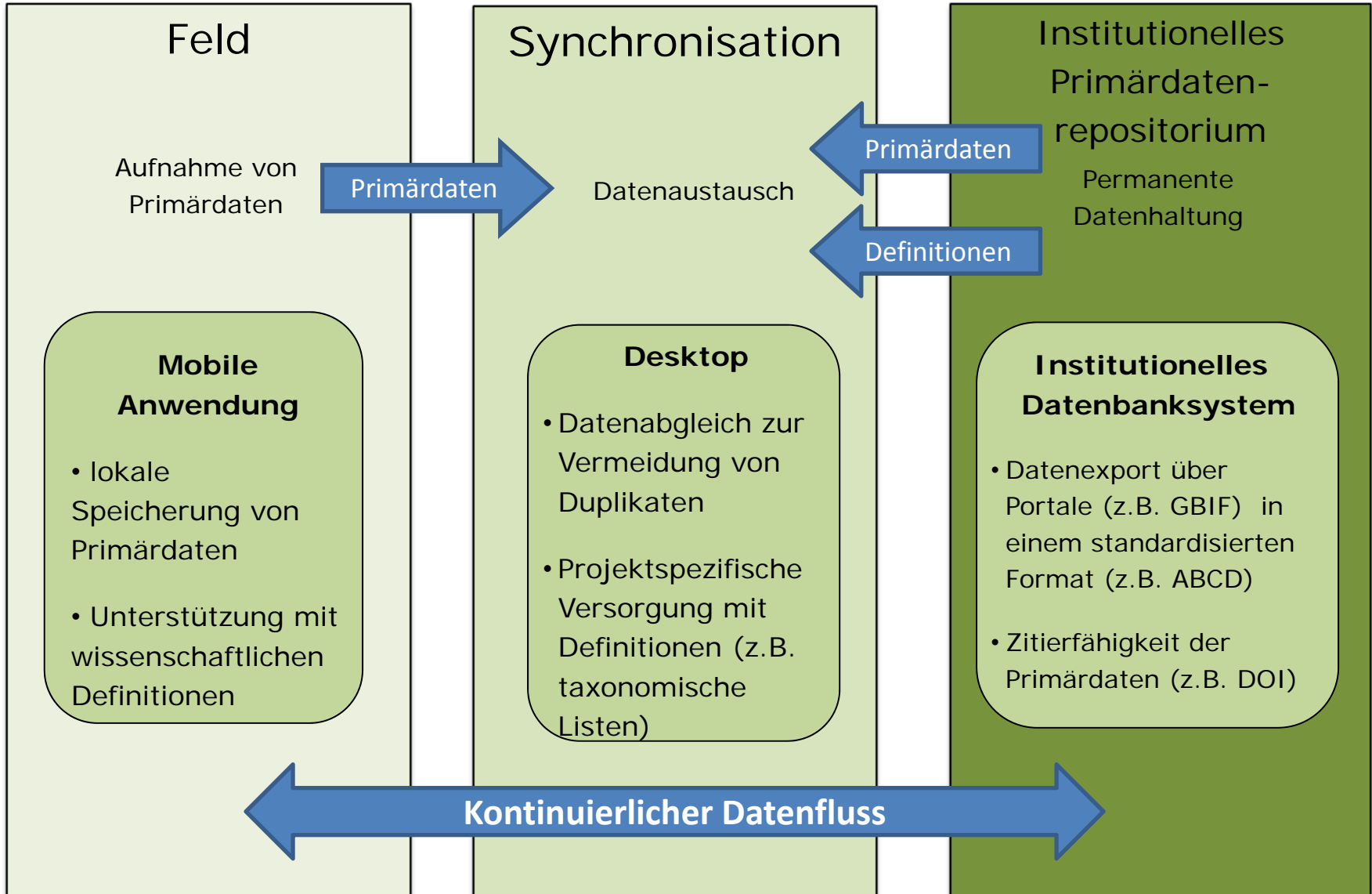
Vorschläge guter wissenschaftlicher Praxis (DFG 1998):

Primärdaten als Grundlagen für Veröffentlichungen sollen auf haltbaren und gesicherten Trägern **in der Institution**, wo sie entstanden sind, **für zehn Jahre** aufbewahrt werden.

Leitbild der Schwerpunktinitiative „Digitale Information“ (Allianzpartnerorganisationen 2008):

Es gilt eine **nachhaltige integrierte digitale Forschungsumgebung** zu schaffen, in der jeder Forschende **von überall** in Deutschland auf das **gesamte publizierte Wissen** und die **relevanten Forschungsprimärdaten** zugreifen kann.

