



Agenda

- | | |
|--------------------|----------------------------------------------------------------------------------------------|
| 9.30 – 9.50 | Aus Vermutungen werden Fakten -
Einführung in genealogische
Datenmodellierung |
| 9.50 – 10.20 | Eine Lösungsmöglichkeit: Gedbas4all
und das Gentech Data Model |
| 10.20 – 10.40 | Diskussion |
| 10.40 – 11.10 | Pause |
| 11.10 – 11.30 | Das CompGen-Erfassungsprogramm -
Quellenorientierte Datenerfassung |
| 11.30 – 12.00 | Diskussion |



Überblick Teil 1

- praktisches Beispiel
- theoretischer Hintergrund
- kleiner Exkurs: Umgang mit Fehlern
- ergebnisorientiert ↔ quellenorientert

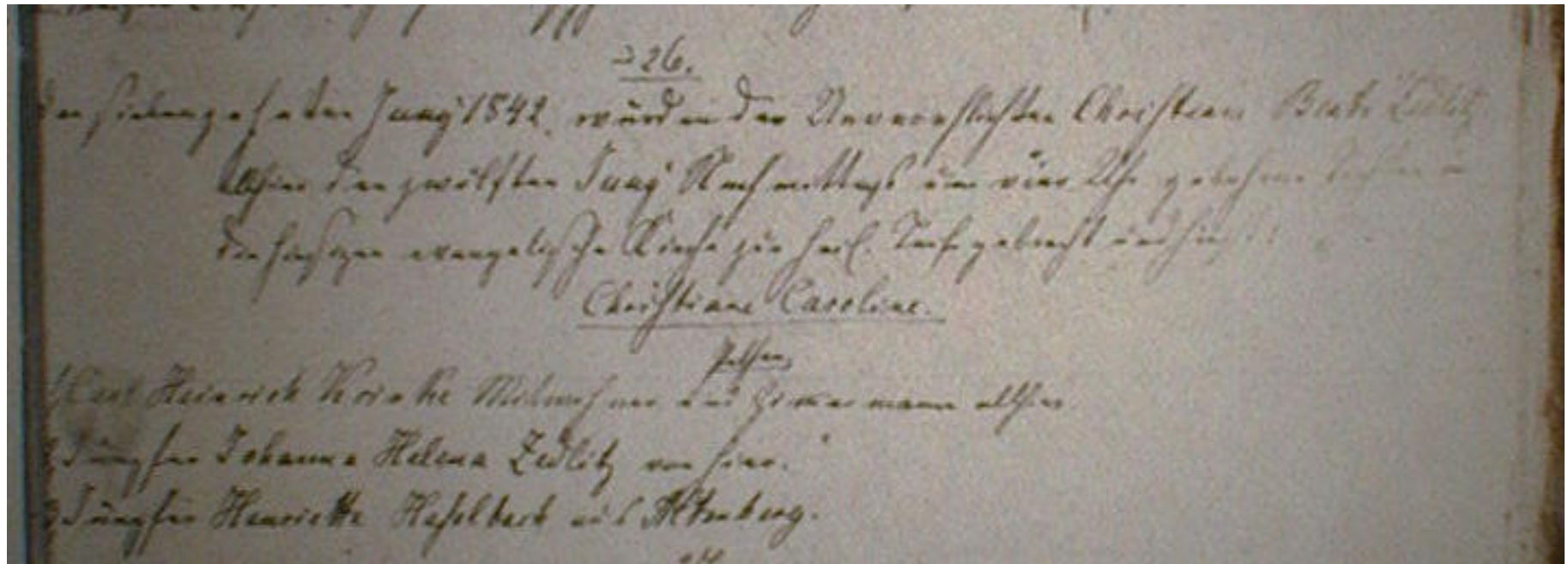


praktisches Beispiel

- ein kleines echtes Beispiel
 - ev. Kirche Seitendorf, Kreis Schönau, Schlesien
 - Kirchenbuch 1794-1870



Taufe 1842



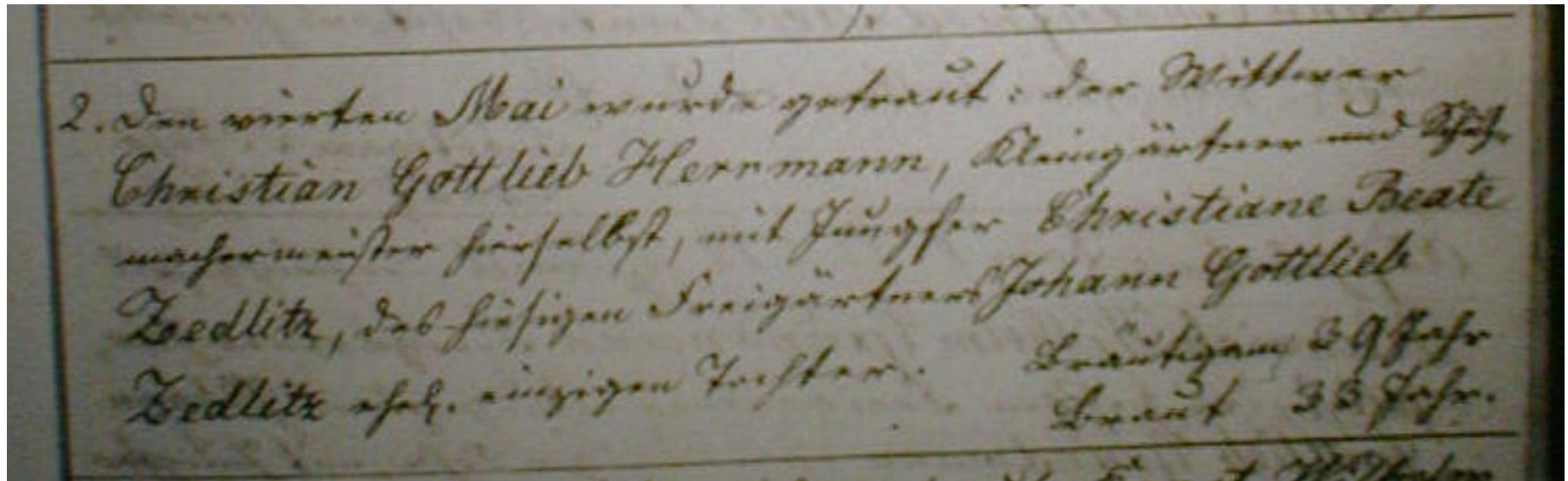
Christiana Beata ZEDLITZ



Christiane Caroline ZEDLITZ
* 1842 Seitendorf



Heirat 1853



Christian Gottlieb HERRMANN
* ca. 1814

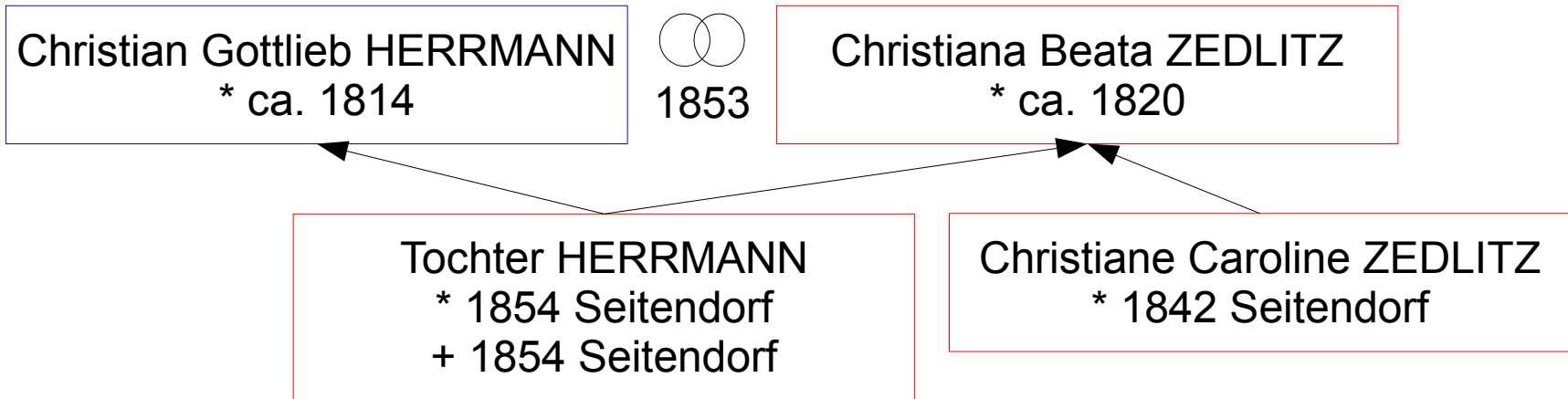
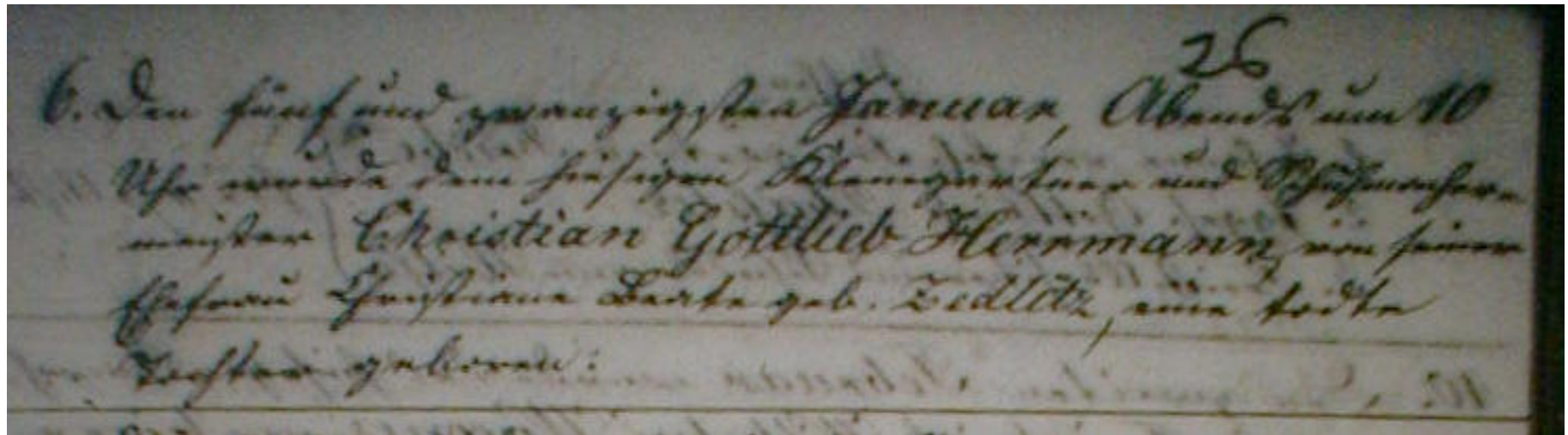
1853

Christiana Beata ZEDLITZ
* ca. 1820

Christiane Caroline ZEDLITZ
* 1842 Seitendorf



Taufe 1854



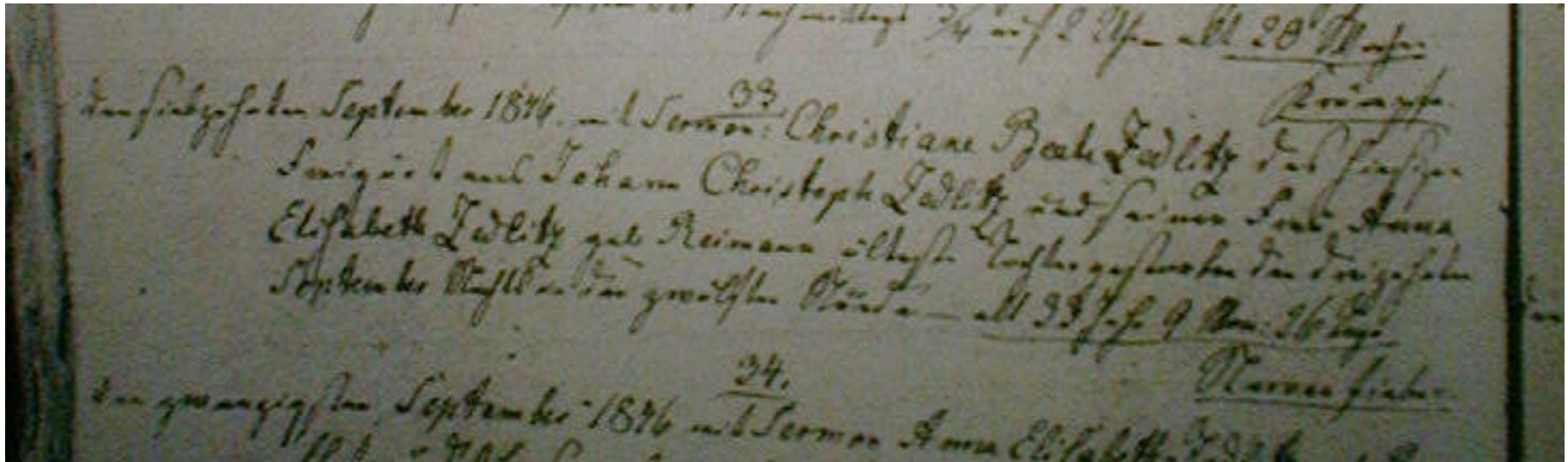


Weitergabe der Daten

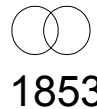
```
0 @I1@ INDI
1 NAME Christiana Beata /ZEDLITZ/
1 FAMS @F1@
0 INDI @I2@
1 NAME Christiane Caroline /ZEDLITZ/
1 FAMC @F1@
0 INDI @I3@
1 NAME Christian Gottlieb /HERRMANN/
1 FAMS @F1@
0 INDI @I4@
1 NAME Tochter /HERRMANN/
1 FAMC @F1@
0 @F1@ FAM
1 WIFE @I1@
1 HUSB @I3@
1 CHIL @I2@
1 CHIL @I4@
```



Beerdigung 1846



Christian Gottlieb HERRMANN
* ca. 1814



1853

Christiana Beata ZEDLITZ
* ca. 1820

Christiana Beata ZEDLITZ
* ca. 1823
+ 1846 Seitendorf

Tochter HERRMANN
* 1854 Seitendorf
+ 1854 Seitendorf

Christiane Caroline ZEDLITZ
* 1842 Seitendorf

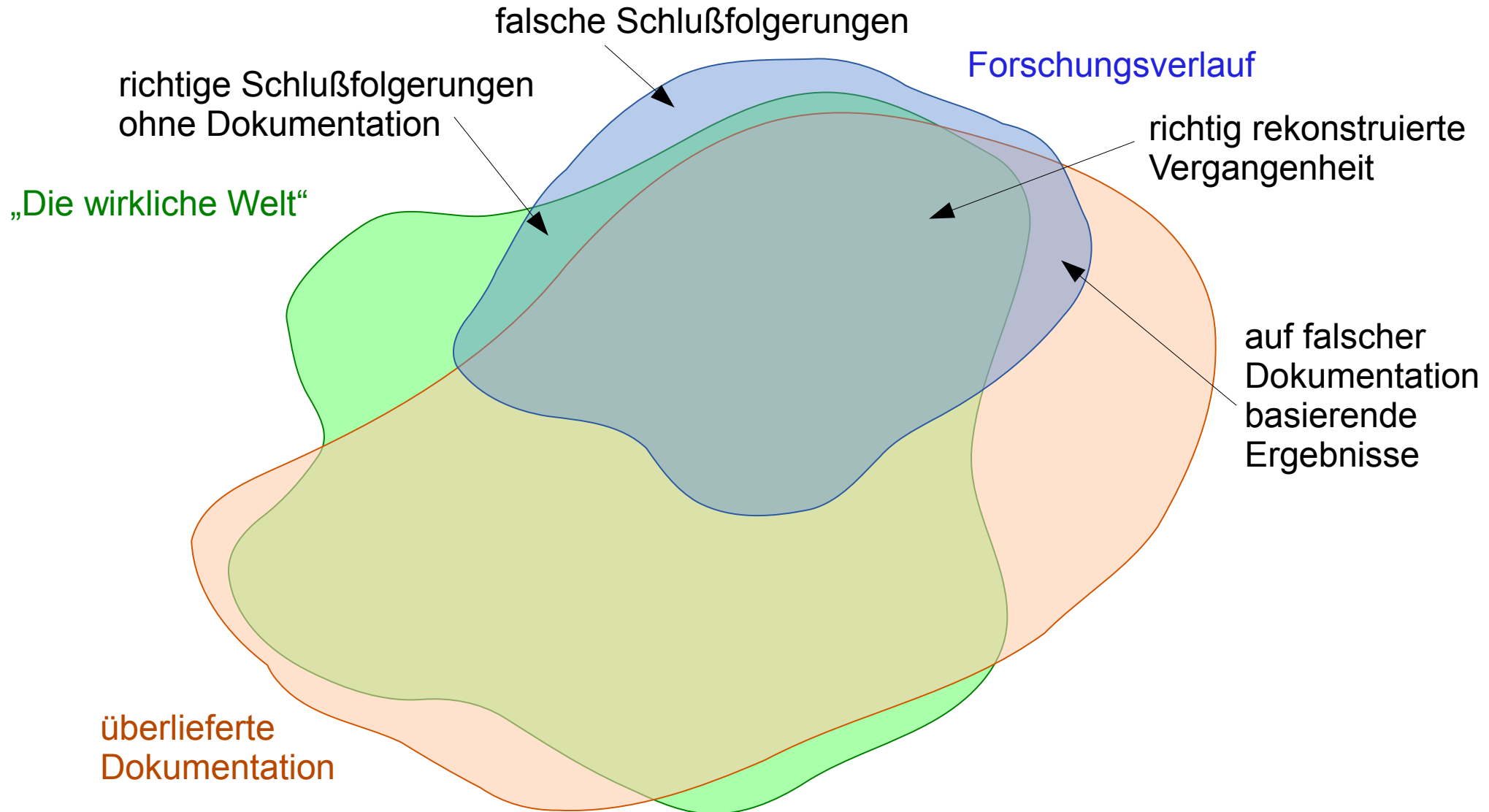


theoretischer Hintergrund

Was ist da passiert?



unser Forschungsgebiet





Fehler sind wichtig

- Falsche Informationen nicht löschen!
- Erspart u.U. erneutes Überprüfen:
 - „Dieses falsche Datum habe ich schon gesehen und es ist wirklich falsch, weil ...“
- Sie können bereits kopiert worden sein.
- Es kann Referenzen auf den Fehler geben.
 - Beim Nachforschen hätte sich die Information geändert.



Fehler und Referenzen

- Eintrag wird in einem Buch referenziert.
- Ein Zahlendreher wird entdeckt und verbessert.
- Beim späteren Überprüfen scheint der Autor des Buches falsch gelesen zu haben.

Ortsfamilienbuch Musterstadt, Person 93.834
Max Maier ist am 6. Juni 1875 geboren.

[...] Max Maier verstarb 1900 im Alter von nur 25 Jahren.¹ [...]

¹ Geburtsdatum laut OFB Musterstadt: 6. Juni 1875
siehe auch <http://ofb.de/Musterstadt/93834>

Ortsfamilienbuch Musterstadt, Person 93.834
Max Maier ist am 6. Juni 1857 geboren.



zwei unterschiedliche Ansätze

ergebnisorientiert

quellenorientiert

typische Begriffe

Person

Familie

Originalquelle

Person(enerwähnung)

elementare Aussage

Schlußfolgerung



ergebnisorientierter Ansatz

- Namen der Person eingeben
- Daten erfassen
- Beziehungen hinzufügen
- Quellen heraussuchen
- Was macht man bei Konflikten?
 - zweites Datum hinzufügen
 - mehrere Eltern, Ehepartner?



quellenorientierter Ansatz

- man beginnt mit Originalquelle
 - digitales Bild oder wortgenaue Abschrift
- elementare Informationen extrahieren
- so entstehen viele „Personen“
- Schlußfolgerungen
 - kennzeichnen
 - lassen sich nachvollziehen („audit trail“)
 - Fehler können leichter korrigiert werden



Nutzen

Wie können wir das für uns nutzen?



