

# Information Engineering und Information Life Cycle

Norbert Fuhr

Universität Duisburg-Essen

Information Engineering  
WS 2008/9

# Inhaltsverzeichnis

Übersicht über die Vorlesung

Information Engineering

Definition

Eigenschaften

Daten – Information – Wissen

Die Begriffe Daten, Information und Wissen

Sichtweisen auf Wissen

Wissensmanagement

SECI-Modell

Information Lifecycle

Der Information Lifecycle im Information Engineering

Phasen des Information Lifecycle

IE-Methoden im Information Lifecycle

Zusammenfassung

# **Übersicht über die Vorlesung**

# Kontakt

## Vorlesung

Norbert Fuhr  
LF 135  
(hinter der Stahltür)

Sprechstunde: nach Vereinbarung  
`norbert.fuhr@uni-due.de`

## Übung

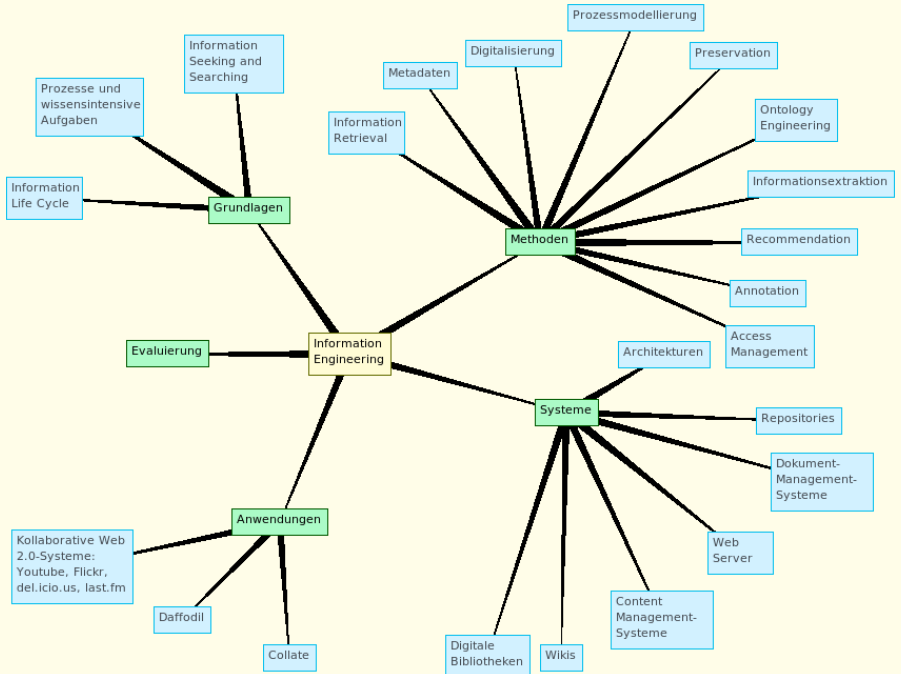
Matthias Jordan  
LF 136  
Sprechstunde: nach Vereinbarung  
`mjordan@is.inf.uni-due.de`

# Organisatorisches

- ▶ Übungen Montags 15:00-15:45 Uhr
- ▶ Übungsbetreuung: Matthias Jordan
- ▶ Übungen manchmal rechnergestützt
- ▶ Notizen zu Vorlesung und Folien auf Webseite

# (Vorläufige) Übersicht über die Vorlesung

1. Grundlagen
2. Methoden
3. Systeme
4. Anwendungen
5. Evaluierung



# Grundlagen

- ▶ Information Engineering
- ▶ Information Lifecycle
- ▶ Daten – Information – