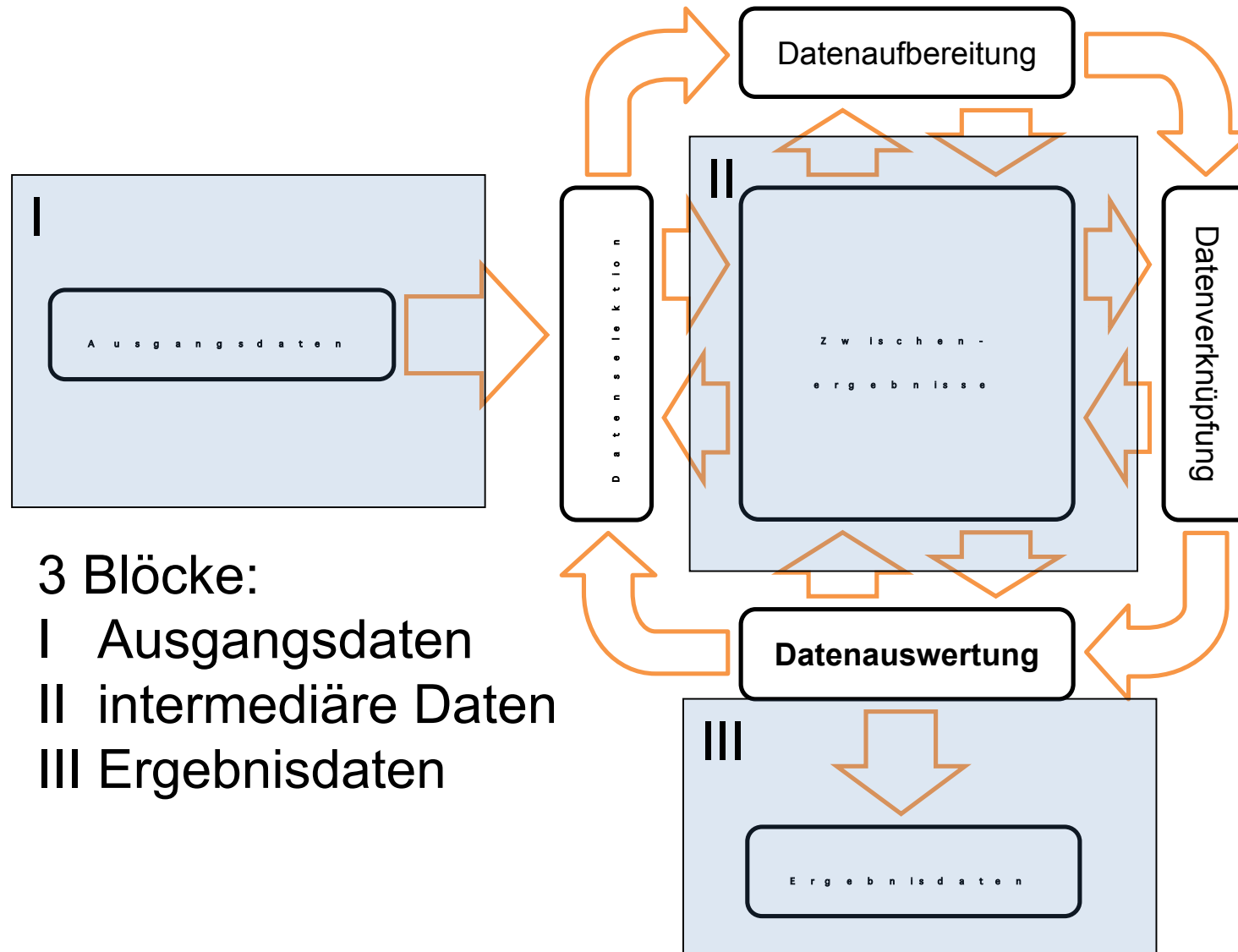

Workshop

Virtuelle Arbeitsumgebung für sozio-ökonomische Forschung und Berichterstattung

Göttingen, 05. November 2010

Harry Enke (AIP)

Datenbereitstellung und –sicherheit



- 3 Blöcke:
I Ausgangsdaten
II intermediäre Daten
III Ergebnisdaten

Datenbereitstellung und –sicherheit: I

Mikrodaten aus

- Personen- und Haushaltssurveys
- Verwaltungsregistern
- Unternehmensbefragungen
- Erhebungen, etc.

