
Data Vault als Kernkomponente operativ eingesetzter Softwarelösungen am Beispiel der Entwicklung des „ID environment managers“

Thomas Mattick,
InfoDyn GmbH



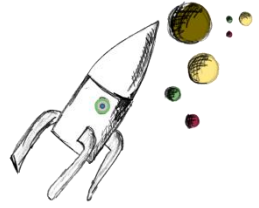
Thomas Mattick
Produkt Manager,
InfoDyn

Stationen (< 5 Jahre)

- 01/15 – Produkt Manager InfoDyn
- 10/14 – Business Development / Beratung InfoDyn / BBF
- 01/14 – 09/14 Senior Consultant BBF
- 05/09 – 12/13 Projektleiter / BI Solution Architekt SALT Solutions

Projekte (Auszug)

- InfoDyn Suite (BI unterstützend) – Produkt
- Auftragsabwicklung/Leistungsbewertung (Luftfahrt/Logistik)
- Wareneingang/Warenausgang/Retouren/Bestand (Logistik)
- Auftragsmanagement (Großhandel)
- Flächenbewirtschaftung (Einzelhandel)



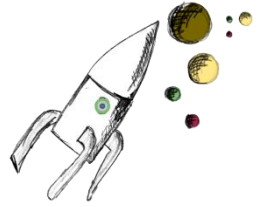
Ausgangslage

Ziele & Erwartungen

Architektur

Entwicklung

Bilanz



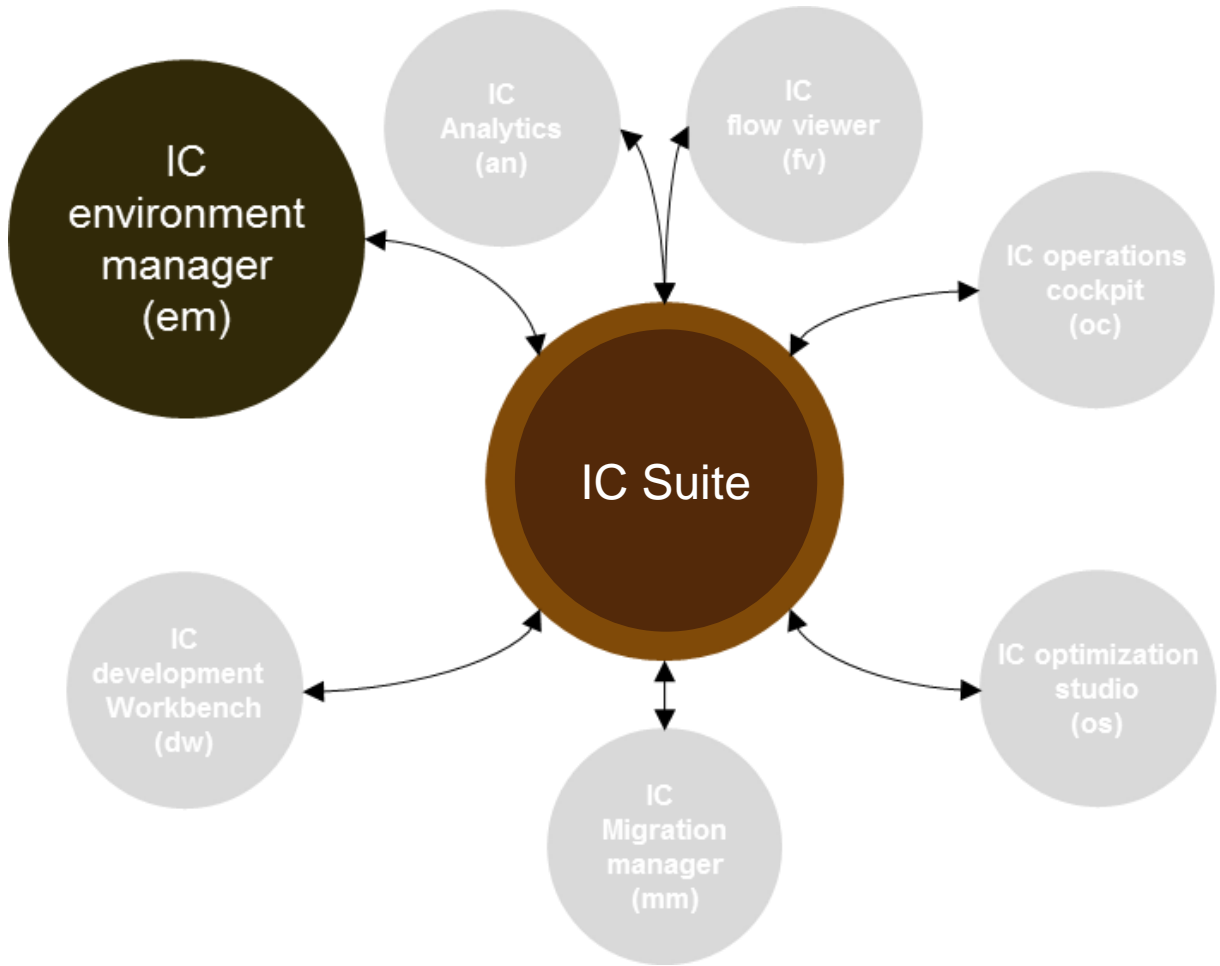
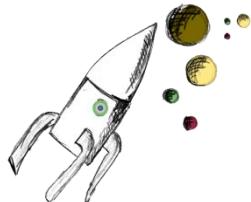
Ausgangslage

