



Management wissenschaftlicher Daten mit dem DataFinder

Softwaretechnologie im DLR – Sistec stellt sich vor.

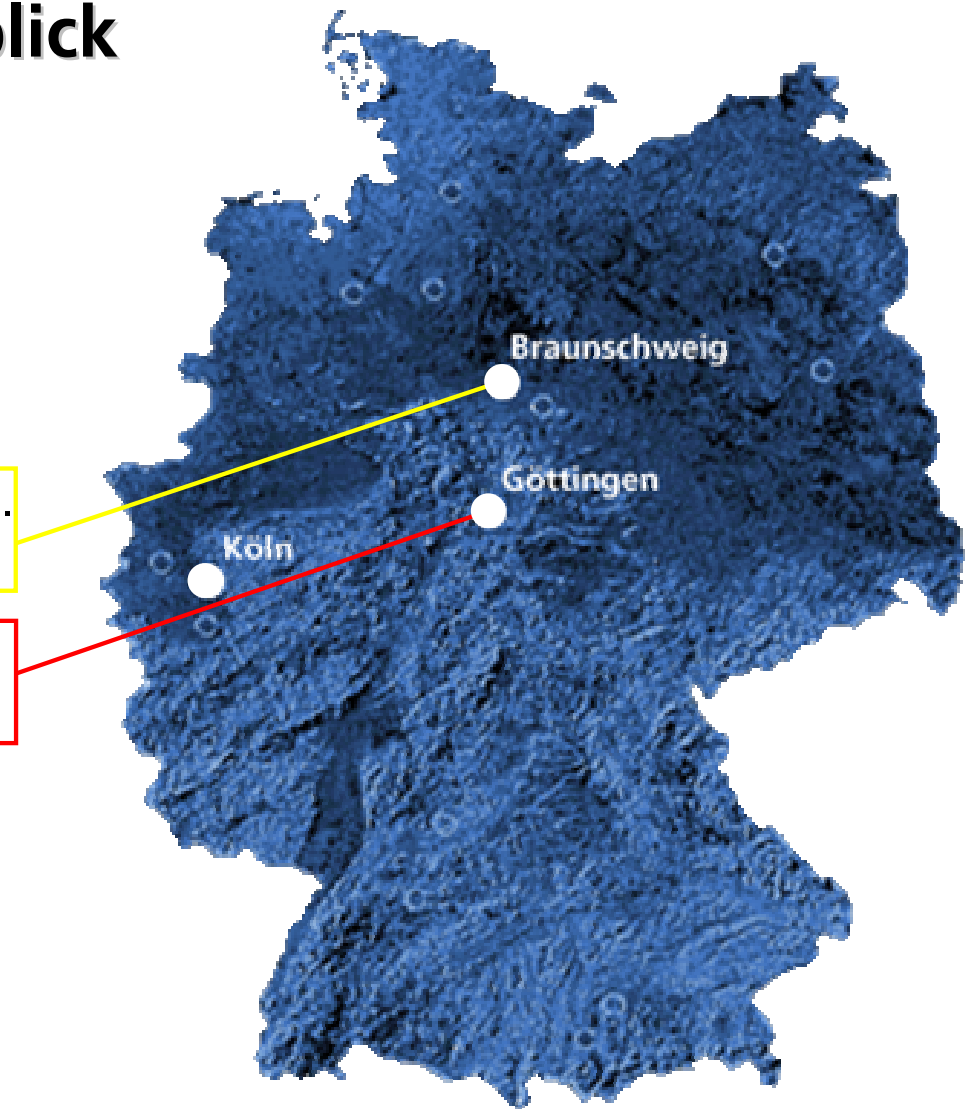


Eike Hoffmann,
DLR-Braunschweig
Institut für Aerodynamik und Strömungstechnik



Das Institut für Aerodynamik und Strömungstechnik (AS) im Überblick

- 3 Standorte (Braunschweig, Göttingen, Köln-Porz)
- Es werden numerische und experimentelle Untersuchungen an Flug- und Raumfahrzeugen durchgeführt.
- Braunschweig: 100 Mitarbeiter, 10 TB, 7 Mio. Dateien (März 2002)
- Göttingen: 140 Mitarbeiter, 4 TB, 7 Mio. Dateien (März 2002)
- Stand September 2003 in Braunschweig: 16 TB, 8 Mio. Dateien.
→ Zuwachs um 50% in einem Jahr (inkl. Aufräumaktion)!





zu viele daten sind zu wenig informationen ...

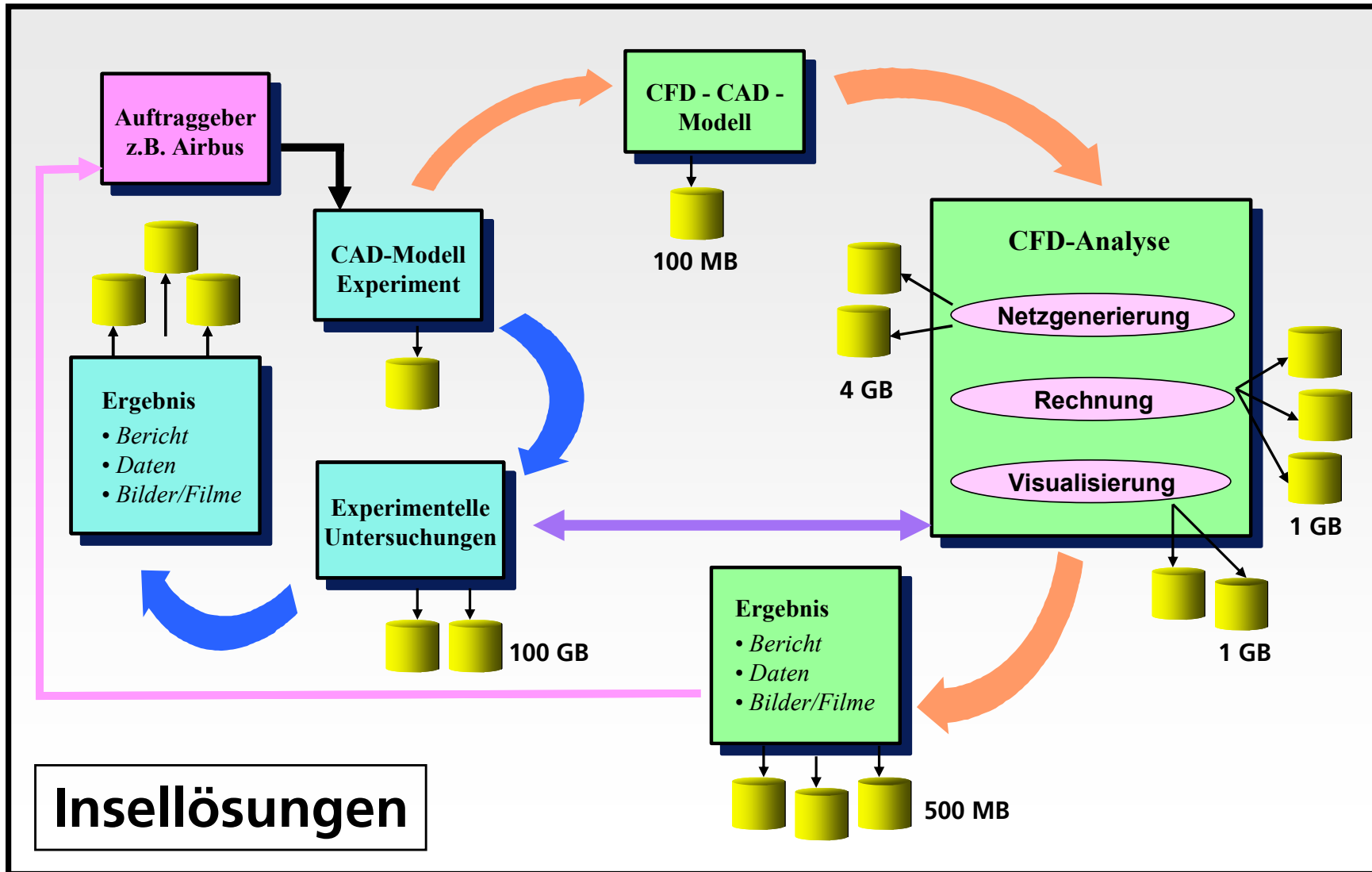
Gerhard Wohland (1995)

Lösungsansatz

- ◆ Daten selektieren
- ◆ Daten strukturieren
- ◆ Zusammenhänge transparent machen
- ◆ Beziehungen herstellen

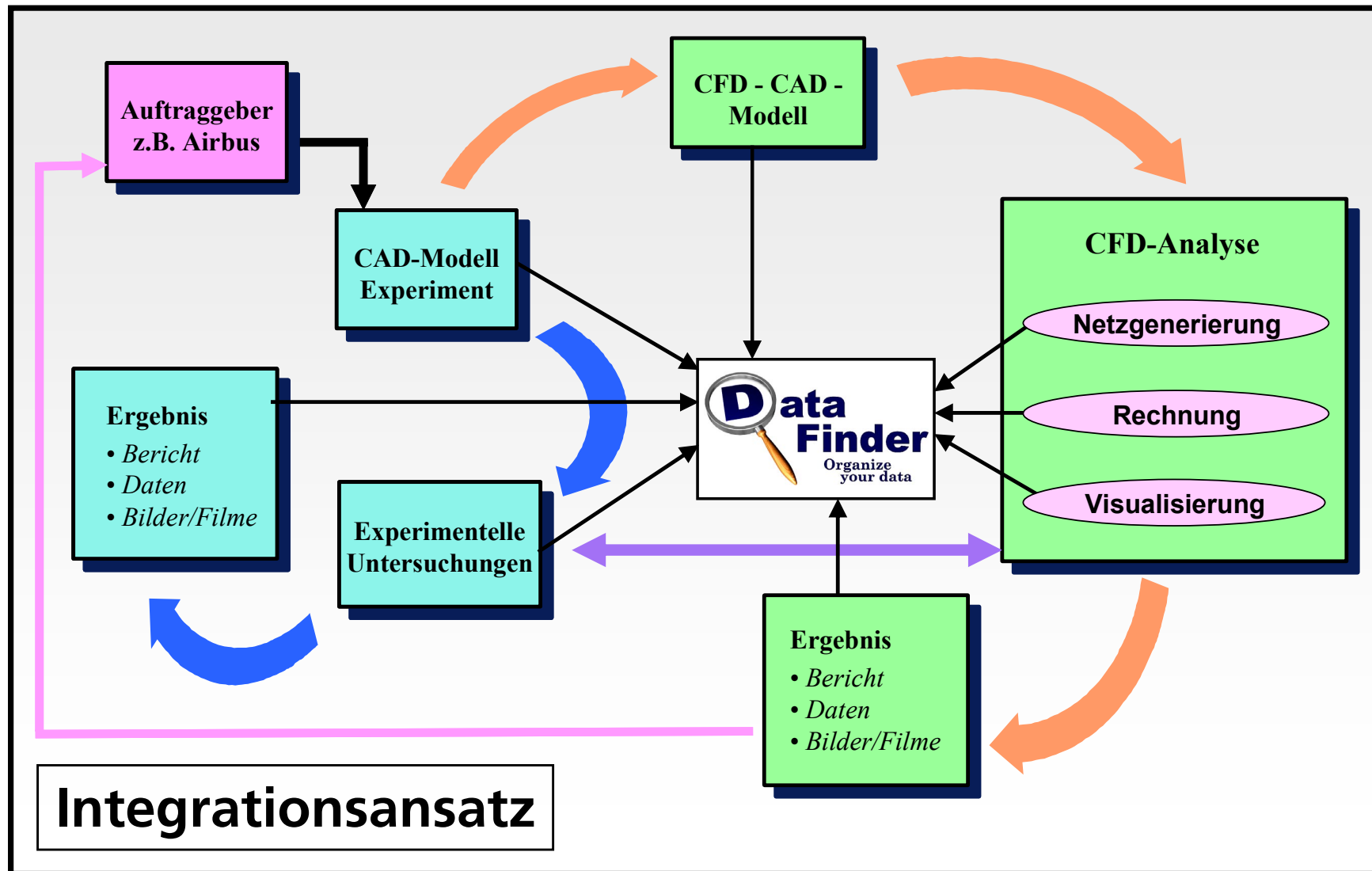


Analyse der Ist - Situation





Entwurf der Soll - Situation





Datenmanagement - Anforderungen von AS

- Verschiedene Standorte und unterschiedliche Medien (Online/Offline)
- Strukturierung der Daten und Kennzeichnung mit Attributen (Metadaten)
- Teamarbeit → Zugriffs- und Rechteverwaltung
- Einfaches suchen und finden von Daten / Verwaltung großer Datenmengen
 - Effizienz bei großen Datenmengen!
 - Auch nach 10 Jahren müssen die Daten noch nachvollziehbar sein.
- Einfache/schnelle Zugänglichkeit (versch. Plattformen: SGI-IRIX, Linux, MS Windows)
- Keine Datenredundanz
- Abbilden einfacher Arbeitsabläufe (Simulationsautomatisierung)
- **Kostenverfolgung, Kostenkontrolle und Kostenbegrenzung**



Erwarteter Nutzen 1: Zeitverkürzung

- Unterstützung paralleler Abläufe
- Unterstützung der gesamten Ablauforganisation (Gleichzeitiges Bearbeiten von Daten, Datenaustausch mit Partnern, etc.)
- Verkürzung der Bearbeitungszeit
- Bereichsübergreifender Zugriff auf alle begleitenden Daten von Allen die am Arbeitsprozess beteiligt sind
- Integration technischer und kommerzieller DV-Anwendungen
- **Reduzierung der Suchzeiten**
Mitarbeiterbefragung: 14% Suchzeiten, Reduzierung auf 7%



Erwarteter Nutzen 2: Kostenreduzierung

- Reduzierung von Doppelarbeit
- Reduzierung von nicht notwendigen Arbeiten (z.B. lange Suchzeiten)
- Verwaltung unterschiedlichster Objekte in einem System
- Zentralisierte Datenhaltung und Verwaltung
→ Reduzierung doppelter Datenbestände
- Langzeitarchivierung nur für die wirklich „wichtigen“ Daten, bei welchen der entsprechende Kostenaufwand vertretbar ist
→ Organisatorisches Konzept wird dafür benötigt
- **Mitarbeiterbefragung: ca. 54% redundante Daten, Reduzierung auf ca. 38% bis 19% möglich**



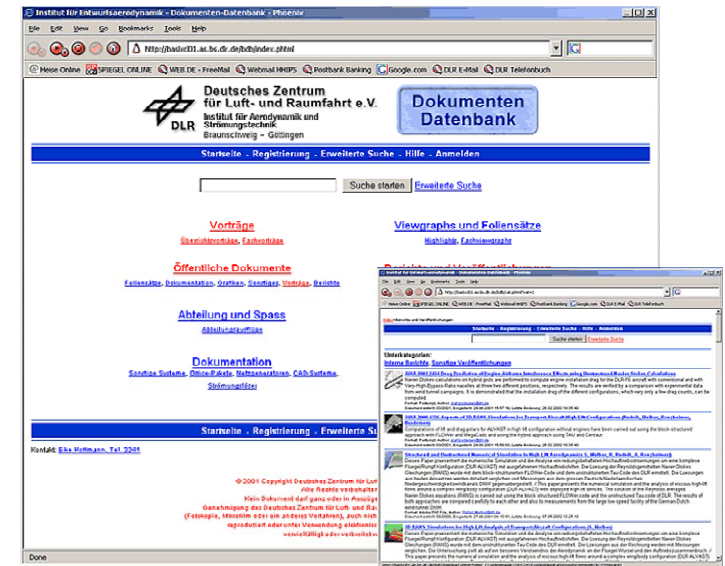
Erwarteter Nutzen 3: Qualitätsverbesserung

- Verbesserung des Projektmanagements durch bessere Übersicht
- Verbesserung der Unterlagenqualität durch bessere Dokumentation (bzw. mehr Dokumentation) – Langzeitdokumentation
→ Produkthaftung, 10 Jahre Aufbewahrungspflicht, ...
- **Know - How Sicherung !!!**
(Transparenz: Abbilden von Zusammenhängen die z.T. nur in den Köpfen Weniger existieren)
- Qualitätsverbesserung im Datenmanagement:
 - Mehr bzw. überhaupt Standardisierung von Verfahren und Prozessen, inkl. Dokumentation
 - Flexibilität der Verfahren
 - Einfache Handhabbarkeit der Verfahren
 - Unterstützung der Anwender/Mitarbeiter (Regeln/Richtlinien)



Kurze Historie der durchgeführten Arbeiten

- Zunächst Entwicklung von WWW-Katalogsystemen am Institut AS
 → Einsatz in bestimmten, eng abgegrenzten Teilbereichen z.B: Dokumentendatenbank
- Evaluierung kommerzieller (Enterprise) Produkt-Daten-Management-Systeme (PDMS) aus dem CAx – Umfeld
- Ein PDM-System wurde dabei in einem sog. Proof of Concept genauer evaluiert
- **Fazit war: System ist geeignet. Aber hohe Investitionskosten bei großem Funktionsumfang der nicht benötigt wird.**





Konzept der DLR-Einrichtung SISTEC zum Management wissenschaftlicher Daten

- Softwaresystem ist auf Anforderungen des Instituts AS zugeschnitten
- Verwendung offener, flexibler Standards mit Zukunftsperspektive und etablierter Standardsoftware, dadurch:
 - Kurze Entwicklungszeiten,
 - Geringe Einführungskosten und
 - Geringe Wartungskosten.
- Entwicklung eines JAVA Prototypen zur Demonstration des Konzepts
→ Implementierungszeit 3 Monate (Proof of Concept durchgeführt)
- Fazit: Vorteil gegenüber kommerziellen PDMS:
DataFinder speziell angepasst = Besseres Kosten/Nutzen-Verhältnis

→ Projekt DataFinder 1



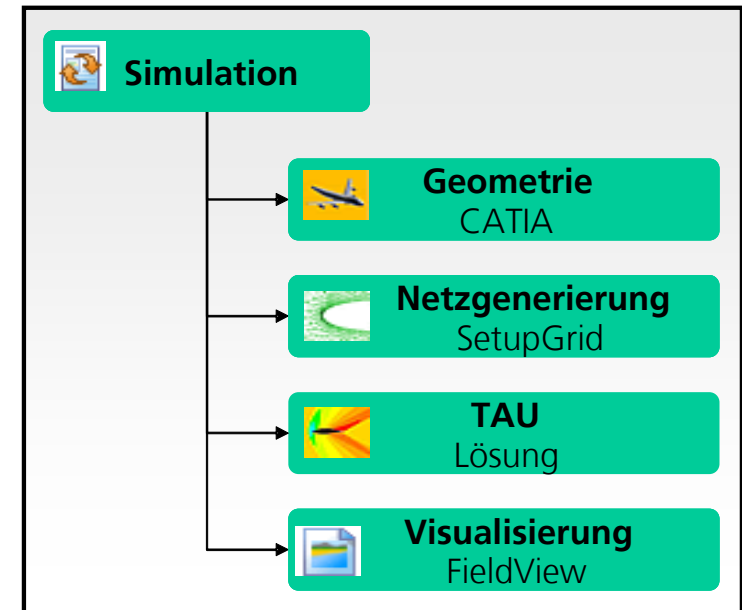
Das Projekt DataFinder 1

- Nur neu entwickeln, was es nicht fertig gibt:
- Verwendung etablierter Standardsoftware wo immer möglich
- DLR-weite Nutzung der entwickelten Software vorsehen
- **Verwendung offener Standards mit Zukunftsperspektive**
 - XML für Metadaten
 - WebDAV für Client/Server-Kommunikation (HTTP-Erweiterung)
 - Fertige Server-Software
 - Python als Implementierungssprache
 - Qt mit Python-Schnittstelle für grafische Benutzeroberfläche
 - Python für Skript-Abarbeitung
- Beginn Anfang 2003 mit Anwenderworkshop → Lastenheft
- Dann Erstellung Pflichtenheft, Implementierung und erste Testphase