Agenda

- 9.30 – 9.50 Aus Vermutungen werden Fakten -
Einführung in genealogische
Datenmodellierung
- 9.50 – 10.20 Eine Lösungsmöglichkeit: Gedbas4all
und das Gentech Data Model**
- 10.20 – 10.40 Diskussion
- 10.40 – 11.10 Pause
- 11.10 – 11.30 Das CompGen-Erfassungsprogramm -
Quellenorientierte Datenerfassung
- 11.30 – 12.00 Diskussion



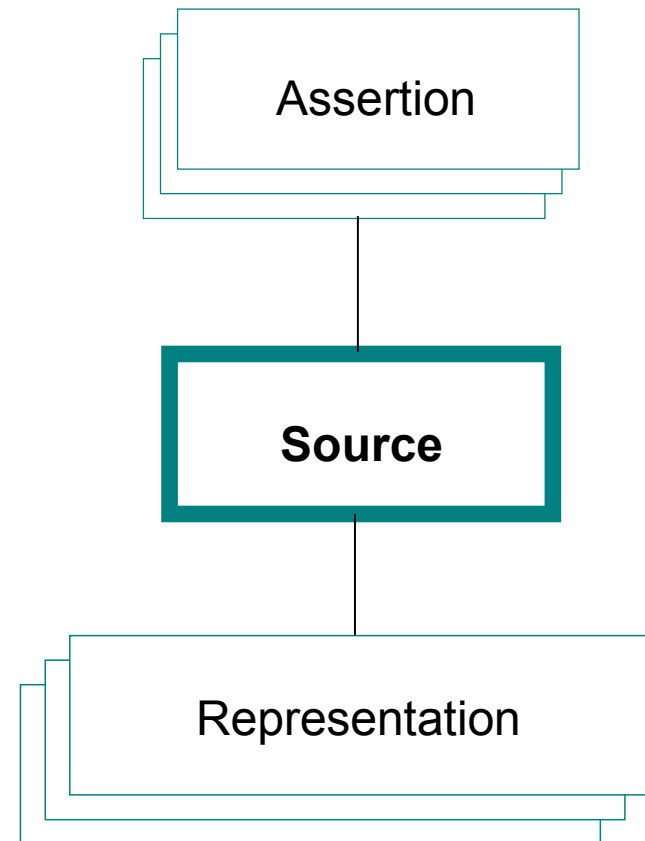
Genealogical Data Model

- GDM – Genealogical Data Model
- quellenorientiert
- entworfen bereits 2000 von GENTECH
- sehr gut durchdacht
- Trennung zwischen
 - administrativen Dingen (Forscher, Projekt, etc.)
 - Nachweisen
 - Schlußfolgerungen
- Teile bereits im Einsatz (bei familysearch.org)



SOURCE

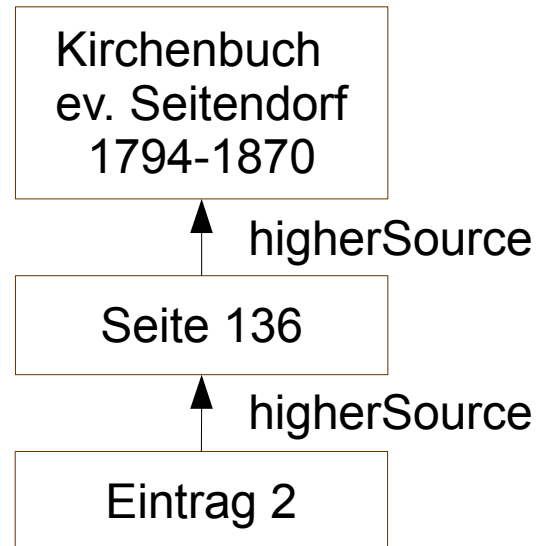
- **Quelle**
- Ausgang jeder Forschung
- z.B. Grabstein, Urkunde, Kirchenbuch
- abstrakt = bezieht sich nicht auf ein konkretes Bild





SOURCE

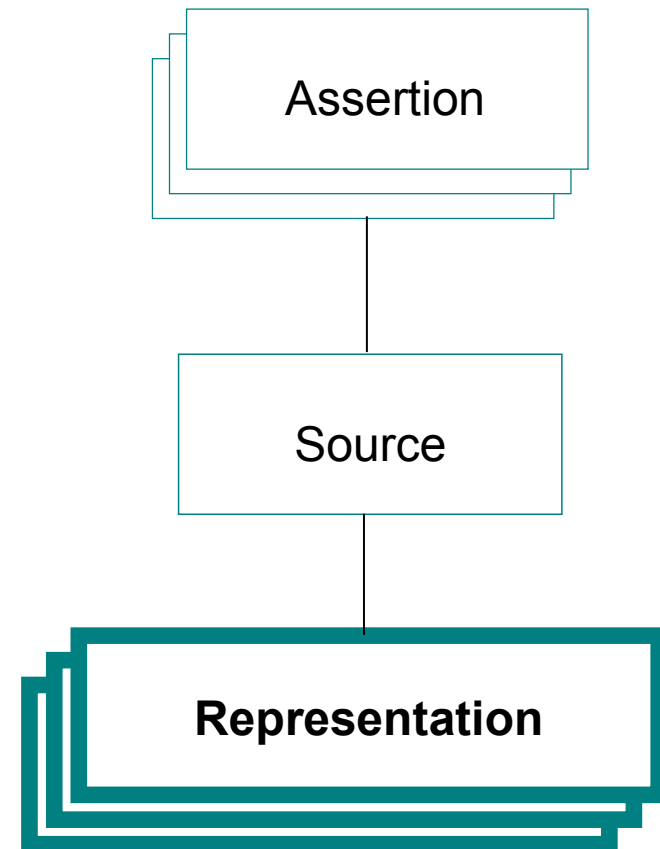
- Quellen sind hierarchisch aufgebaut
 - Buch ← Seite ← Eintrag
 - Akte ← Jahrgang ← Blattnummer ← Zeile
- Beispiel aus der Praxis





REPRESENTATION

- **Abbildung** einer SOURCE
- z.B. Bild, elektronische Abschrift
- auch mehrere zu einer SOURCE





REPRESENTATION

- praktisches Beispiel

Kirchenbuch
ev. Seitendorf
1794-1870

Seite 136

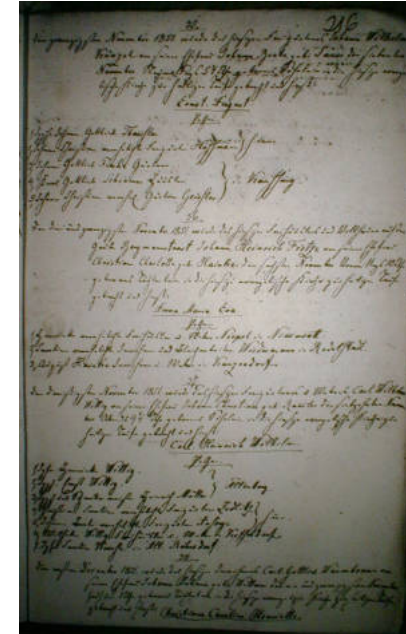
Eintrag 2

source

img_136.jpg

source

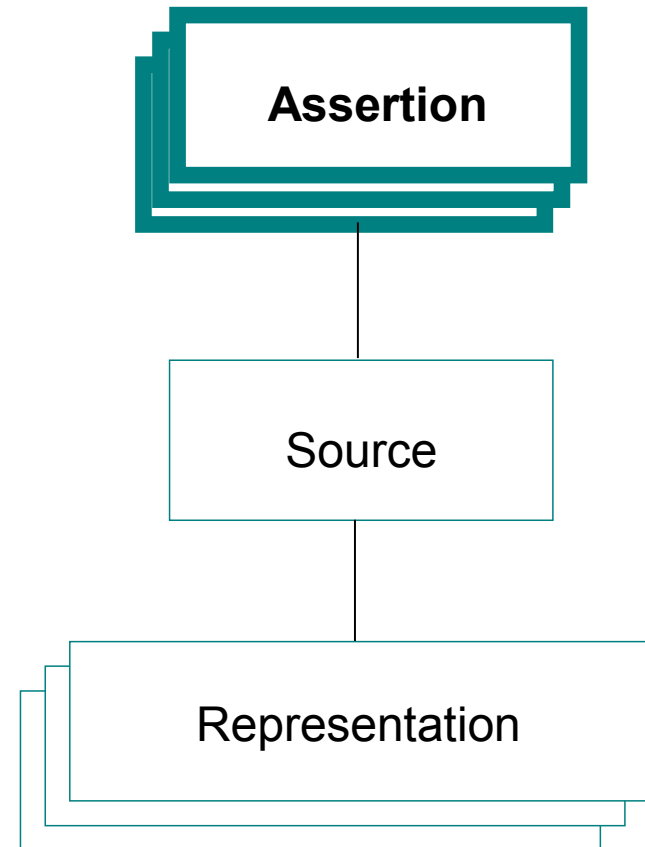
2. Den vierten Mai wurden getraut: der Wittwer Christian Gottlieb Herrmann, Kleingärtner und Schuhmachermeister hieselbst, mit Jungfer Christiane Beate Zedlitz, des hiesigen Freigärtners Johann Gottlieb Zedlitz ehel. einzige Tochter. Bräutigam 39 Jahr, Braut 33 Jahr.