Ziele & Erwartungen

Architektur

Entwicklung

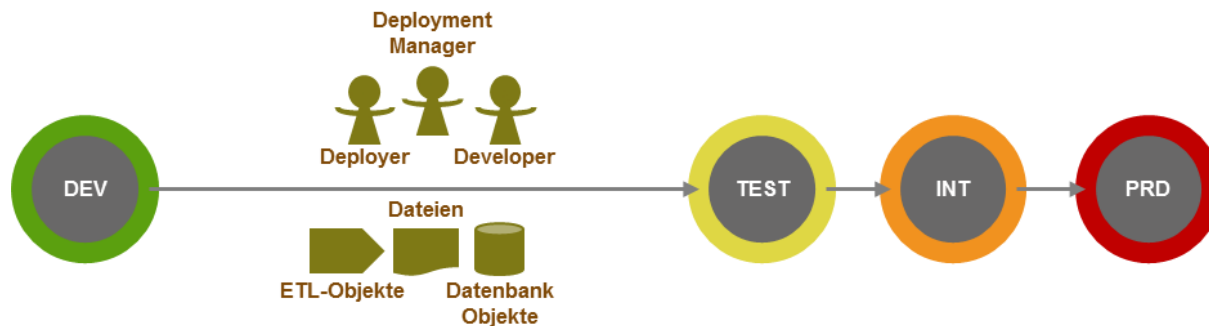
Bilanz





Allgemein

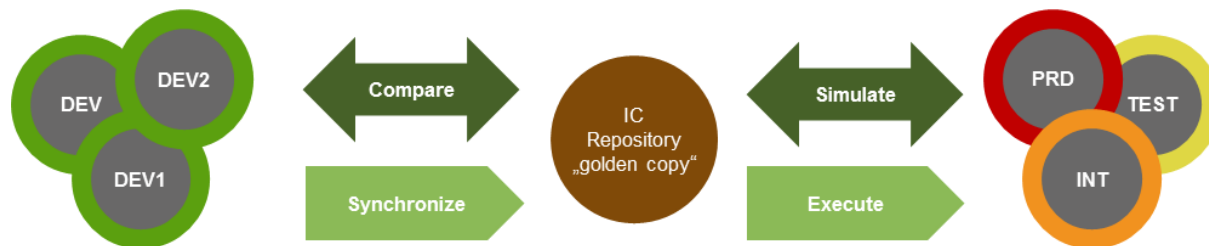
- Deploymentwerkzeug
- Standardisierung des Softwareverteilungsprozesses für verschiedene Technologien
- Sicherstellung der Einhaltung von Compliance Richtlinien
- Konfiguration, Steuerung und Überwachung des Release-Managements
- Trennung von Verantwortlichkeiten und Arbeitsschritten
- Organisation und Verwaltung unterschiedlicher Umgebungen
- Automatisierung von Softwareverteilungsprozessen





Zentrales Repository

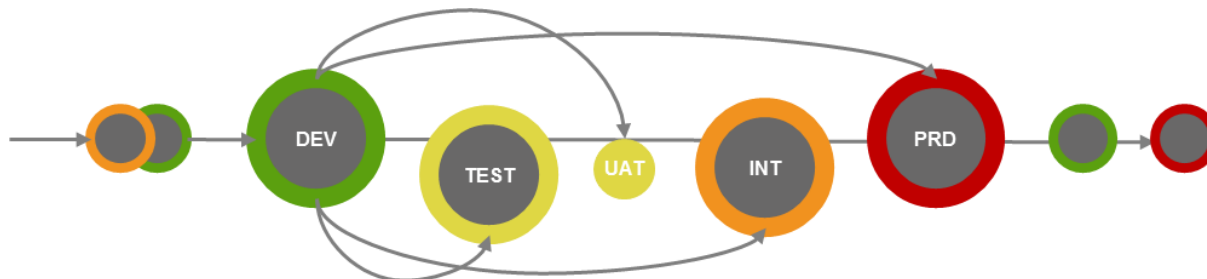
- Datenbank
 - Metadaten zu ETL-Objekten
 - Metadaten zu Datenbank-Objekten
 - Metadaten zu Dateisystemen
- Dateisystem
 - XML's zu ETL-Objekten
 - DDL's zu Datenbank-Objekten
 - Dateien (Skripte, Log-Dateien, ...)
- Vollständige ICem-Versionshistorie
- Vollständige und durchgängige Protokollierung (Aktionen und Prozesse)