Ziel: Kontinuierlicher Datenfluss



Aufbau eines **Informationsnetzes** für **biologische Forschungsdaten** von der Erhebung im Feld bis zur nachhaltigen Sicherung in einem Primärdatenrepositorium (**IBF**)

Mit besonderem Dank an:

Deutsche Forschungsgemeinschaft
Förderbereich LIS – Informationsmanagement
(GZ: INST 106535/1-1, INST 21946/1-1, INST
2850/1-1, INST 747/1-1)

DFG

Die Arbeitsgruppen des IBF-Projekts

Arbeitsgruppe
Pilze und Flechten



Arbeitsgruppe
Gefäßpflanzen



Arbeitsgruppe
Ökologie



IT-Arbeitsgruppe



Arbeitsgruppe
Pilze und Flechten



Arbeitsgruppe
Gefäßpflanzen



Arbeitsgruppe
Ökologie



Kartierung von Pilzen und Flechten in Regensburg und Umgebung



Arbeitsgruppe
Pilze und Flechten



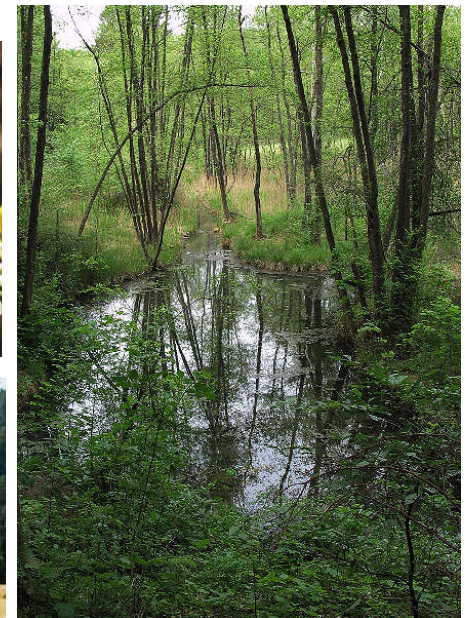
Arbeitsgruppe
Gefäßpflanzen



Arbeitsgruppe
Ökologie



Kartierung von Gefäßpflanzen in Regensburg und Umgebung



Arbeitsgruppe
Pilze und Flechten



Arbeitsgruppe
Gefäßpflanzen



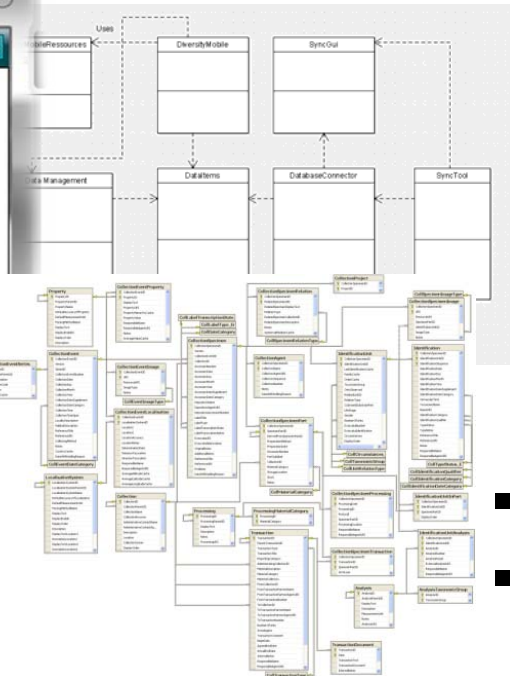
Arbeitsgruppe
Ökologie



- Verteilung von Gallwespen auf *Salix fragilis*
- Versuchsfläche von 40 Pflanzen in Bayreuth