A u s g a n g s d a t e n

Quellen:

- Forschungsdatenzentren
- Statistische Ämter
- Forschungseinrichtungen
- Verwaltungen, etc.

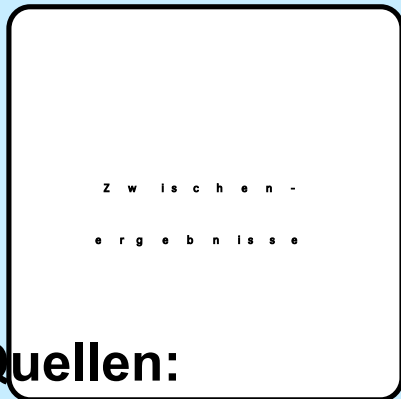
Sicherheit der Daten in der VAU:

1. Nutzer-basierte Zugriffsrechte
 - Individuelle Rechte
 - Gruppenzugehörigkeiten
 - keine rollenbasierten Zugriffe
 - Zertifikats basierte Infrastruktur
2. Metadaten mit Informationen über Beschränkungen
 - zeitlich
 - personengebunden
3. Lokalität der Daten
 - Service Level Agreements mit einzelnen Ressourcenprovidern für
 - Storage
 - Computing

Datenbereitstellung und –sicherheit: II

Typen:

- Arbeitsdatensätze
- Syntax-Dateien
- Grafiken, Dokumente
- Programm-spezifische Dateien



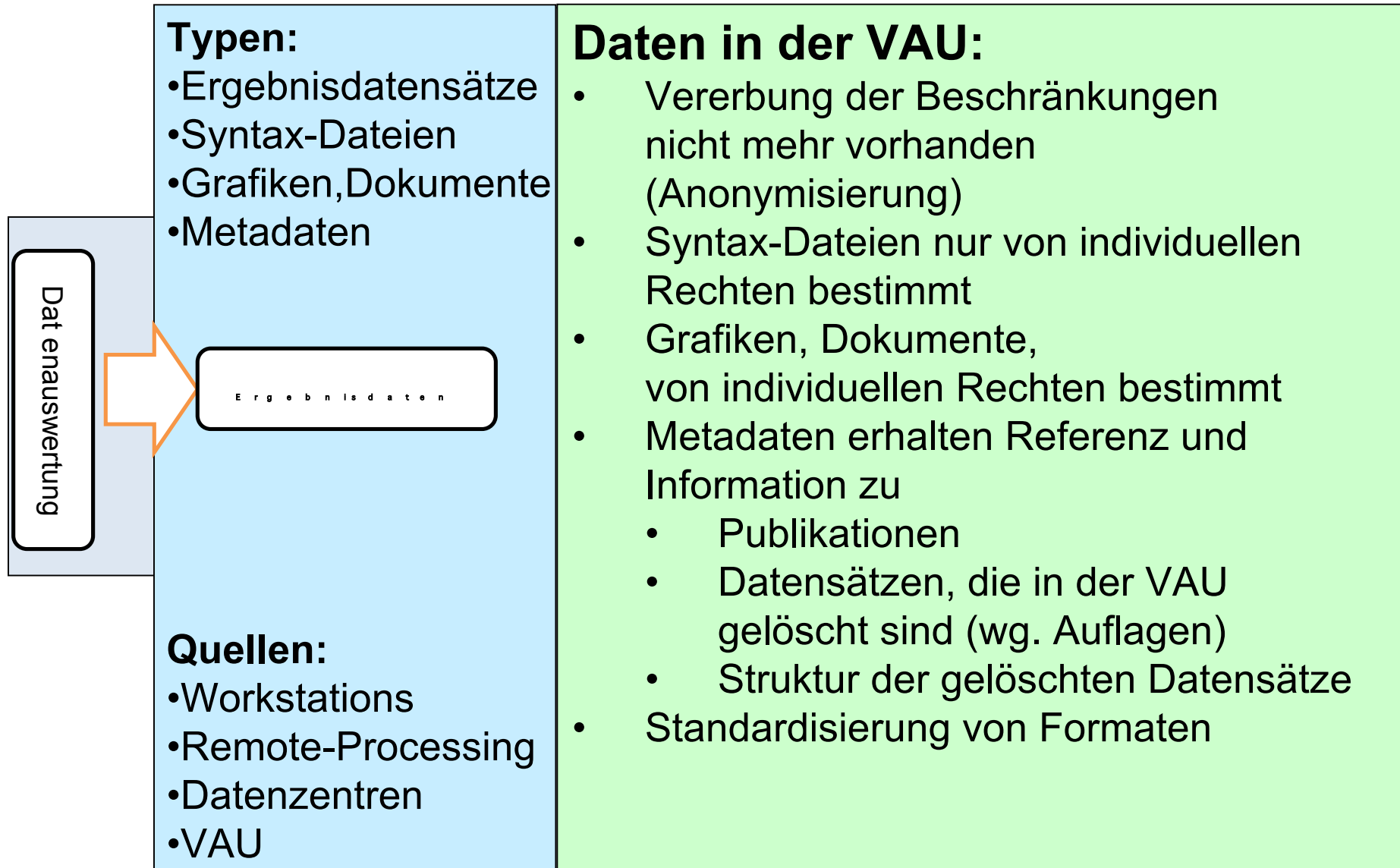
Quellen:

- Workstation
- Remote Processing
- OnSite Processing
- Datenzentren
- VAU

Daten in der VAU:

- Vererbung der Beschränkungen
Problem: wann sind Daten
hinreichend unkenntlich, um
Beschränkungen wegzunehmen?
- Syntax-Dateien:
nur von individuellen Rechten bestimmt
- Grafiken, Dokumente,
Programm-Output
von individuellen Rechten bestimmt
oder
noch fortwirkende Beschränkungen?
- MetaDaten
- Lokalität der Daten
Umgang mit Replikaten

Datenbereitstellung und –sicherheit: III

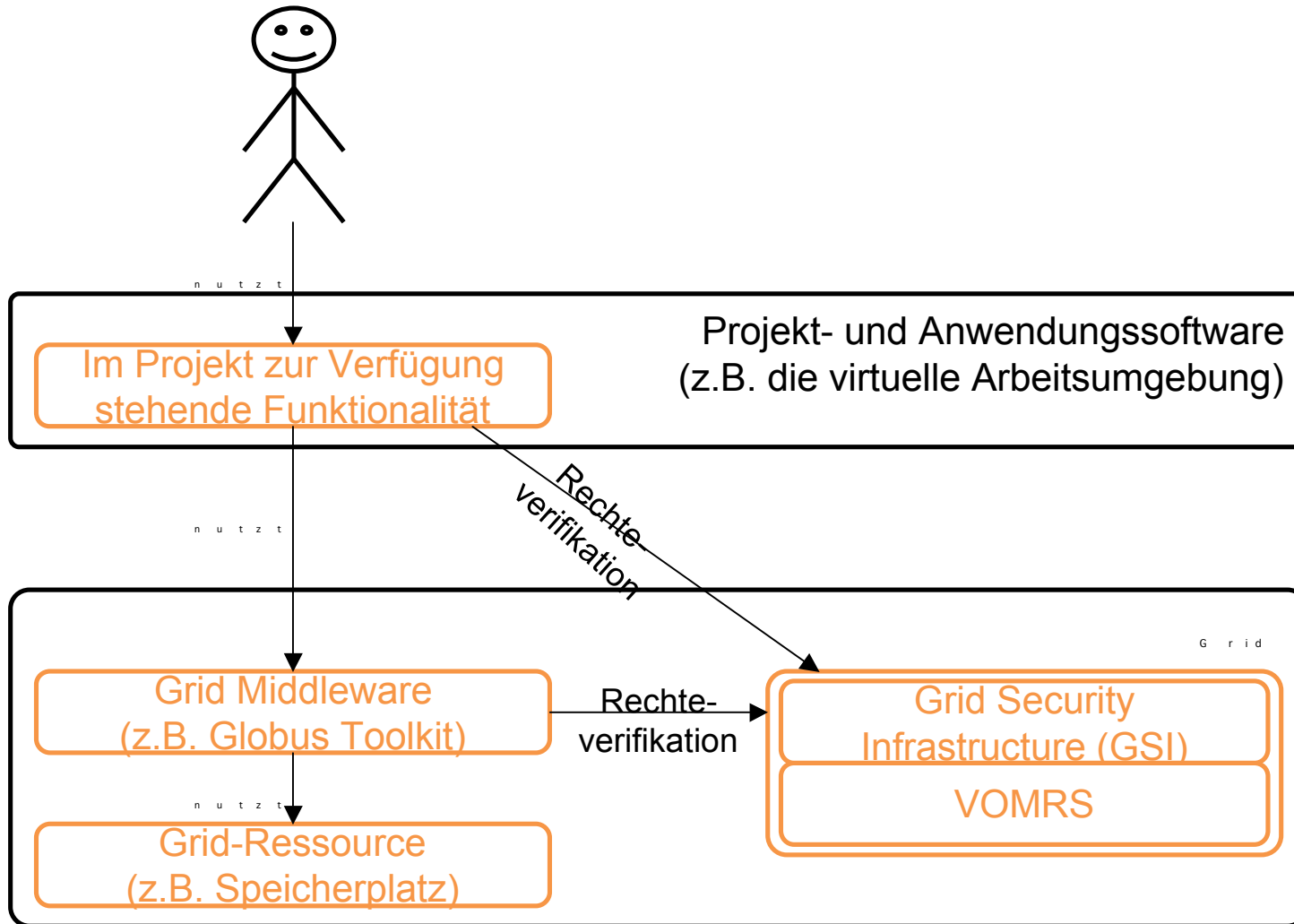


Datenbereitstellung und –sicherheit: DGrid

Grid Security Infrastruktur : (Middleware)

- Zertifikatsbasierte Authentifizierung
 - Durch zeitlich limitierte Proxy-Zertifikate Single-Sign-On
- Vorhandener „Chain of Trust“ durch GridKA und DFN mit (lokalen, universitären oder institutsweiten) Registration Authorities (RA) für Zertifikate
- Nutzer basierte Autorisierung auf (lokaler) Grid-Resource
- Gruppen- und Nutzerrechte bei lokaler Autorisierung
- Virtuelle Organisationen (VO) mit einer (hierarchischen) Struktur
- Einfaches Management der Mitgliedschaft in verschiedenen Substrukturen der VO über den Membership Registration Service (VOMRS)
- Verwaltung des VOMRS durch SOEB/Community