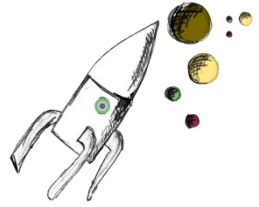




Allgemein

- Unterstützung
 - Betriebssysteme – Windows, Linux, AIX, Solaris
 - Informatica Powercenter – 8.6, 9.1, 9.5, 9.6
 - Informatica Repositories – Oracle, DB2, SQL-Server
 - Datenbanken - Oracle
 - Dateien – Windows, Unix
- Rechte-Management
 - Basierend auf Nutzern und Gruppen
 - LDAP-Integration
- Integrierter Freigabeprozess für produktive Deployments
- Integrierter Restore vollständiger Releases



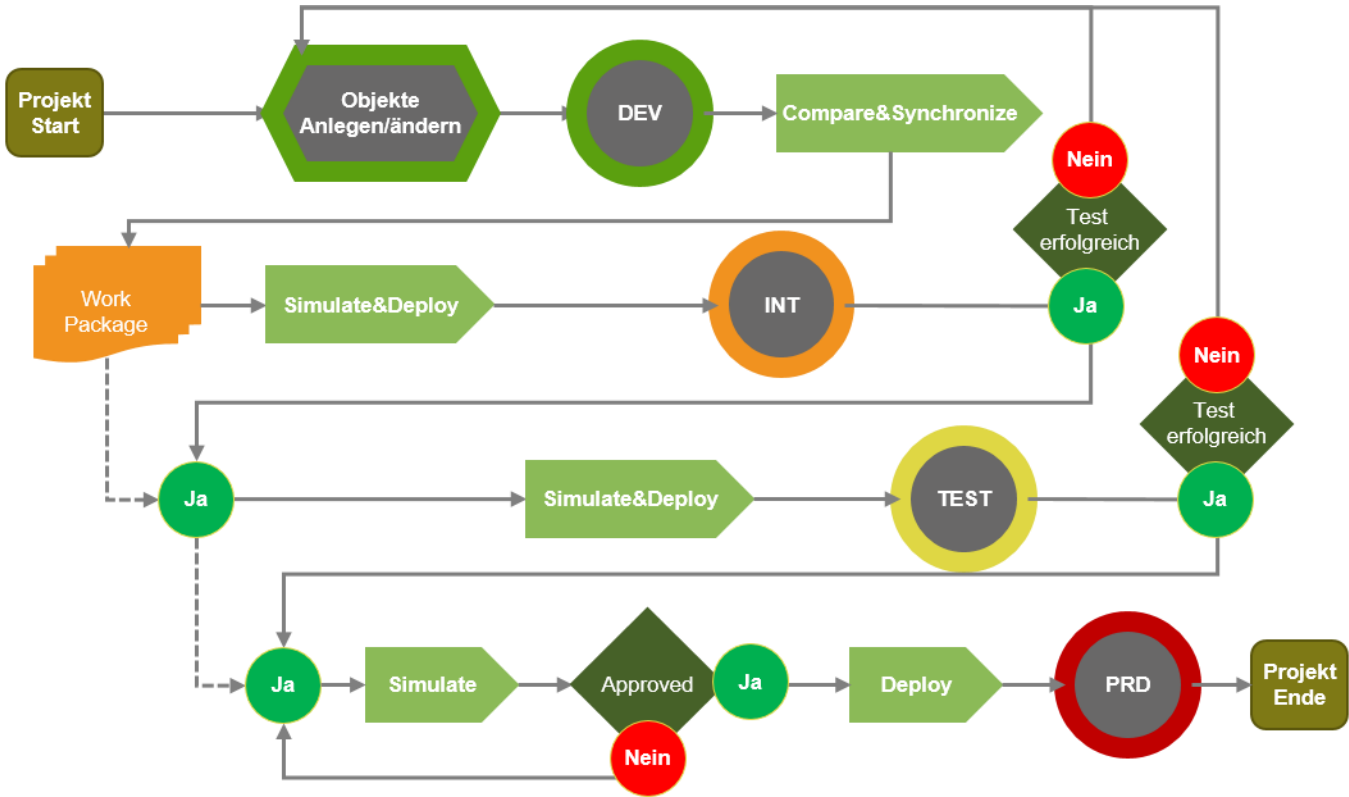


Funktional

- Applikationseigener Zugriff
 - Informatica – PMREP + Metadaten-Repository
 - Oracle – DDL und DML
 - Dateien – SFTP
- Skripting
 - Unterstützung von beliebigen Programmen in Form von Shell- bzw. Batchskripten
 - Umbenennen von Ordnern
 - Austausch von Parametern
 - Verschieben von Objekten nach dem Deployment
 - Ausführung von DDL-Skripten nach Deployment
 - ...
 - Können vor bzw. nach Aktionen in Quell- und oder Zielumgebungen ausgeführt werden
 - Können im Deployment selbst vor Ausführung in die Gastumgebung deployed werden