ASSERTION

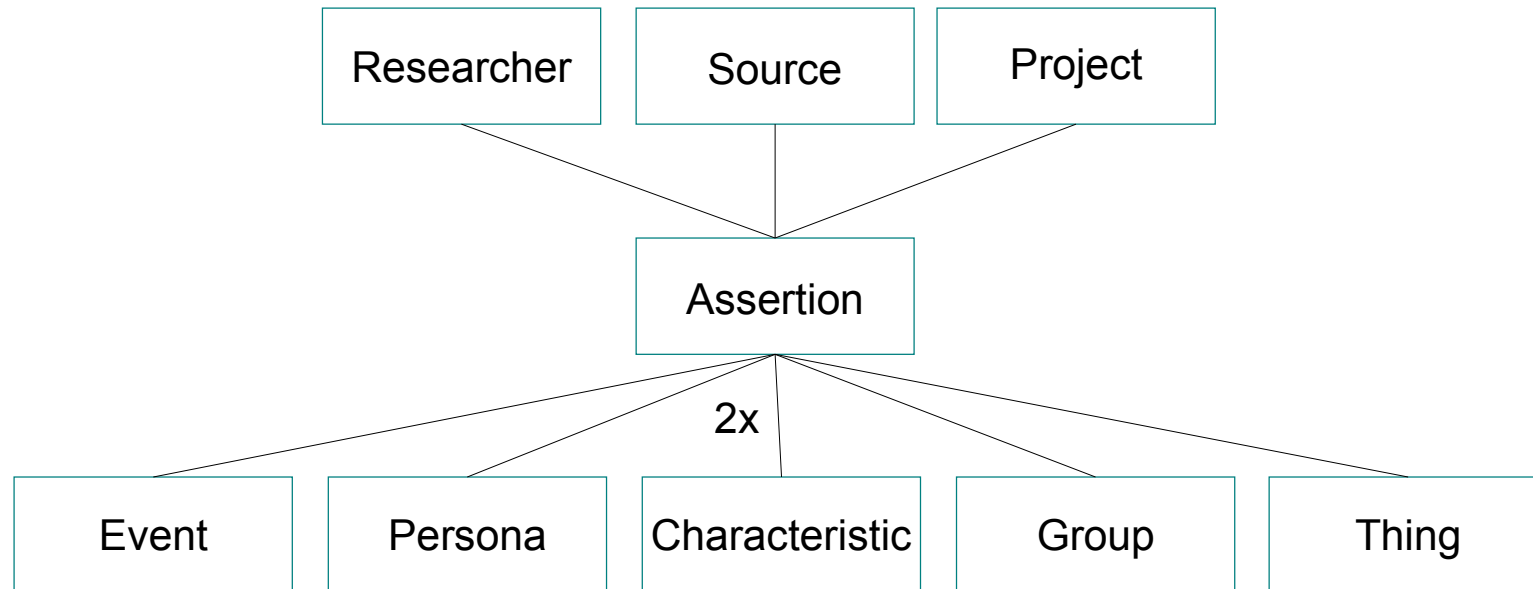
- ausgehen von von Quellen werden **Schlußfolgerungen** gezogen
- aus einer SOURCE typischerweise mehrere ASSERTIONS





ASSERTION

- eine ASSERTION verknüpft zwei Objekten (untere Reihe)
- gibt Auskunft, wer/wann/warum die Schlußfolgerung gezogen hat





EVENT

- Ein **Ereignis** hat zu einer bestimmten Zeit an einem bestimmten Ort stattgefunden.
- Es hat einen genau definierten Typ.
 - Geburt, Heirat, Tod
 - Abfahrt/Ankunft eines Schiffes
 - Wohnen, Arbeiten
- Ereignistypen sind erweiterbar.
- Ereignistypen stehen in Beziehung zueinander.



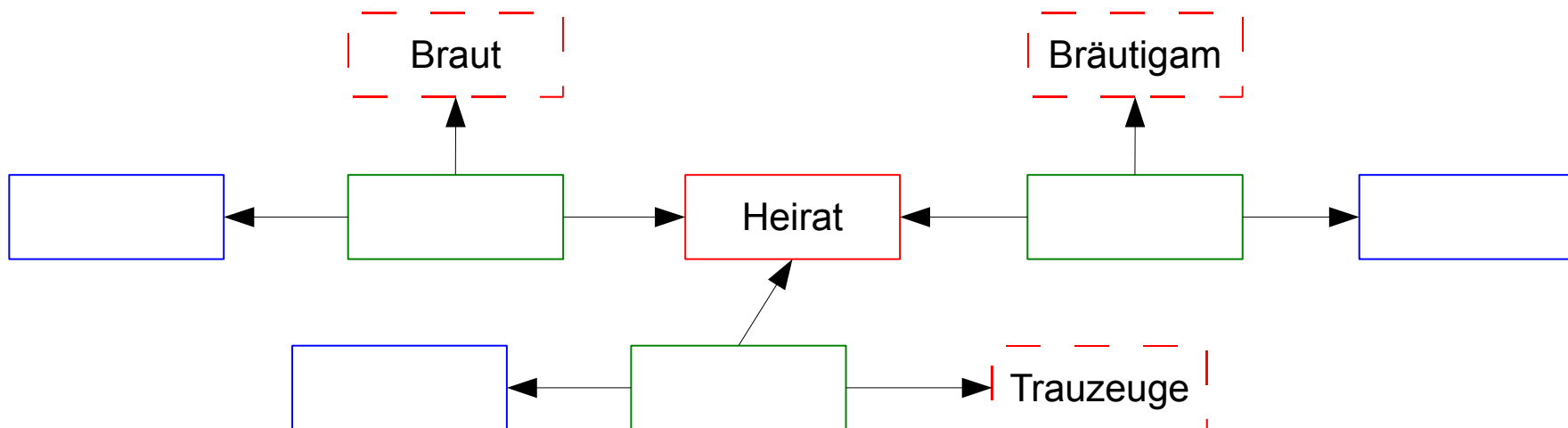
EVENT-TYPE

- Ein Ereignistyp kann eine Spezialisierung eines anderen sein.
 - „kirchliche Heirat“ und „standesamtliche Heirat“ ist eine Spezialisierung von „Heirat“
 - „gefallen“ ist eine Spezialisierung von „gestorben“
- Zur Qualitätssicherung kann man Ereignistypen weitere Informationen mitgeben.
 - „Heirat benötigt zwei Personen.“
 - noch viel Arbeit



EVENT

- PERSONA nehmen bei einem EVENT eine Rolle ein: EVENT_ROLE
 - Bräutigam
 - Braut
 - Hauptperson (engl. „principal“)





PERSONA

- für jedes Vorkommen einer Person in einer Quelle wird eine eigene PERSONA erzeugt

2. Den vierten Mai wurden getraut: der Wittwer Christian Gottlieb Herrmann, Kleingärtner und Schuhmachermeister hieselbst, mit Jungfer Christiane Beate Zedlitz, des hiesigen Freigärtners Johann Gottlieb Zedlitz ehel. einzige Tochter. Bräutigam 39 Jahr, Braut 33 Jahr.

- Christian Gottlieb Herrmann
 - vorherige Frau von Christian Gottlieb Herrmann
 - Christiane Beate Zedlitz
 - Johann Gottlieb Zedlitz
-



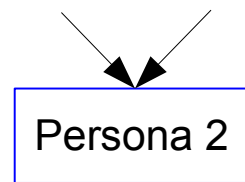
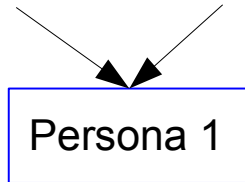
PERSONA

- Nie werden an eine bestehende PERSONA Informationen angehängt.
- Es wird eine neue PERSONA erzeugt.
- Vermutlich identische PERSONAs werden mit einer GROUP verknüpft.
- Aus der GROUP wird eine neue PERSONA erzeugt.
 - Sie ist so durch keine Quelle belegt.
- Fehler lassen sich so leicht korrigieren.



PERSONA

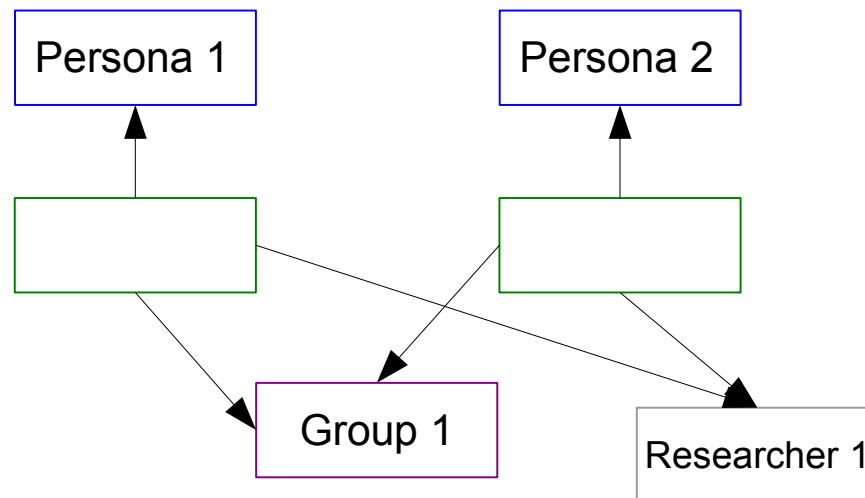
- eine Person taucht in zwei Quellen auf und soll als identisch markiert werden:





PERSONA

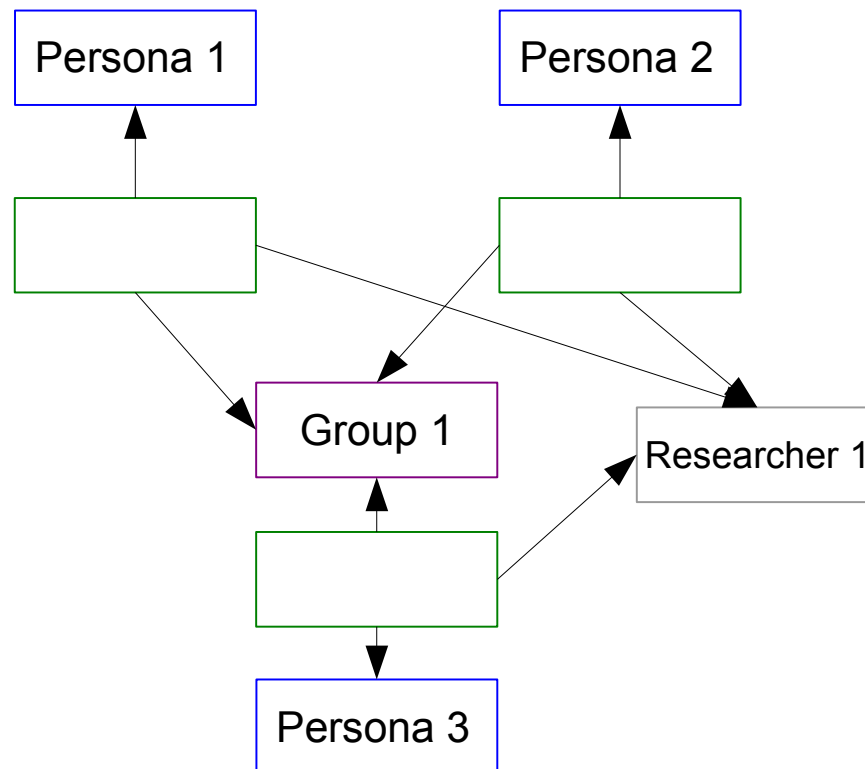
- beide PERSONAs werden zu einer Gruppe zusammengefaßt:





PERSONA

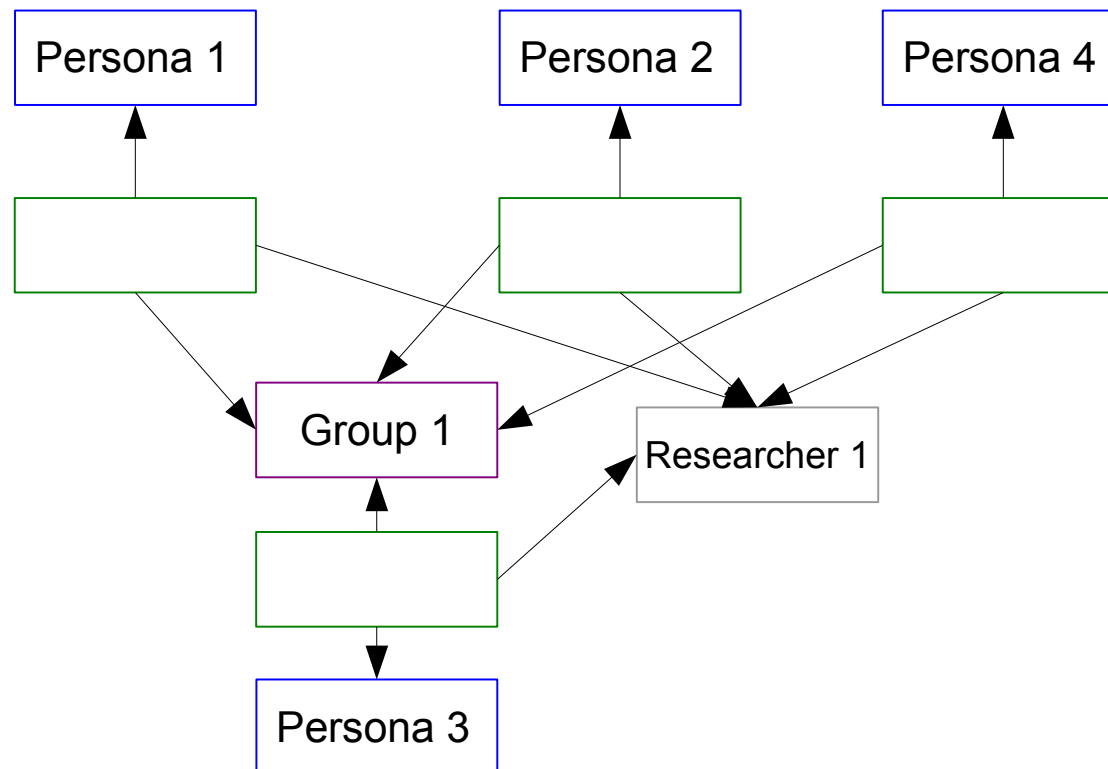
- aus der Gruppe wird eine PERSONA erzeugt:





PERSONA

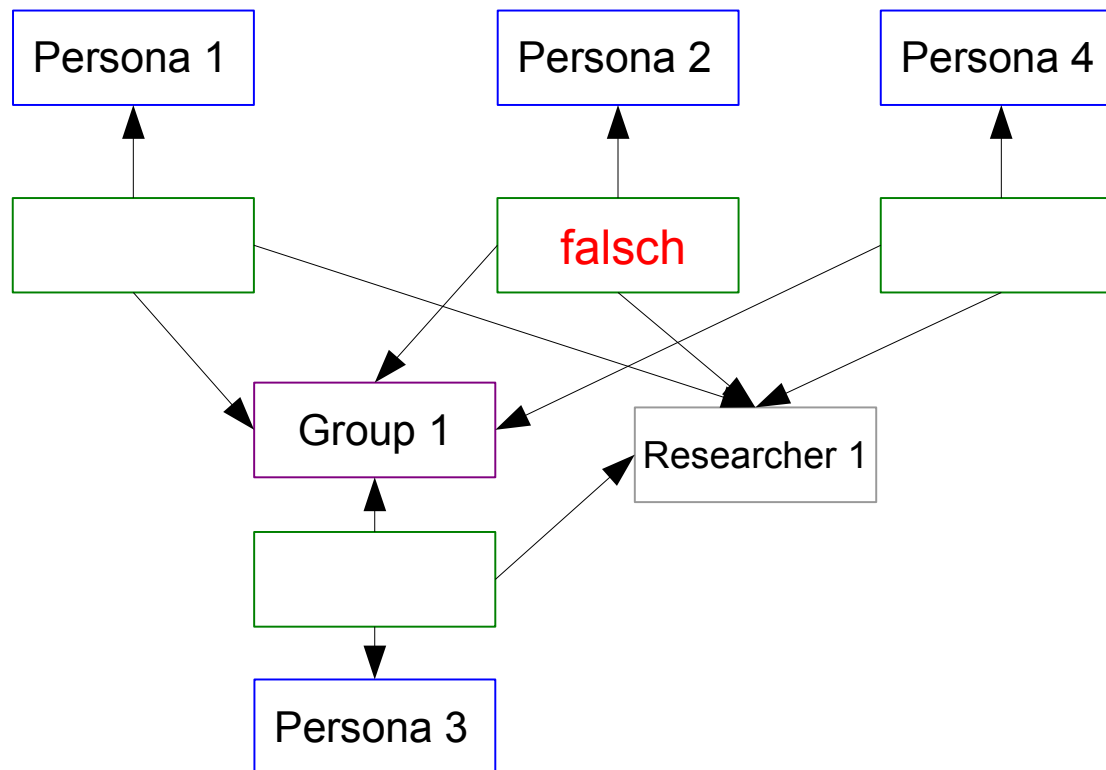
- PERSONAs aus neuen Quellen werden der Gruppe hinzugefügt:





PERSONA

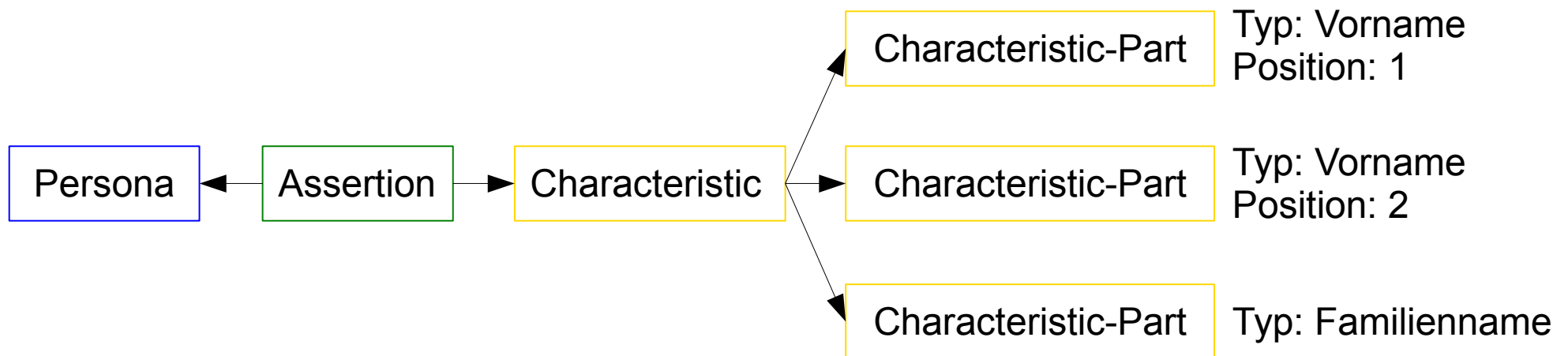
- bemerkt man später einen Fehler, wird die ASSERTION als falsch markiert