IT-Arbeitsgruppe



Entwicklung des mobilen Clients

Entwicklung von generischen

Daten Im- und
Exportschnittstellen

Entwicklung von
Authentifizierungsmechanismen
und -strukturen

- Aufbau eines Datenrepositoriums
- Aufbau eines
Softwarerepositoriums

Technische Umsetzung des IBF-Projekts

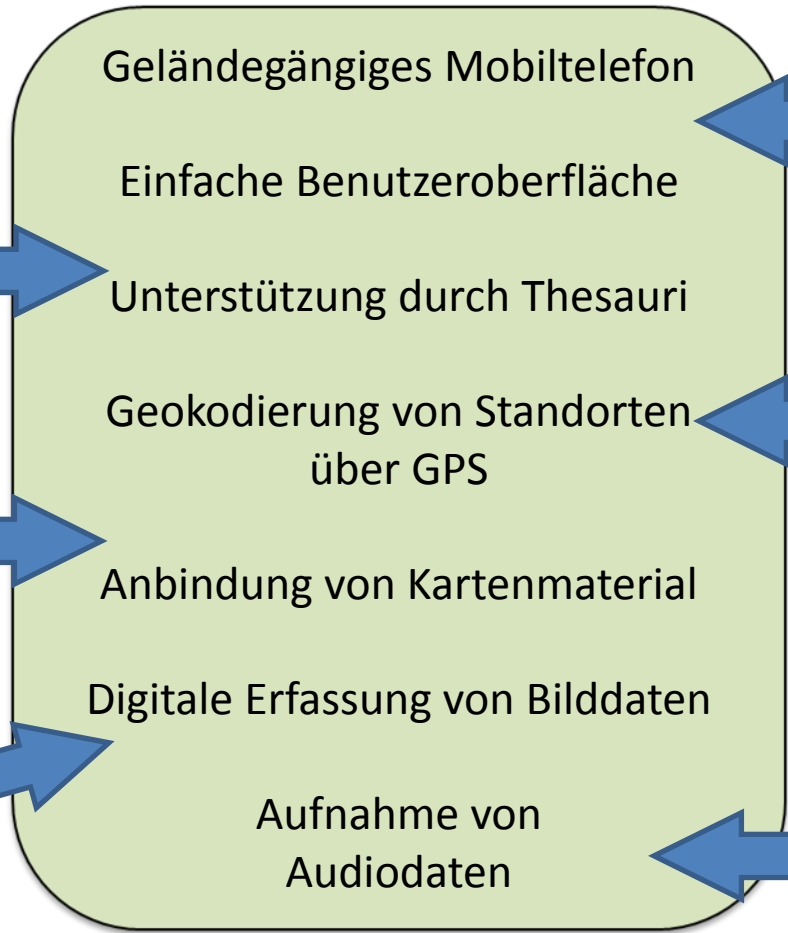
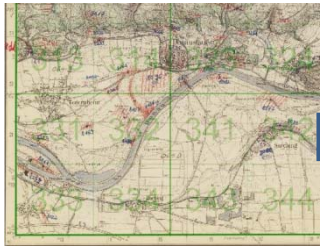
- **Mobiler Client (*DiversityMobile*) als Teil der *Diversity Workbench***
 - Implementierung Benutzeroberfläche zur Datenaufnahme für mobile Geräte ab Windows Mobile 5.0