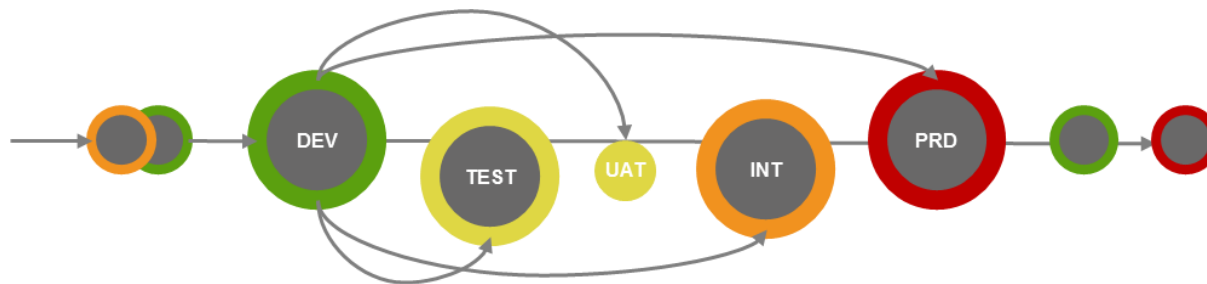
Prozessual

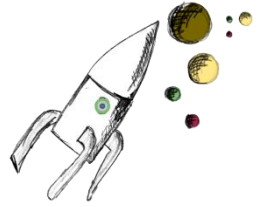




Vorteile

- Klare Verantwortlichkeiten
- Einheitliche Benutzeroberfläche
- Benachrichtigungssystem
- Sicherheit beim Deployment
- Darstellung abhängiger Objekte





Ausgangslage

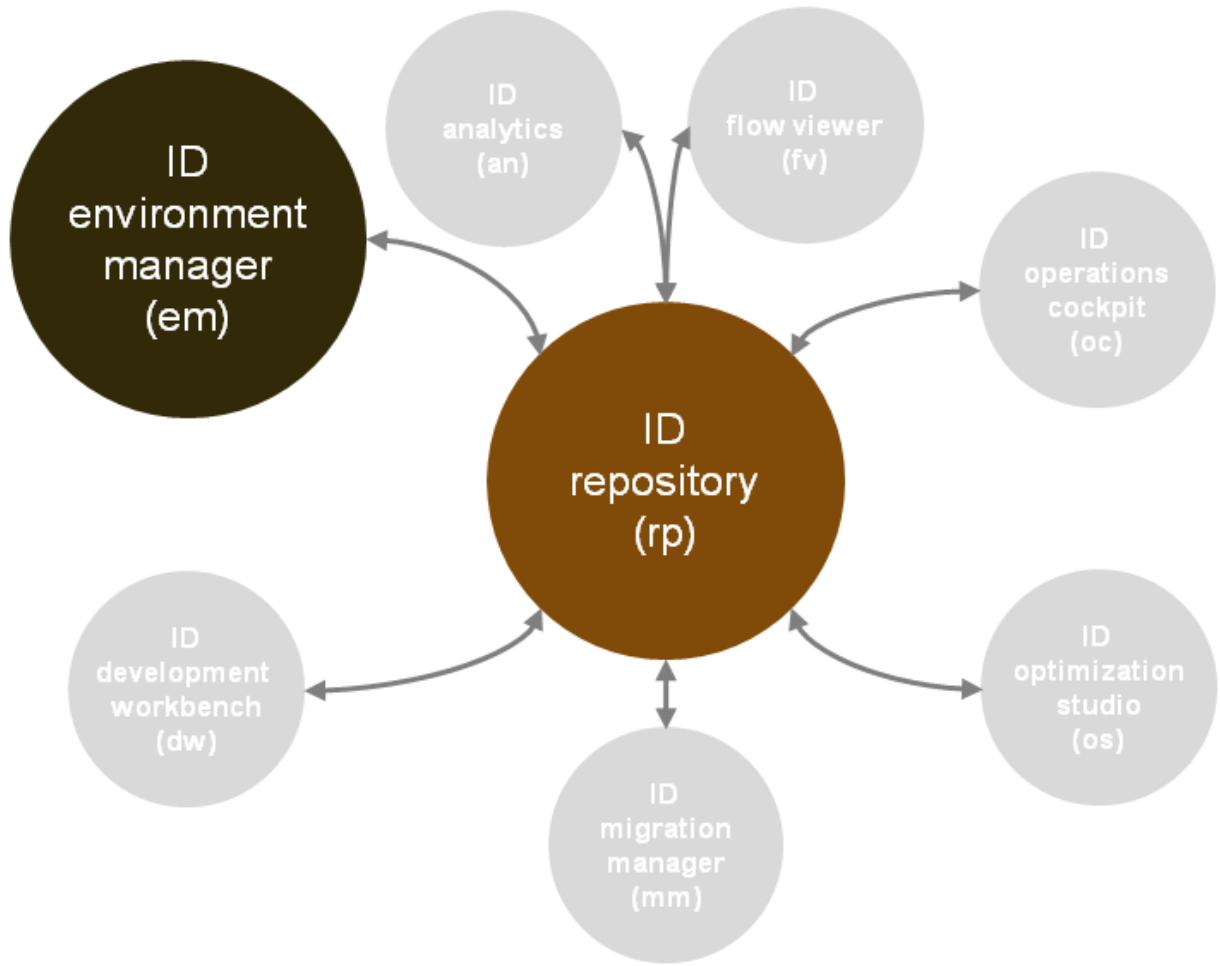
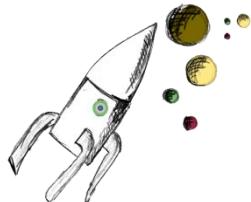
Ziele & Erwartungen