Datenbereitstellung und –sicherheit: DGrid

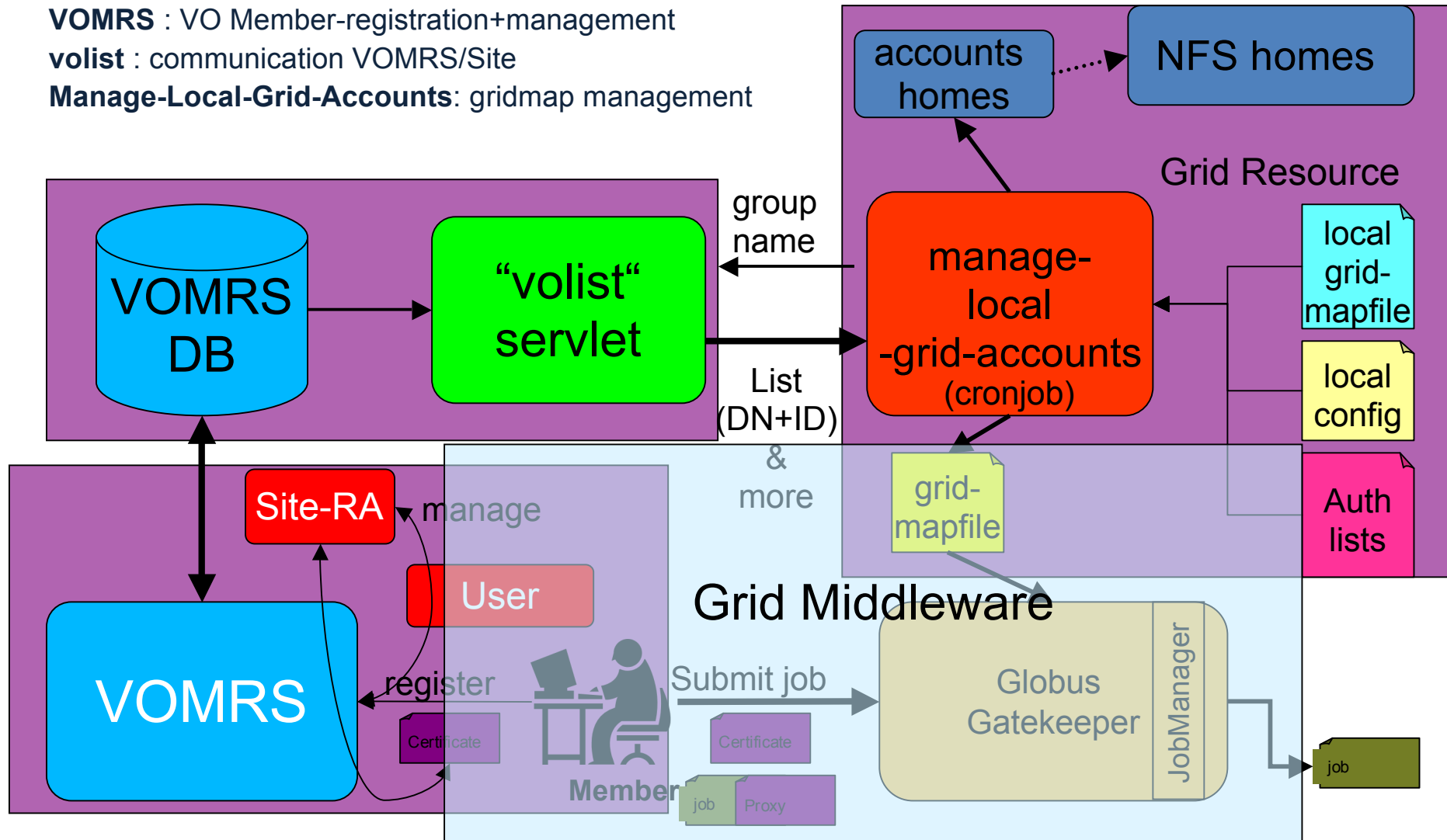


VO/Grid Account Management (Beispiel AstroGrid)

VOMRS : VO Member-registration+management

volist : communication VOMRS/Site

Manage-Local-Grid-Accounts: gridmap management



D-Grid-Infrastruktur

- 3 Grid Middlewares:
 - Globus, gLite, Unicore
 - unterschiedliche Architekturen, einheitliche, PKI basierte Grid-Sicherheits-Infrastruktur (GSI)
- 2 Zertifizierungs-Organisationen (Certificate Authorities, CA)
 - GridKA (KIT, ehem. FZK), DFN
 - Einheitliche Organisation des Virtuellen Organisations Managements
- Ressourcenanbieter (Rechen- und Speicherkapazitäten) aus
 - Instituten und Hochschulen
 - Hochschul- und Landesrechenzentren
 - HPC Einrichtungen

Einschub: Infrastruktur DGrid II



D-Grid

[AstroGrid](#) [BW-Grid](#) [C3-Grid](#) [DGI](#) [FinGrid](#) [HEP](#) [InGrid](#) [MediGrid](#) [PartnerGrid](#) [PT-Grid](#) [TextGrid](#) [WISENT](#)

D-Grid gesamt

[Drill-Down](#)

[Computing
Elemente](#)