 - Speicherung der Primärdaten und Anbindung taxonomischer Datenbanken mit einem selbstentwickelten Objekt-Relationalen Mapper

- **Datentransfer an das Repository und Datenhaltung**
 - Synchronisation mit *DiversityCollection* über eine Desktop-Oberfläche
 - Nachbearbeitung von Daten in *DiversityCollection*
 - *Datenspeicherung am Leibniz-Rechenzentrum*
 - www.diversityworkbench.net

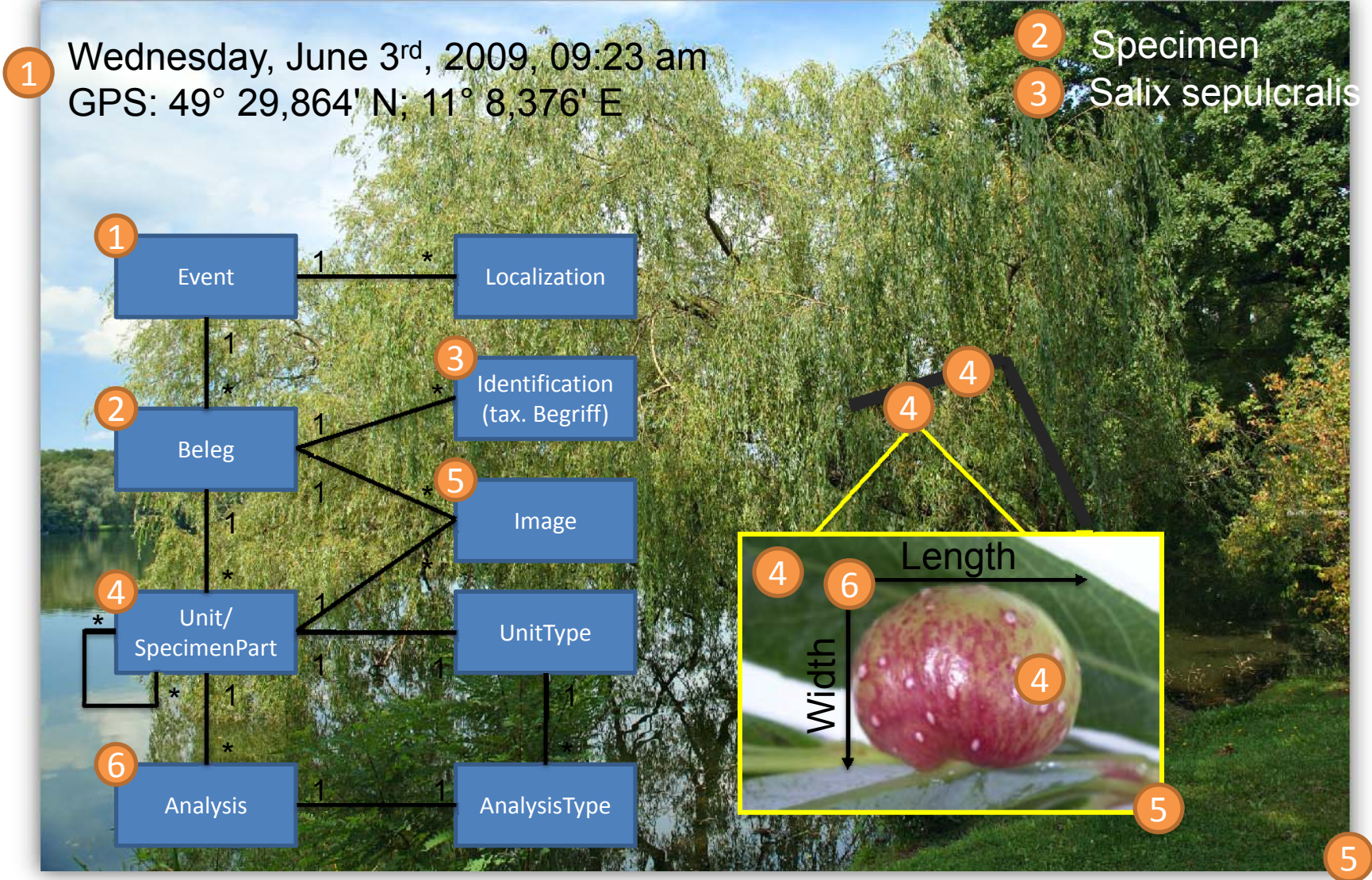
- **Datenverteilung an Endnutzer**
 - Datenexport im ABCD-Format und speziellen Formaten
 - Anbindung an BioCase, GBIF, LifeWatch etc.