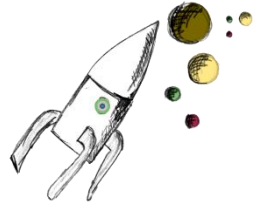
Architektur

Entwicklung

Bilanz

Ziele & Erwartungen > IC Suite wird ID Suite



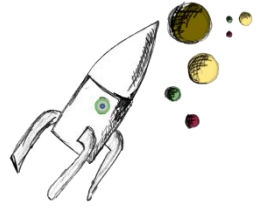


Ziele

- Modulübergreifendes Metadaten-Repository
- Unterstützung neuer Domänen
 - z.B. SQL-Server, Teradata, Talend, ...
- Implementierung Kommandozeilenwerkzeug
- Integration erweiterter Berichtsfunktionalitäten

Erwartungen

- Trennung der Domänen
 - Logik
 - Datenmodell
- Variable Beziehungen zwischen Domänen
- Kapselung von Logik
 - Datenmodell
 - Businesslogik
 - Frontendlogik

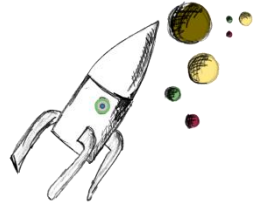


Ziele

- Parallele Entwicklung
 - Domänenunabhängigkeit
- Model Driven - Domänenmodellorientiert
 - Datenmodell -> Generator
 - CRUD-Operationen -> Generator
 - REST-Interfaces -> Generator

Erwartungen

- Implementierung von Generatoren
- Erweiterung um neue Domänen
 - Keine Implikationen auf bestehende Entitäten
 - Nutzung bereits bestehender Standardfunktionalitäten
- Erweiterung um neue Attribute bestehender Entitäten
 - Anpassung am Domänenmodell
 - Auswirkung über Generatorenprinzip auf alle Technologiebestandteile (DM, REST, ...)



Ziele

- Unterstützung unterschiedlicher
 - Workflows
 - Layouts
- Kardinalitätsverschiebungen zwischen Entitäten
 - $1:n > n:m > 1:1 > \dots$

Erwartungen

- Umstellung auf Modale Dialoge
- Referenzen zwischen Modalen Dialogen
 - Querverweise
 - Automatisches Ausfüllen nach Anlage
- Erweiterte Kontextmenüs
- Intuitivere Benutzerführung
- Komfortfunktionen
- Interaktion angemeldeter Nutzer