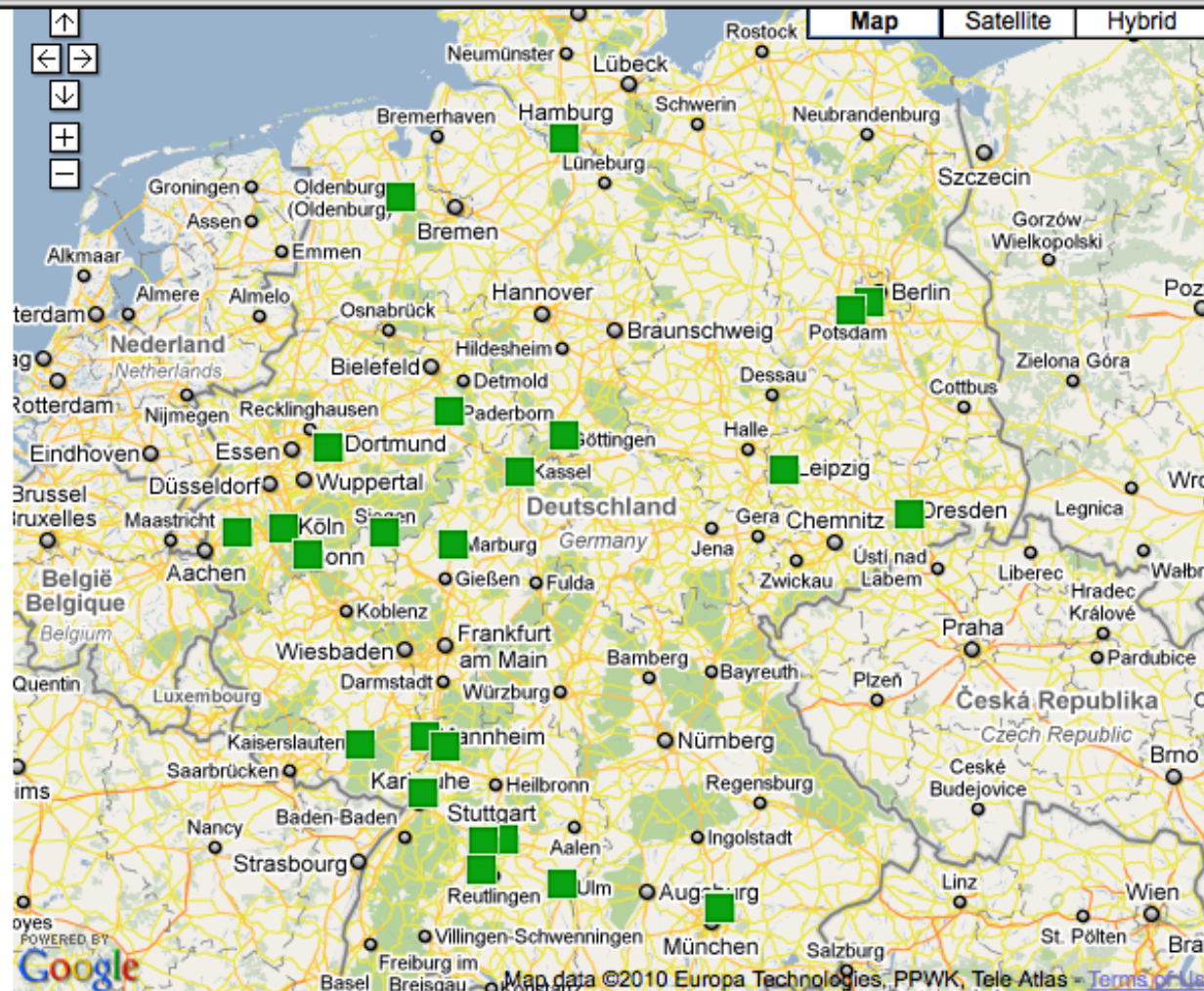
[Cluster
Workernodes](#)

[Siteinfo und
Wartung](#)

[Topologie](#)

[Jawari
Benchmarking](#)