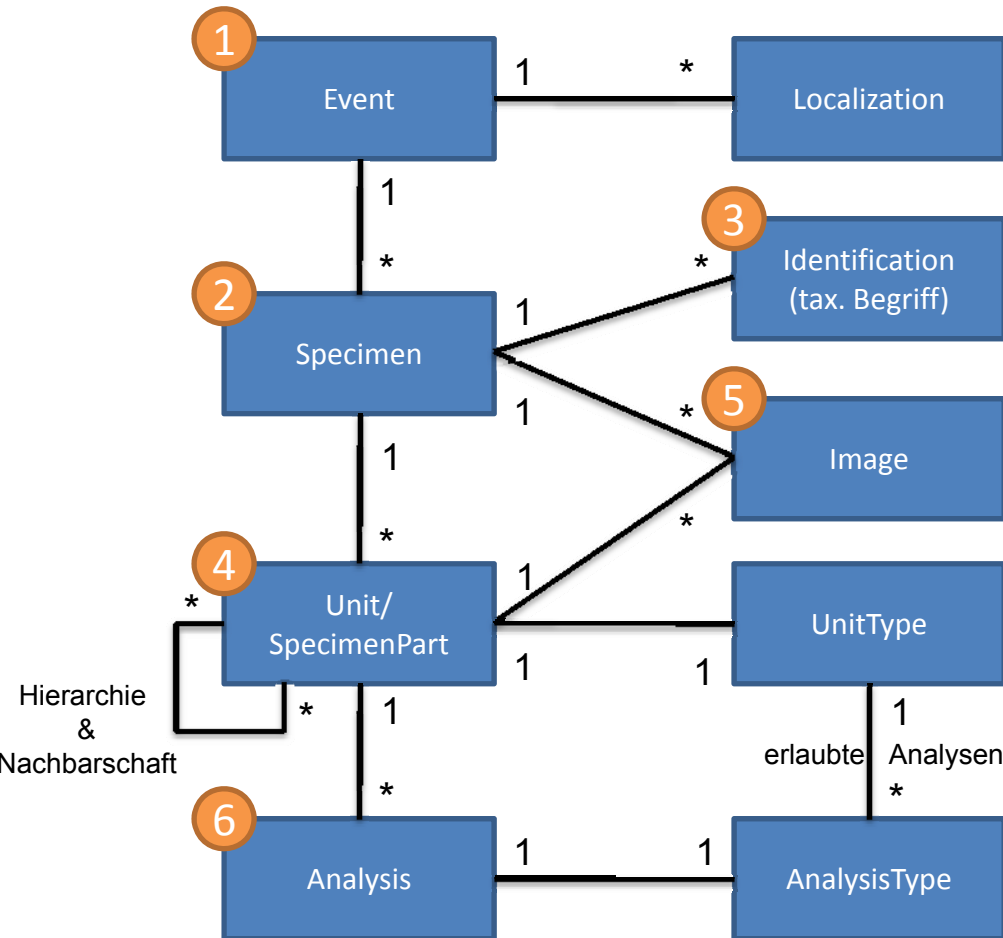
Features von *DiversityMobile*



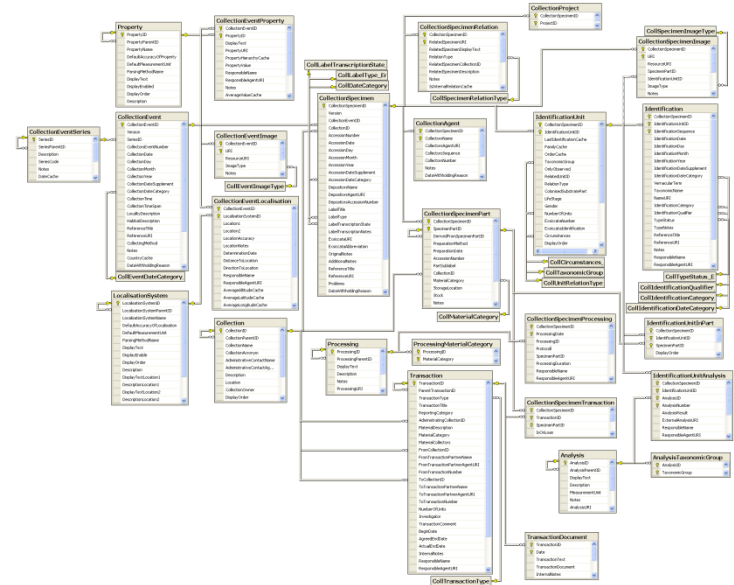
Beispiel aus der Praxis



Übersicht



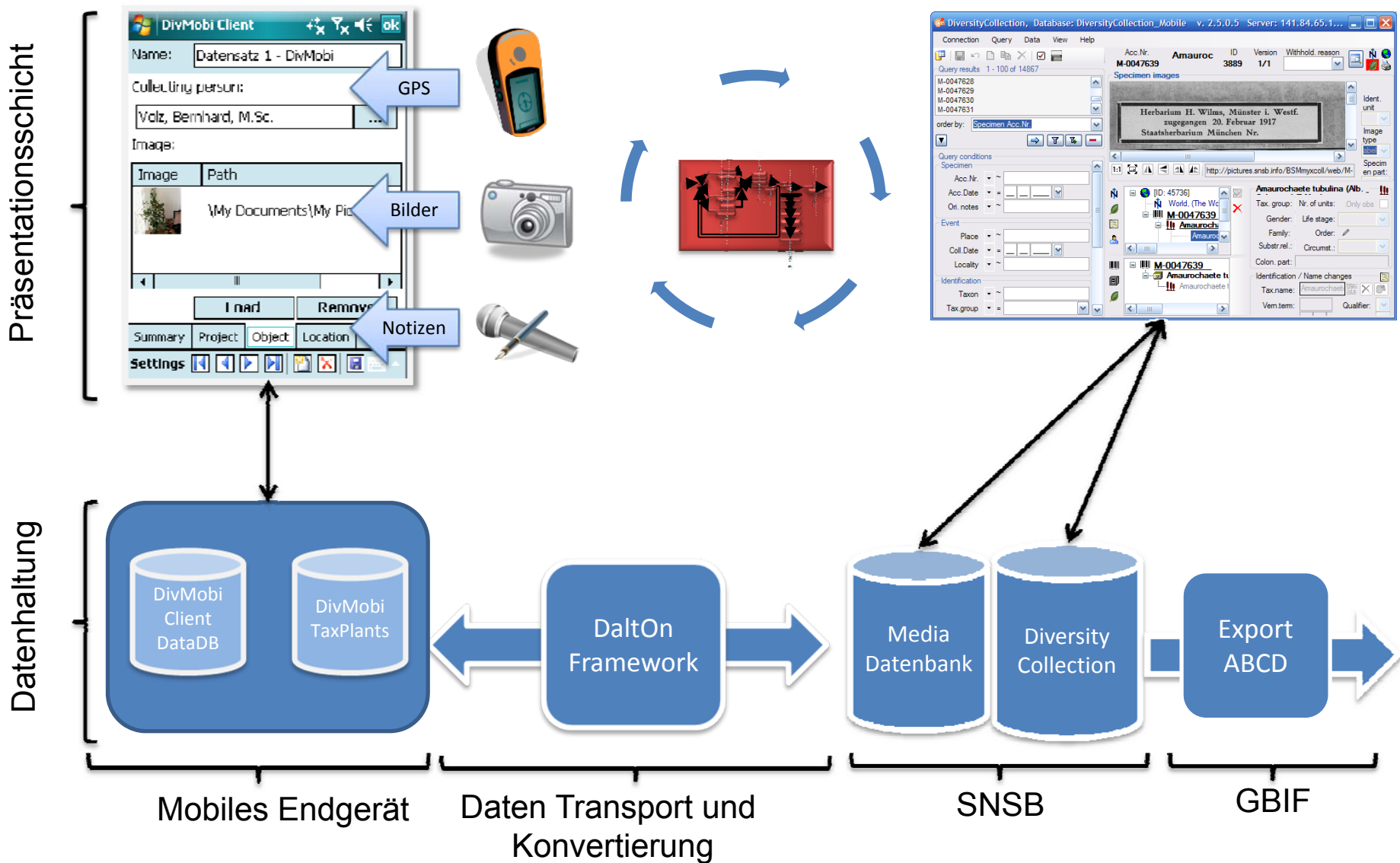
Implementiertes Datenschema



http://www.diversityworkbench.net/Portal/wiki/CollectionModel_v2.05