DataFinder Funktionen

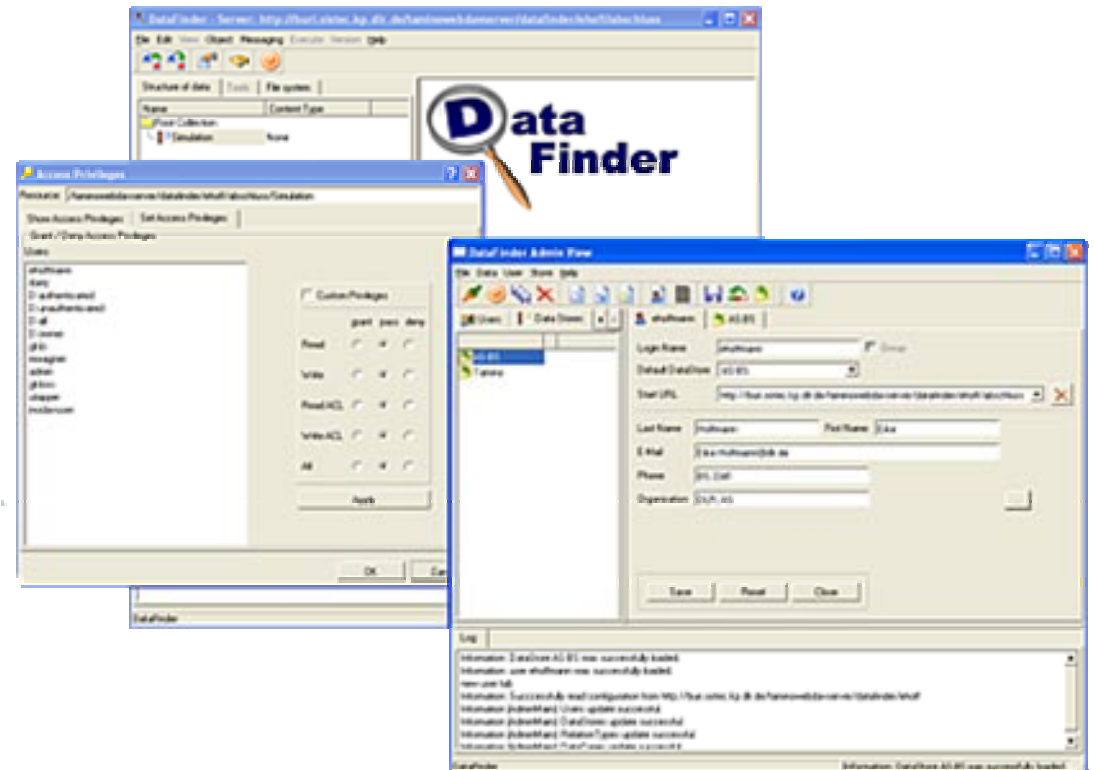
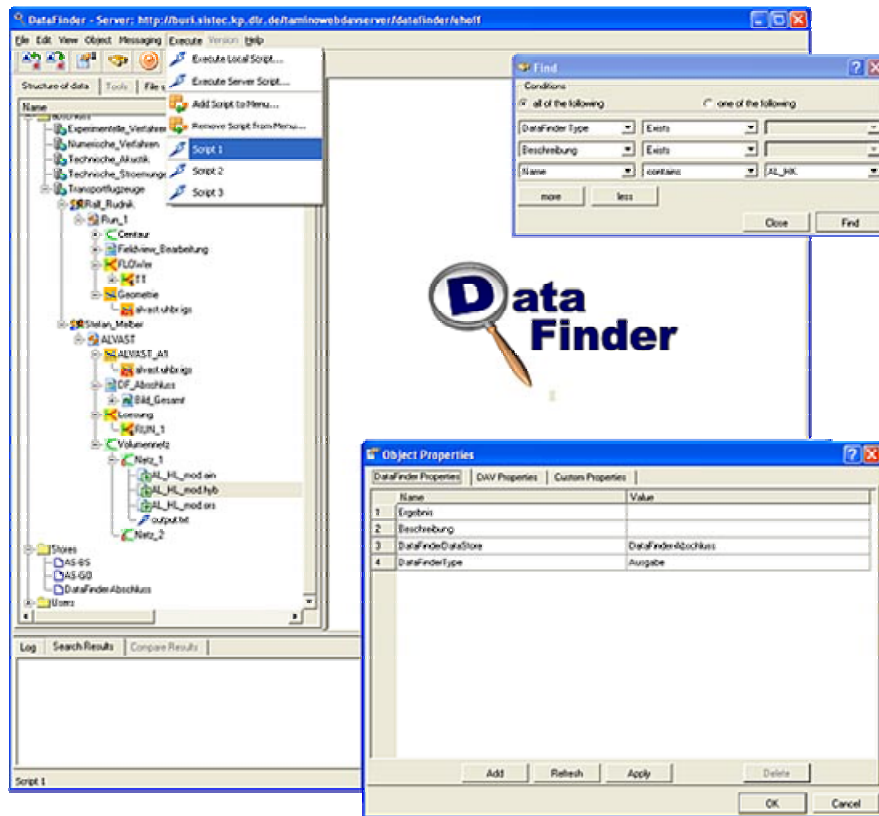
- Clients für **verschiedene Plattformen** (SGI-IRIX, Linux, Windows ...)
- **Freie Datenmodellierung** (Objekte, Relationen, Attribute)
- Up- und Download von Daten in **verschiedene DataStore-Typen** (z.B. FTP, WebDAV, Dateisystem) – Trennung von Metainformationen/Struktur und Daten
- **Benutzer-, Gruppen-, Rollen- und Rechteverwaltung**
 - Authorisierung über LDAP möglich
- **Komplexe Suchmechanismen** (jedoch zur Zeit keine Volltextsuche), Integriertes **einfaches Benachrichtigungssystem**
- **Erweiterbar durch Skripte** in Python (Programmierschnittstelle)





Anwender Client

Administrator Client



Realisierung der Oberflächen mit Python und Qt



DataFinder - Server: http://bsasdf01.as.bs.dlr.de:4000/taminowebdavserver/datafinder/config/proj/s...

File View Object Messaging Execute Help

Structure of data File system

Name	DF Data Type	Content Type	Length	Modified	Created
sikma					
SikMa					
Experimentelle Untersuchungen	Projekt	None		0 10. Feb. 16:22	10. Feb. 16:22
Experiment	Experiment	None		0 10. Feb. 16:23	10. Feb. 16:23
TWG Februar 2004	Messkampagne	None		0 10. Feb. 16:23	10. Feb. 16:23
Eike Test	Messreihe	None		0 13. Feb. 15:27	13. Feb. 15:27
12_06530	Messpunkt	None		0 13. Feb. 15:37	13. Feb. 15:37
D033_01.dat	Fmt_SikMa_DAT	applic			
D033_01.dp	Fmt_SikMa_DP	applic			
D033_01.dpe	Fmt_SikMa_DPE	applic			
D033_01.ida	Fmt_SikMa_IDa	applic			
D033_01.twg	Fmt_SikMa_TWG	applic			
3_00030	Messpunkt	None			
D034_01.dat	Fmt_SikMa_DAT	applic			
D034_01.dp	Fmt_SikMa_DP	applic			
D034_01.dpe	Fmt_SikMa_DPE	applic			
D034_01.ida	Fmt_SikMa_IDa	applic			
D034_01.twg	Fmt_SikMa_TWG	applic			
7_00000	Messpunkt	None			
00001.dat	Fmt_SikMa_DAT	applic			
00001.dp	Fmt_SikMa_DP	applic			
00001.dpe	Fmt_SikMa_DPE	applic			
00001.ida	Fmt_SikMa_IDa	applic			
00001.twg	Fmt_SikMa_TWG	applic			

Log Search Results

19:39:41: Connected to http://bsasdf01.as.bs.dlr.de:4000/taminowebdavserver/datafinder/config/proj/s...

Resource Properties

Resources: 12_06530

DataFinder Properties | DAV Properties | Custom Properties

Name	Value
1 Ist_Rollwinkel	45.28 [deg]
2 Giernoment	-185.2534 [Nm]
3 Messzeit	hh:mm:ss
4 Staudruck	25001.4 [Pa]
5 Analogkarte_Status_Code	
6 Klappenmoment_links	2.2534 [Nm]
7 Y-Kraft	122.2534 [N]
8 Soll_Machzahl	
9 X-Kraft	2.2534 [N]
10 Inkrementalkarte_2_Status_Code	
11 Beschreibung	
12 Inkrementalgeberkarte_2	
13 statischer_Druck	30000.5 [Pa]
14 Anzahl_Posttrigger	
15 Taktrate	
16 Plenumsdruck	60145.5 [Pa]
17 IDa_Nummer	
18 Reynoldszahl	3.687 [Mio]
19 Dauer_Pretrigger	0.010 [s]
20 Inkrementalkarte_1_Status_Code	
21 Ruhedruck	60086.6 [Pa]
22 Klappenmoment_rechts	-4.2534 [Nm]