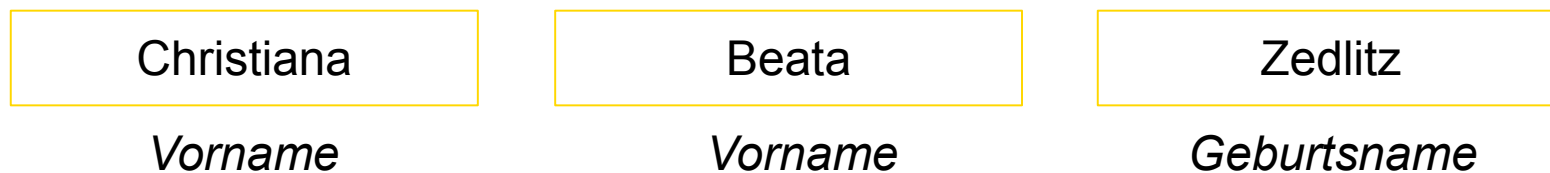
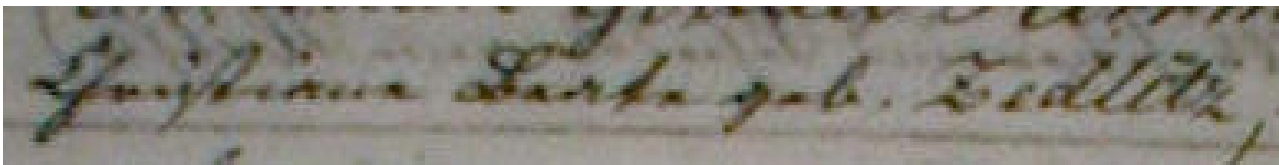
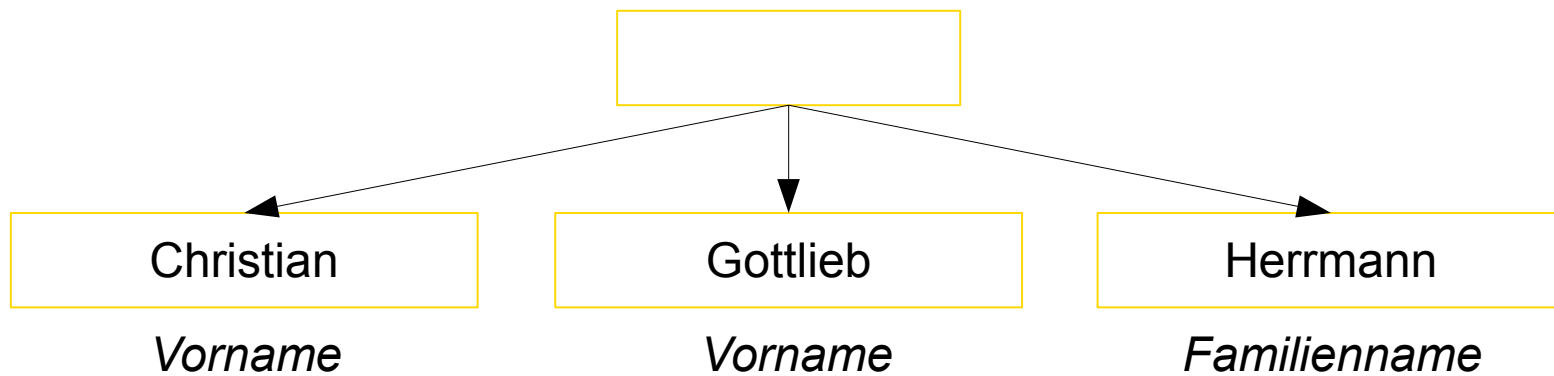
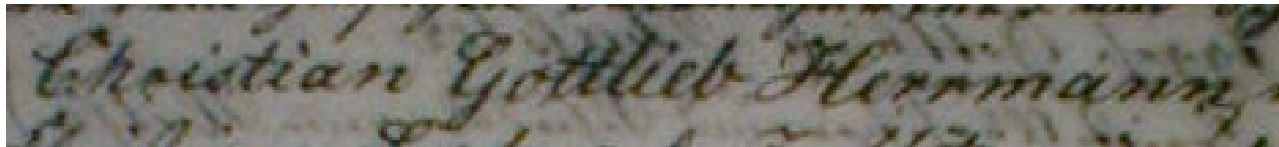
CHARACTERISTIC

- **Eigenschaft** von Personen, Dingen, etc.
- grob gesagt: Textinformationen
- Sie hat einen genau definierten Typ.
- besteht aus mehreren (geordneten) Teilen, den **CHARACTERISTIC-PARTs**





CHARACTERISTIC





GROUP

- Gruppen bestehen aus PERSONA, GROUP, CHARACTERISTIC, EVENT oder THING
- Beispiele für Gruppen:
 - PERSONA: Kinder eine Person
 - THING: Häuser, die bei einer Feuer zerstört wurden
- Funktion innerhalb der Gruppe wird über GROUP_ROLE beschrieben:
 - Kind
 - Mutter

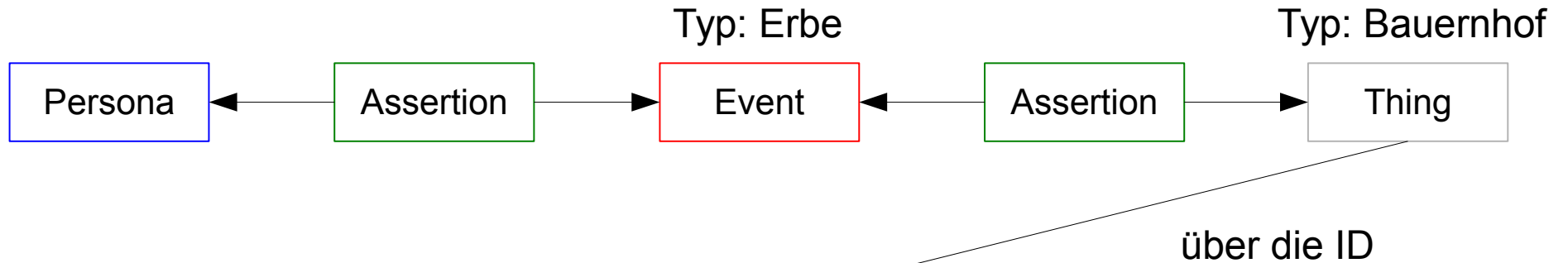


THING

- nicht im GDM
- ein „Gegenstand“, z.B. Schiff
- nimmt an Ereignissen teil (Abfahrt, Ankunft)
- erweiterte Angaben im GenWiki, auch Fotos
- über einen definierten Typ `THING_TYPE` wird die Art des Gegenstands festgelegt
- kann evtl. auch für juristische Personen genutzt werden



Verweis ins GenWiki



Schönberg (Schleswig-Holstein)/Vollhufe 3 – GenWiki - Konqueror

Home News Regional Vereine Datenbanken Software Mailinglisten Links Hilfen/FAQs

jzedlitz Eigene Diskussion Einstellungen Beobachtungsliste Eigene Beiträge Abmelden

Seite Diskussion Bearbeiten Versionen/Autoren schützen Löschen Verschieben

Schönberg (Schleswig-Holstein)/Vollhufe 3

< Schönberg (Schleswig-Holstein)

Ehemalige Hausnummer: 9 und 10, jetzt Niederstraße 11 bzw. Aussiedlung Höndorfer Tor.

1. um 1483 Tewes Wiese
2. seit 1550 Albert Wiese
3. seit 1607 Sievert Wiese
4. seit 1642 Jochim Wiese
5. seit 1681 Claus Wiese
6. seit 1742 Peter Wiese
7. seit 1770 Claus Wiese
8. seit 1812 Hinrich Lage jun., gest. 1825
9. seit 1828 Lochim Arp 1, gest. 1870, und Frau Trien Arp, geb. Mundt
10. Jochim Arp 2, gest. 1902
11. Johannes Arp, gest. 1930
12. Reinhold Arp, gest. 1966

ca. 1910



mögliche ASSERTIONS

		SUBJECT 2				
		PERSONA	EVENT	GROUP	CHARAC- TERISTIC	THING
SUBJECT 1	PERSONA		✓	✓	✓	
	EVENT		✓	✓		
	GROUP	✓		✓		✓
	CHARAC- TERISTIC			✓		
	THING		✓	✓	✓	



mögliche Assertions

		SUBJECT 2				
		PERSONA	EVENT	GROUP	CHARAC- TERISTIC	THING
SUBJECT 1	PERSONA		✓	✓	✓	
	EVENT		✓	✓		
	GROUP	✓		✓		✓
	CHARAC- TERISTIC			✓		
	THING		✓	✓	✓	

PERSONA → PERSONA wäre denkbar („P1 ist Vater von P2“), es ist jedoch besser, dies über Gruppen abzubilden



mögliche Assertions

		SUBJECT 2				
		PERSONA	EVENT	GROUP	CHARAC- TERISTIC	THING
SUBJECT 1	PERSONA		✓	✓	✓	
	EVENT		✓	✓		
	GROUP	✓		✓		✓
	CHARAC- TERISTIC			✓		
	THING		✓	✓	✓	

Ereignis im Leben einer PERSONA: „Ist Braut bei Hochzeit.“



mögliche Assertions

		SUBJECT 2				
		PERSONA	EVENT	GROUP	CHARAC- TERISTIC	THING
SUBJECT 1	PERSONA		✓	✓	✓	
	EVENT		✓	✓		
	GROUP	✓		✓		✓
	CHARAC- TERISTIC			✓		
	THING		✓	✓	✓	

PERSONA ist Mitglied der GROUP: „gehört zum Dritten Hannoverschen Infanterie-Regiment“



mögliche Assertions

		SUBJECT 2				
		PERSONA	EVENT	GROUP	CHARAC- TERISTIC	THING
SUBJECT 1	PERSONA		✓	✓	✓	
	EVENT		✓	✓		
	GROUP	✓		✓		✓
	CHARAC- TERISTIC			✓		
	THING		✓	✓	✓	

PERSONA hat eine bestimmte Eigenschaft: „Die Person hatte rote Haare.“



Anzeige

- Wird die Anzeige nicht zu kompliziert?
- Lösung: Informationen zusammenfassen
- solange Schlußfolgerungen nicht widersprüchlich sind, kann kombiniert werden:

Zeitungsanzeige: „Paul wurde am 28. Mai 1985 geboren.“

Geburtsurkunde: „Paul Heinrich wurde am 28. Mai 1985 geboren.“

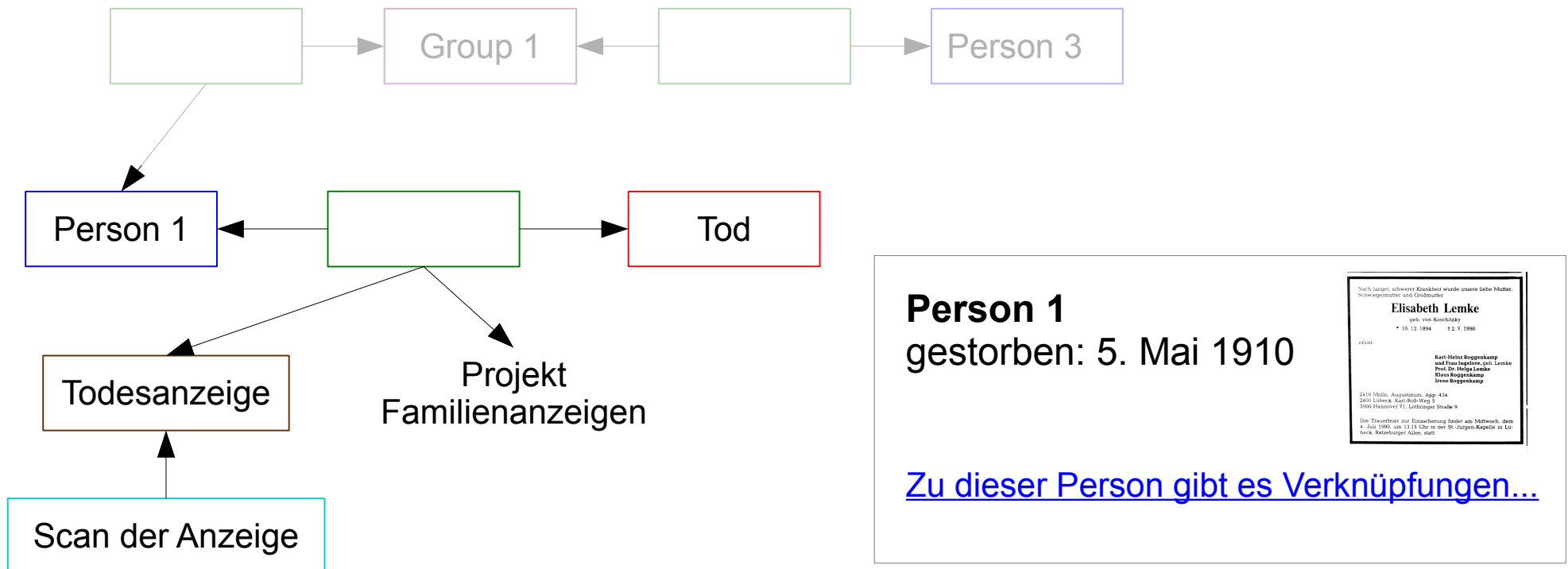
Zeitung 2003: „Wir gratulieren unserem Sohn Paul zum 18. Geburtstag.“