Umfang: ca. 40 Tabellen mit je 10-30 Spalten

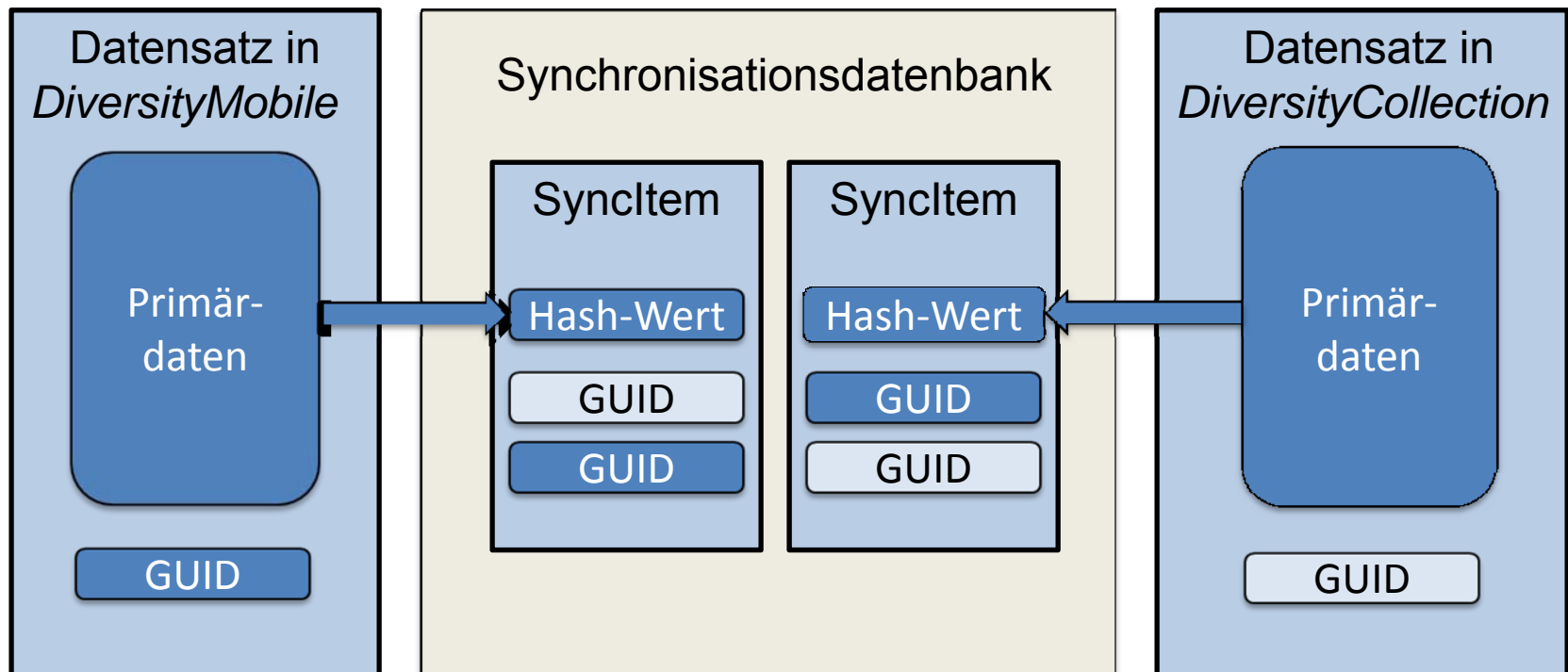
Architektur-Überblick



Konzept Synchronisation

Identifizierung von Datensätzen über GUIDs im Microsoft Universally Unique Identifier (*UUID*) Standard