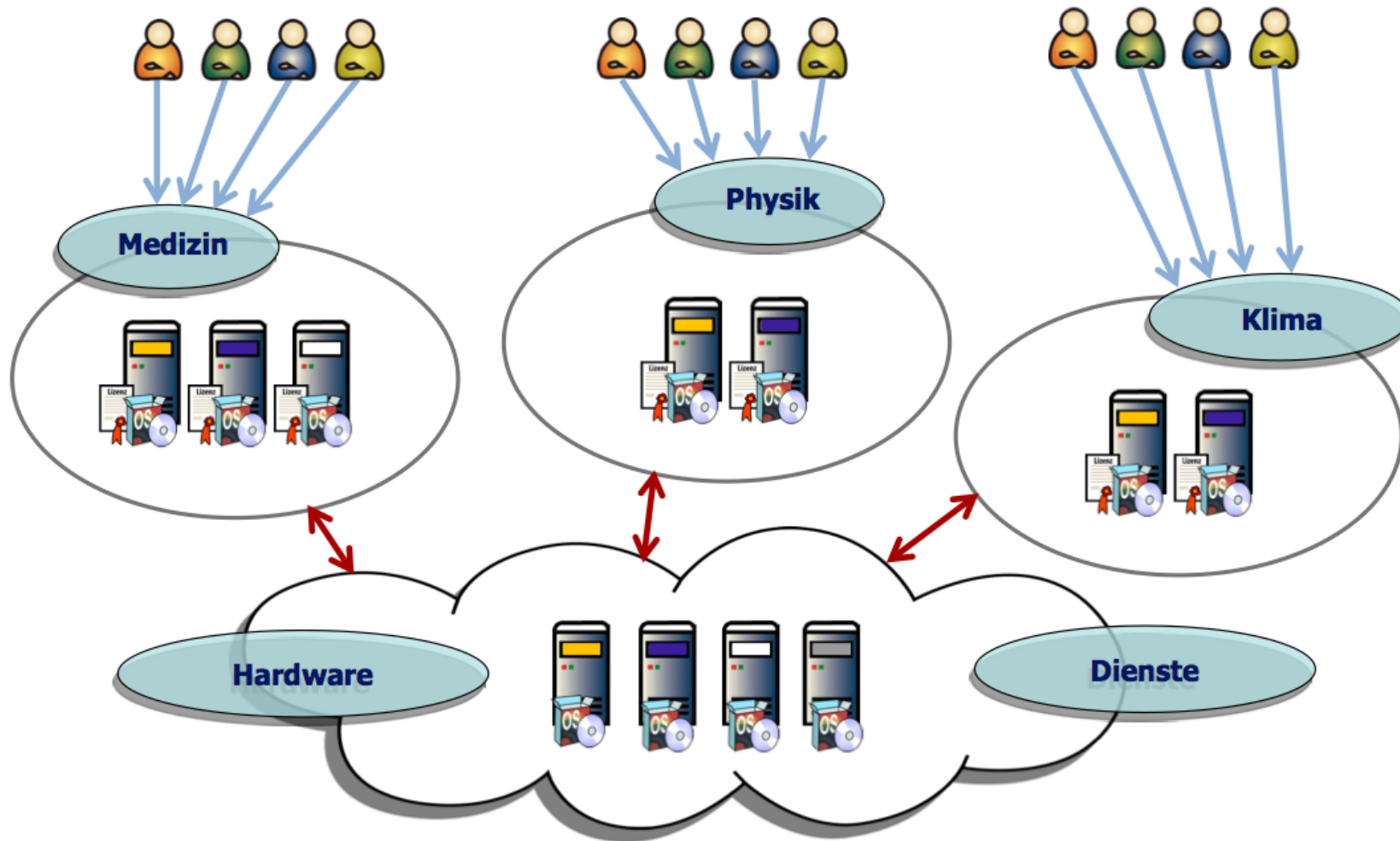
[Details XML-Dump](#)



- normaler Betrieb
- geplante Wartung
- Wartung
- Fehler

[Geomaint MDS Dynamic Map](#)





U. Schwiigelshohn, DGrid AHM 2010

Datenbereitstellung und –sicherheit: Nutzung DGrid I

Nutzung der GSI und der D-Grid-Infrastruktur:

- Service Level Agreements mit DGrid-Providern für
 - die Bereitstellung von dedizierten Ressourcen
 - Datenspeicherplatz
 - Software für Datenzugriff / -bearbeitung
 - Rechenkapazität „on demand“, aber auf vereinbarten Ressourcen (Trust)
 - Professionelle Maintenance der Ressourcen
- Derzeit OpenSource R für Auswertungen auf Grid Ressourcen
- Langfristiger: Lizenzen für SPSS, Stata,.... auf Grid Ressourcen anwendbar

Datenbereitstellung und –sicherheit: Nutzung DGrid II

Nutzung der GSI und der D-Grid-Infrastruktur:

- Verwaltung der VO-Mitgliedschaft und der VO-Gruppen durch den VO-Management-Prozess
 - Community basierte Entscheidungsstrukturen
 - Schnelle und leichte Bildung von Arbeitsgruppen
 - Zertifikatsbasierte Authentifizierung
- Erfüllung von Sicherheitsanforderungen der FDZ, z.B. für die Online-Auslieferung von SUF, möglich
 - Langfristig ggf. weitere Vereinfachungen beim Fernrechnen?
 - Bessere Referenzierung von gelöschten SUF?
- Speicherung von und Zugriff auf Daten in einer Umgebung, die verschiedene Sicherheitskriterien zusammenführt und erfüllt
- Für den einzelnen Nutzer mit weniger Aufwand verbunden