⇒ **Geburt: 1985-05-28 [Quelle1] [Quelle2] [Quelle3]**



Anzeige

- Begrenzung auf einzelne Projekte
- Suche nur Daten aus Projekt Familienanzeigen





Agenda

9.30 – 9.50	Aus Vermutungen werden Fakten - Einführung in genealogische Datenmodellierung
9.50 – 10.20	Eine Lösungsmöglichkeit: Gedbas4all und das Gentech Data Model
10.20 – 10.40	Diskussion
10.40 – 11.10	Pause
11.10 – 11.30	Das CompGen-Erfassungsprogramm - Quellenorientierte Datenerfassung
11.30 – 12.00	Diskussion



Agenda

9.30 – 9.50	Aus Vermutungen werden Fakten - Einführung in genealogische Datenmodellierung
9.50 – 10.20	Eine Lösungsmöglichkeit: Gedbas4all und das Gentech Data Model
10.20 – 10.40	Diskussion
10.40 – 11.10	Pause
11.10 – 11.30	Das CompGen-Erfassungsprogramm - Quellenorientierte Datenerfassung
11.30 – 12.00	Diskussion



Agenda

- | | |
|----------------------|-------------------------------------------------------------------------------------|
| 9.30 – 9.50 | Aus Vermutungen werden Fakten -
Einführung in genealogische
Datenmodellierung |
| 9.50 – 10.20 | Eine Lösungsmöglichkeit: Gedbas4all
und das Gentech Data Model |
| 10.20 – 10.40 | Diskussion |
| 10.40 – 11.10 | Pause |
| 11.10 – 11.30 | Das CompGen-Erfassungsprogramm -
Quellenorientierte Datenerfassung |
| 11.30 – 12.00 | Diskussion |



Eingabeprogramm

Wie kommen Daten in Gedbas4all?



Import bestehender Daten

- Zahlreiche erfaßte Daten lassen sich gut übernehmen.
 - Adreßbücher
 - Kirchenbuchverkartungen (z.B. der WGfF)
- strukturierte Daten
- klarer Bezug zur Quelle
- keine Schlußfolgerungen bzw. deutliche gekennzeichnet

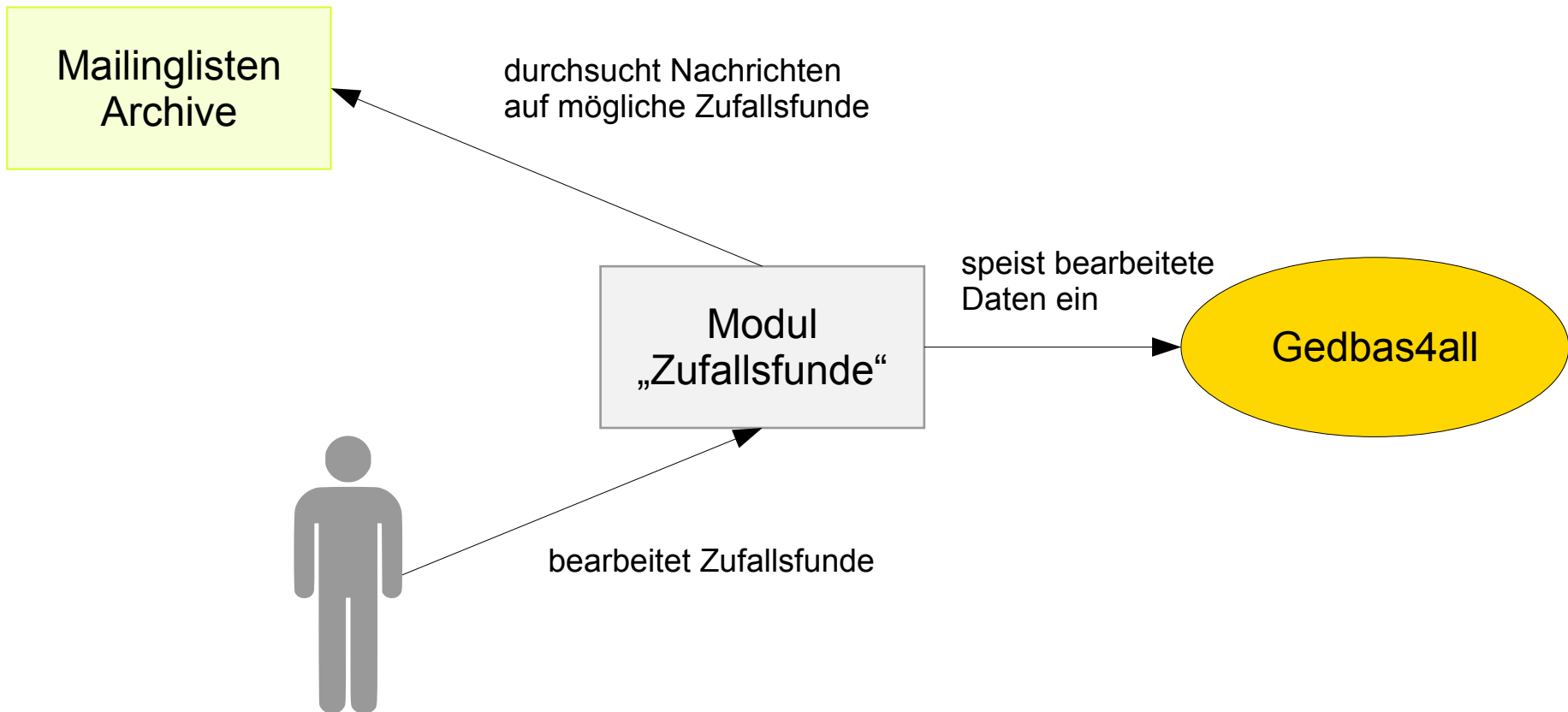


Und was ist mit GEDCOM?

- gut geeignet für die elektronische Speicherung von Stammbäumen als Forschungsergebnis
- Vorsicht beim Import
 - keine Schlußfolgerungen
 - klarer Bezug zur Quelle
 - keine Informationen aus verschiedenen Quellen mischen
- ein Versuch in der DigiBib: Buch in eine GEDCOM-Datei zu übersetzen
- sonst: Datei ist Quelle (wie gedrucktes Buch)



Eingabe von Zufallsfunden





Erfassungsprogramm

- parallele Bearbeitung von gescannten Dokumenten
- aus Bildern werden (tabellarische) Informationen abgetippt
- ähnlich wie FamilySearchIndexing
 - nur Personendaten, keine Ortsverzeichnisse etc.
 - keine DjVu-Dateien
 - nur große Arbeitspakete
 - Rechte an gescannten Bildern



Bearbeitungsschritte

Client

Server

1. Bearbeiter wählt Projekt aus

2. Programm fordert neues Bild an



3. unbearbeitetes Bild suchen

4. als „in Bearbeitung“ markieren



5. Server schickt Bild

5. Bearbeiter tippt eine Seite ab

6. Programm sendet Daten an Server



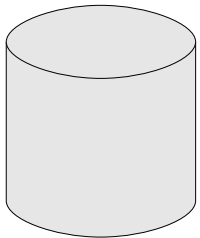
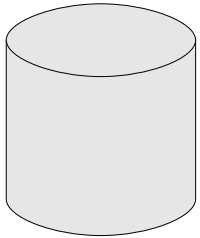
7. Daten abspeichern

8. Bild als „bearbeitet“ markieren



Architektur

Arbeitspakete,
erfaßte Daten



Bilder

Arbeitspaket-
verwaltung

Server

Webanwendung
(HTML)

Webanwendung
(Flash)