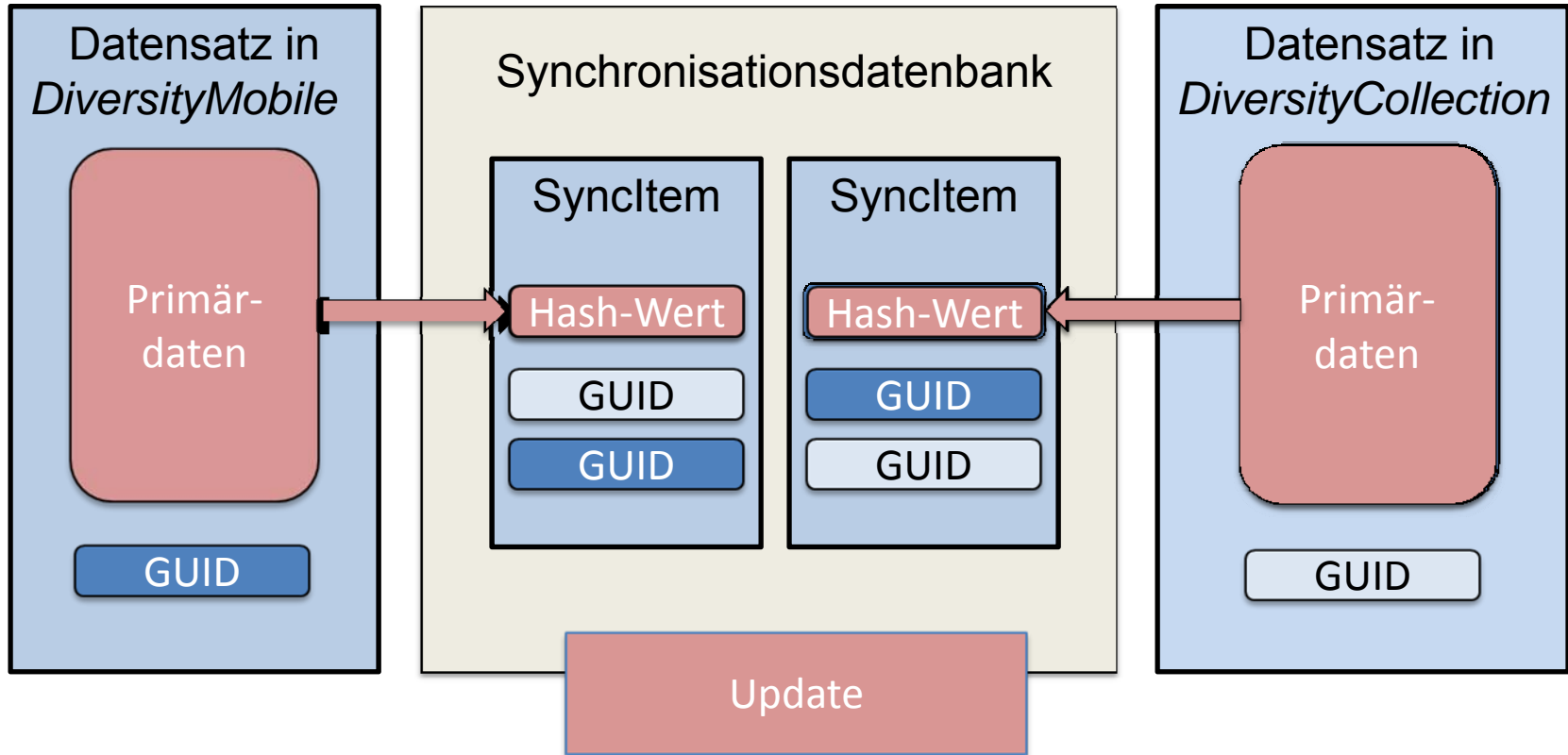
Berechnung von Hashwerten aus den Primärdaten



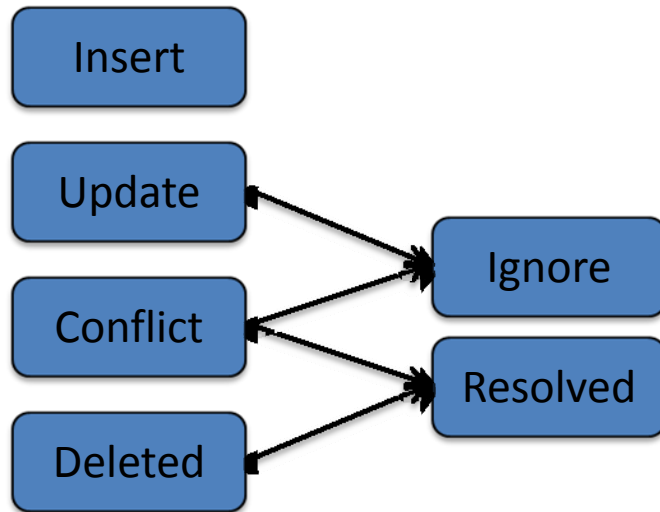
Synchronisationszustände

- Durch Vergleich der Hashwerte können verschiedene Synchronisationszustände identifiziert werden

SyncItem	Zustand	Aktion
Kein Eintrag	Insert	Objekt wird eingefügt
Eintrag vorhanden, Hashwert unverändert	Ignore	Keine Aktion
Eintrag vorhanden, Hashwert der Quelle verändert	Update	Objekt wird aktualisiert
Eintrag vorhanden, Hashwert des Ziels verändert	Conflict	Benutzerdefinierte Aktion
Eintrag vorhanden, Zielobjekt gelöscht	Deleted	Benutzerdefinierte Aktion