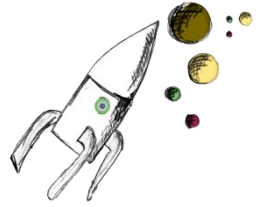
Ausgangslage

Ziele & Erwartungen

Architektur

Entwicklung

Bilanz

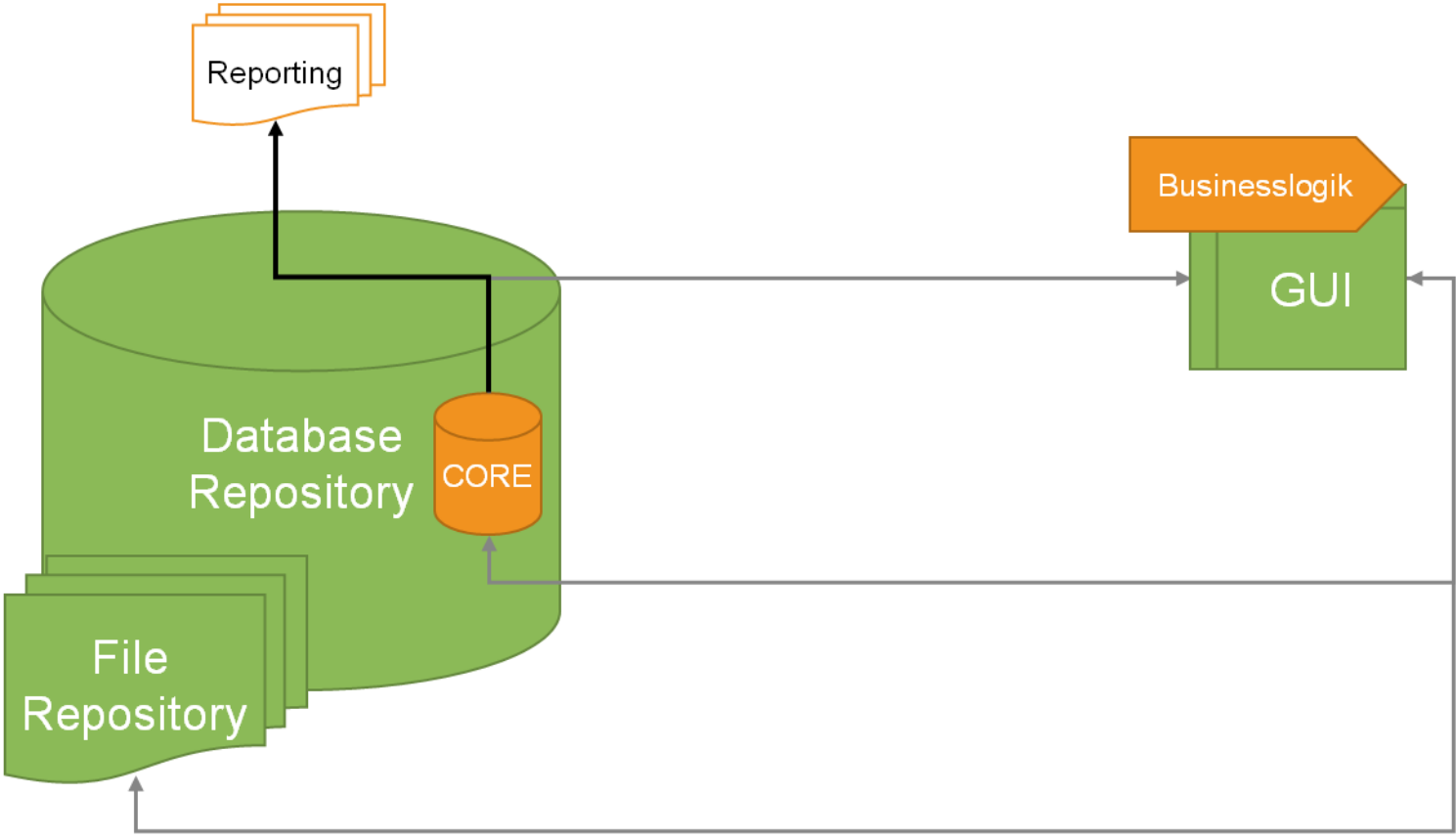


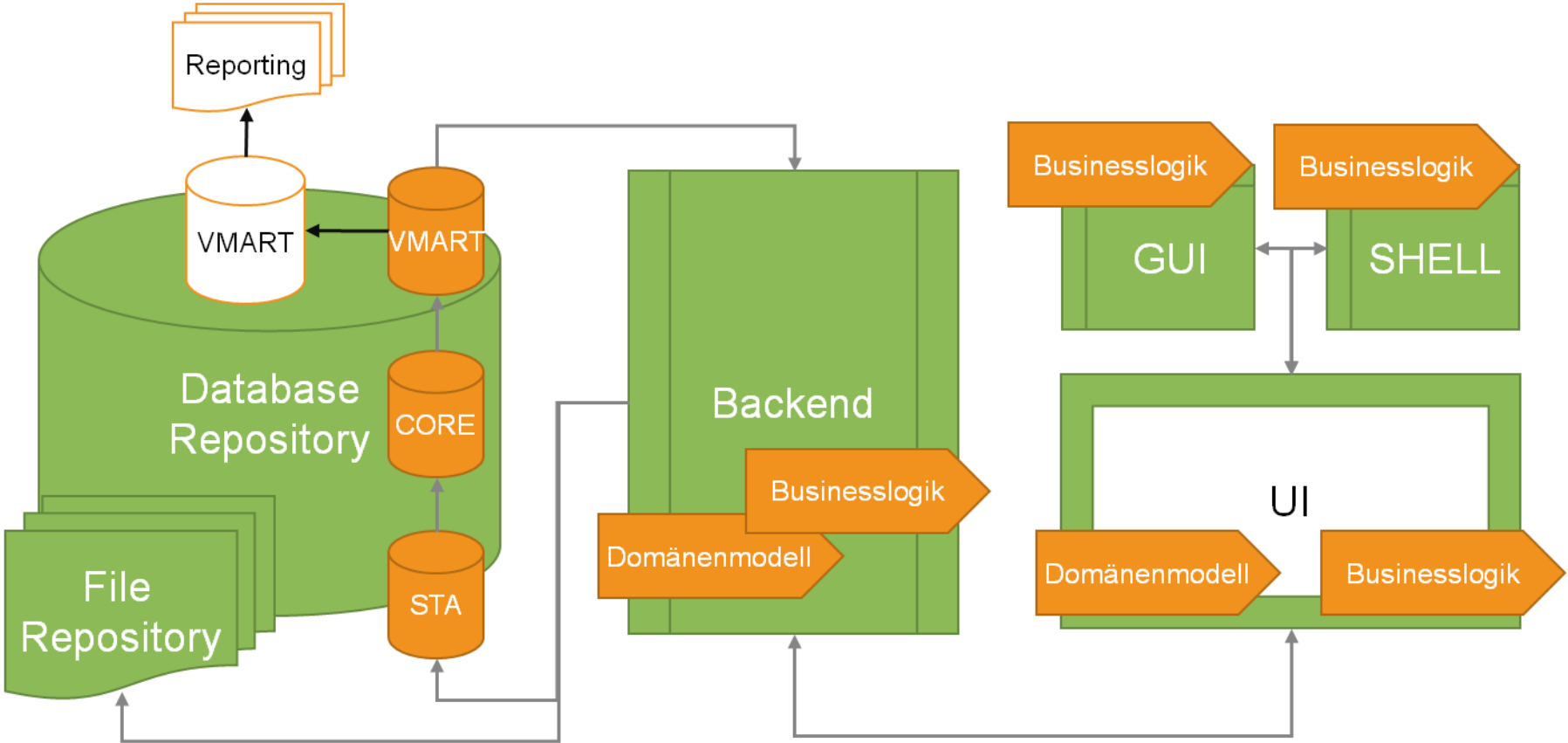
Software

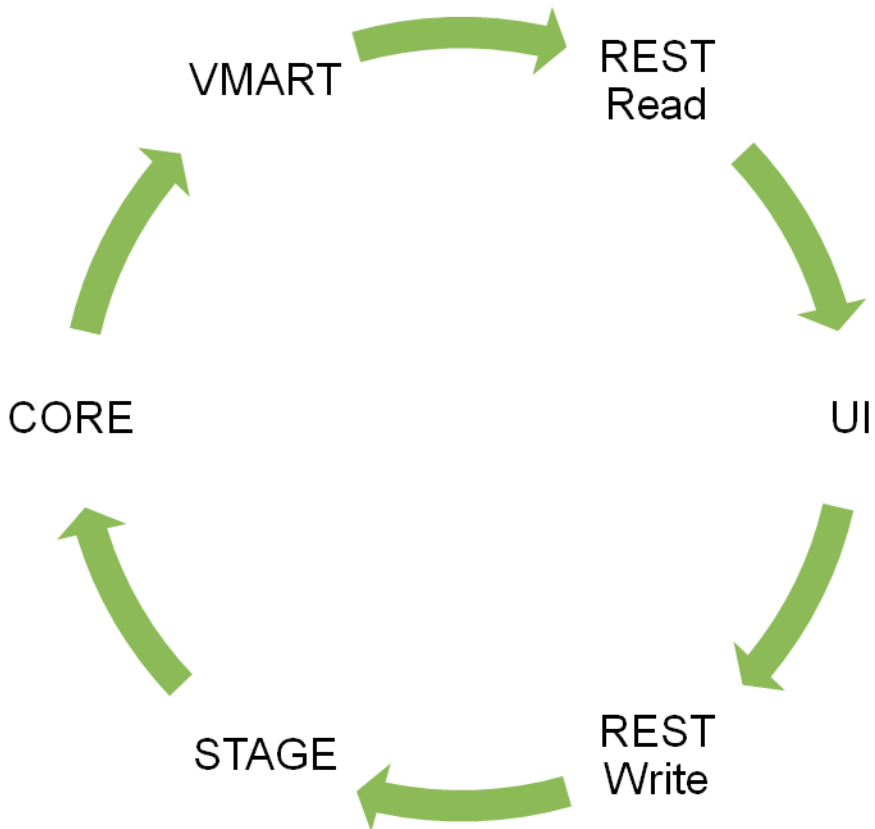
- Backend-Domänenmodell
- Backend-Businesslogik
- REST-Interfaces
- Frontend-Domänenmodell
- Frontend-Businesslogik
 - Anbindung unterschiedlicher UI-Module (GUI, Kommandozeile)

Zugriff

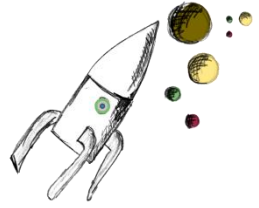
- STA-Schema
- CORE-Schema
- VMART-Schema







Agenda



Ausgangslage

Ziele & Erwartungen

Architektur

Entwicklung

Bilanz



Domainmodell

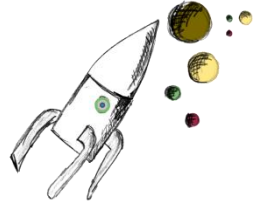
- Model Driven Development
- Pflege des Domainmodels über Eclipse
 - Hauptdomain
 - Subdomains
- Generierung von 80% folgender Objekte
 - JAVA Entitäten - Klassen und Methoden
 - JAVA Rest-Interfaces
 - DB Objekte
 - STA
 - CORE
 - VMART
 - [Bewirtschaftungsprozeduren]

```

  ▾ IReference
    > eu.infodyn.models.documentation
    > status : EIntegerObject
  > IVersionedReference -> IReference
  > IStaEntity
  > IDeployableObject
  > IComparableObject
  > IPhysicalObject
  > IBusinessObject
  > IConnectionFactory
  > IConnection
  ▾ System -> INamedElement, IRestElement
  ▾ systemName : EString
    > eu.infodyn.models.generation.hint
    > eu.infodyn.models.documentation
    > description : EString
    > environments : SystemEnvironment
    > applications : SystemApplication
    > projects : SystemProject
  ▾ SystemEnvironment -> IReference
    ▾ system : System

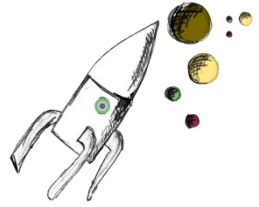
```