Add Refresh Apply Delete

OK Cancel

DataFinder Admin View

File Data User Store Utilities Help

Data Types Relation Types

Fmt_FLDWer Fmt_IGES Fmt_PDF Fmt_SikMa_DPE Fmt_TAU Messpunkt

DataType Name: Messpunkt

Attribute Name	Type	Default Value	Required
1 Beschreibung			
2 Messreihennummer			
3 Messpunktnummer			
4 IDa_Nummer			
5 Messungsart			
6 Soll_Machzahl			
7 Ist_Machzahl			
8 Ruhedruck			
9 statischer_Druck			
10 Plenumsdruck			
11 Staudruck			

New Attribute Remove Attribute

Save Reset Close

Log

19:46:16: new datatype tao
19:46:16: Datatype Fmt_SikMa_DPE was successfully loaded.
19:46:16: new datatype tab
19:46:16: Datatype Fmt_TAU was successfully loaded.

DataFinder

DataType Messpunkt was successfully loaded.

DataFinder: Messdatenerfassung im TWG



Weiterführung der Arbeiten & Ausblick

- Zielsetzungen von DataFinder 2 sind das **Bekanntmachen des Produkts** im DLR und die **Einführung an zwei DLR-Einrichtungen**
- **Recht hohen Bekanntheitsgrad im DLR** erreicht
- **Einführungs- und Erprobungsphase** an zwei DLR-Einrichtungen steht **kurz vor einem erfolgreichen Abschluss**
- Weitere **Proof of Concepts in Durchführung**
- Weiterentwicklung des DataFinder:
 - **WWW-Oberfläche, Erweiterte DataStore-Schnittstellen, Batch-Fähigkeit**, Anbindung Microsoft SharePoint Services (**Intranet**), Anbindung **TSM/ADSM** (SfR), Workflows, Anbindung **DLR-Projektmanagementportal**, ...
- Externe Vermarktung des DataFinder: Derzeit gibt es Kontakte mit einer Firma in Braunschweig.



Die Philosophie des DataFinder in Kürze

DataFinder ist eine Art Standardsoftware - aber kein System von der Stange.

Um den **DataFinder** optimal zu nutzen, muss er an die Einrichtung angepasst werden.

DataFinder stellt Zusammenhänge zu Daten her, die oft nur in den Köpfen der Mitarbeiter bestehen.

DataFinder eröffnet so die Chance, dieses unerschöpfliche Potential zu sichern und mit modernsten Mitteln zu erschließen, zu pflegen und für **alle** Mitarbeiter zur Verfügung zu stellen.

Ein **DataFinder** Datenmodell erlaubt die Abbildung beliebiger, komplexer Strukturen.

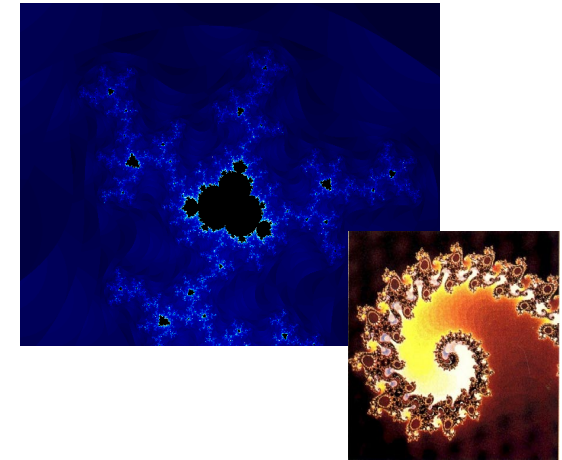
Die Daten der Einrichtung werden mit Hilfe des **DataFinder** Datenmodells modelliert / beschrieben.

Durch Nutzung offener Standards mit Zukunftsperspektive (z.B. XML und WebDAV) leicht erweiter- und anpassbar.



DataFinder

- Stellt Zusammenhänge her
 - macht Zusammenhänge sichtbar
- und**
- **schafft so Ordnung.**



Vielen Dank für Ihre Aufmerksamkeit!

<http://www.dlr.de/DataFinder>