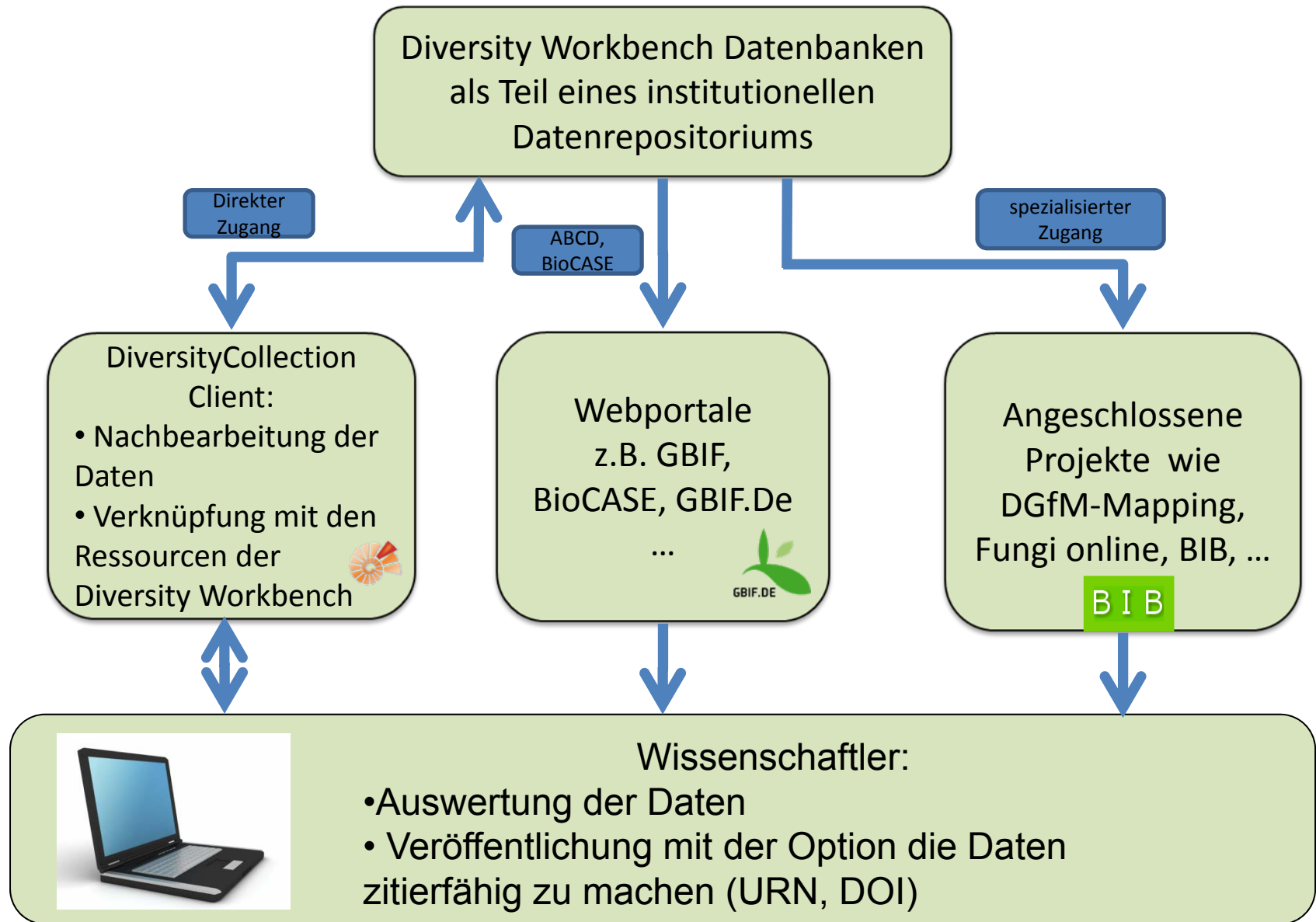
Zustandsübergänge und Konfliktmanagement



- Resolved
 - Benutzer bestimmt selber wie ein Konflikt gelöst wird
 - Daten werden auf beiden Seiten aktualisiert

- Lokalisierung von Konflikten in der Tabelle Fieldstate
- Lokalisierung von Konflikten über Hashwerte für jede Spalte

Zugang zu Daten über die Diversity Workbench



DiversityMobile mit der ***Diversity Workbench***:

- Schaffung eines kontinuierlichen Datenflusses
- Erhöhung der Qualität der Datenaufnahme
- Speicherung von ergänzenden Daten
- Erfüllung der neuen Anforderungen an die Primärdatenspeicherung