Server

Java Webstart
Programm

Windows-
Programm

Browser

Client

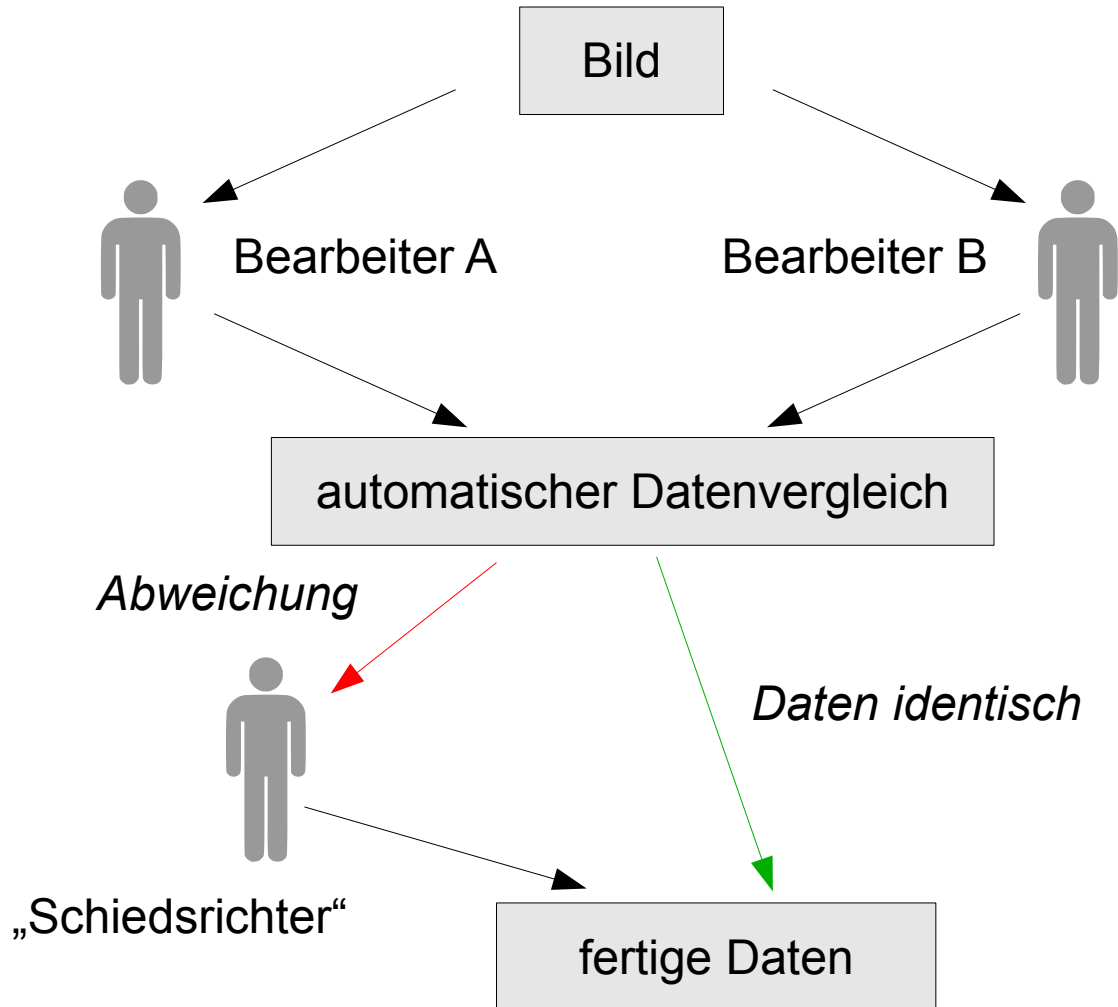


Eingabeverfahren

Wie erreichen wir eine hohe Qualität der abgetippten Informationen?



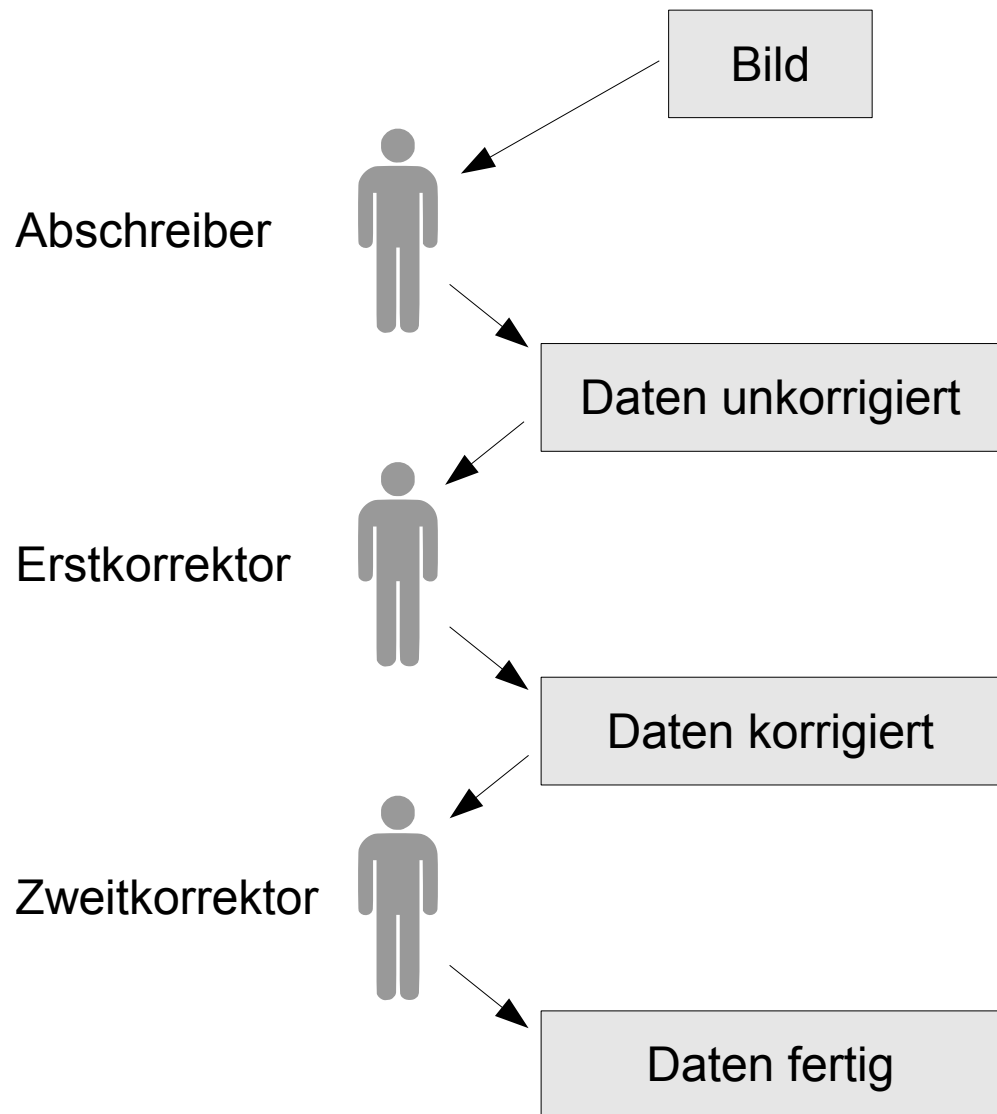
„Double Keying“



- extrem genau
- funktioniert auch mit unzuverlässigen Bearbeitern
- mehr als doppelter Aufwand



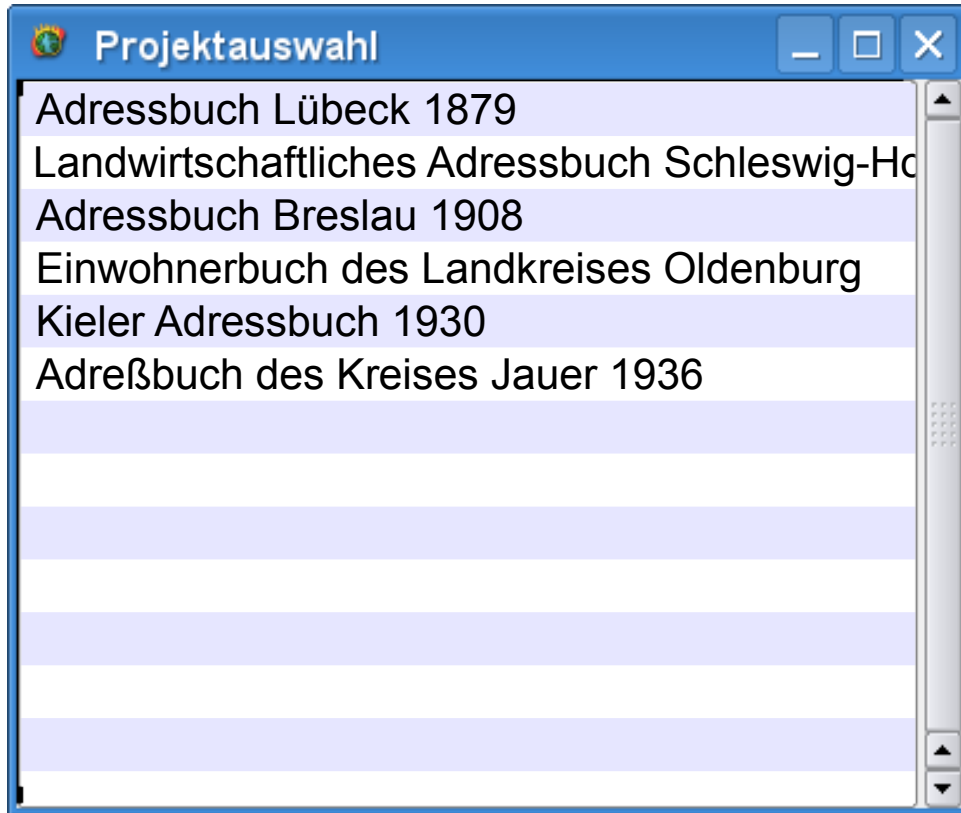
Doppeltes Korrekturlesen



- bekannt aus Wikisource und DigiBib
- keine „doppelte Arbeit“
- evtl. schneller
- jeder Korrektor kann neue Fehler eingeben
- schneller durch einfache Korrektur → schlechtere Qualität



Projektauswahl



- aus der Liste der Projekte wählt der Bearbeiter eines aus



Nutzen

- Wer profitiert davon?
- alle Projekte mit (einfach) strukturierten Daten
- Daten müssen bereits digital vorliegen
- bei uns sind das...



Nutzen

- Adreßbuch-Datenbank
 - komplizierter Umgang mit Dateien entfällt
 - für Geleichenheitshelfer
 - bessere Qualität
- DigiBib
 - bei Büchern mit Tabellen
- Familienanzeigen
 - gescannte, aber noch nicht getippte Anzeigen
 - Arbeit ließe sich trennen: Scannen und Tippen



andere Vereine

- Grabsteinprojekt der Maus
- Kirchenbuchverkartung der WGfF
- Haltermannkartei der Maus
- Volkszählungen der AGGSH
- Karteikarten in Vereinsarchiven
- ...



neue Möglichkeiten

- Verlustlisten 1. WK
- alle Arten existierender genealogischer Register
 - Schuberts Trauregister
- historische Verzeichnisse
 - Lehrer, Schüler, Pastoren, Soldaten
- ...