Property	Value
Changeable	true
Container	false
Containment	false
Default Value Literal	
Derived	false
EKeys	
EOpposite	environments : SystemEnvironment
EType	System -> INamedElement, IRestElement
Lower Bound	1



Zeitstempel

- Ablage operationaler Zeitstempel an Satelliten
 - LoadDts
 - ValidFrom
 - ValidTo
- Ablage prozessorientierter Zeitstempel im Log > Zugriff über Prozessschlüssel an Satelliten
 - CreateProcessId
 - ChangeProcessId



Generic Dependencies

- Abhängigkeiten zwischen Domänen
 - Domäne
 - Entität
- Bestimmung im CORE auf CORE-Daten
- Realisierung über Applikationslogik
- Bewirtschaftung nach regulärem Lauf parallel zur Versionslogik

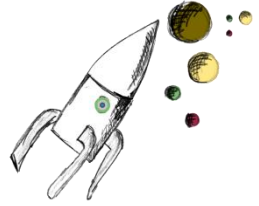
Domänenübergreifende Satelliten

- Domänenübergreifende Entitäten
- Standardmäßig für Haupt- > Subdomänenentitäten
- Spezialisierte Informationen in Subdomänensatelliten
 - IC_H_Entität
 - IC_S_Entität
 - FILE_S_Entität
 - IF_S_Entität



Schlüsselbildung

- Standard Hash
 - ID_SRC basierend auf externen Entitätsnamen
- UUID Hash
 - ID_SRC basierend auf UUID der Anwendung (Namensflexibilität) für interne Entitäten
- Mixed Hash
 - ID_SRC basierend auf Standard und UUID Hash -> Parent>Child
 - Definition überschreibt Linkbeziehung



Status

- Definition der Beziehung
 - 1:1, 1:n, n:m
- Definition im Domainmodell
- Generierte Bewirtschaftung anhand Beziehungstyp

Version

- Berechnung nach CORE-Ladung
- Basiert auf LoadDts
- Kann überschrieben werden

Delete

- Attribut Delete
- Housekeeping

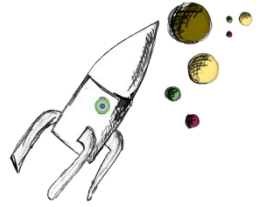


Activity Log

- Standard Entität
- Verlinkung zu Activity-Log relevanten Entitäten
- Aktive Einträge
- Generierte Einträge

Technical Log

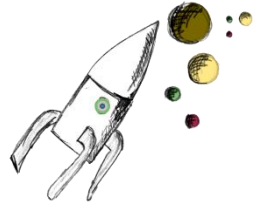
- Prozessverarbeitung und Laufzeitsteuerung
- Logging jeglicher Datenbewegung
 - STA
 - CORE
- Job-Ketten
 - JAVA als Auslöser



Transaktionen

- Bildung übergreifender/abhängiger Verarbeitungsketten
 - (STA->VAULT)
- Rollback bei Fehler
- Paketierung über JAVA
- Parallel
- Sequentiell nach Bewirtschaftungstyp (Keine Link auf Links)
 - Schritt 1: Alle Hubs
 - Schritt 2: Alle Links, Alle Satelliten (DEL) an Hubs
 - Schritt 3: Alle Satelliten an Links (DEL, STS, VRS)
 - Schritt 4: offene VRS Satelliten
- Ein Ziel eine Quelle

Agenda



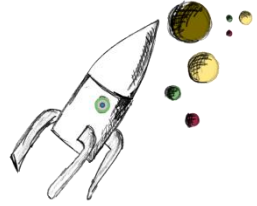
Ausgangslage

Ziele & Erwartungen

Architektur

Entwicklung

Bilanz



Vorteile

- Model Driven durch Standardisierung
- Hoher Anteil an Automatisierung > DDL, JAVA
- Kontextorientierte Verteilung der „Businesslogik“
- Automatische Versionskontrolle und Protokollierung aller Entitäten und Verarbeitungen

Nachteile

- Hoher Modellierungsaufwand
- Fehlende Fragmentierung von Satelliten aufgrund Generierung
- Lange Kommunikationswege
- Viele Transaktionen

Offen

- Change Management
- Datenoverhead in operativen Systemen
- Verarbeitungs- und Abfragezeiten bei steigendem Datenwachstum
- Überblick über Entitäten

Vielen Dank!

InfoDyn GmbH
Wendelsteinstraße 16
D-82166 Gräfelfing

www.infodyn.eu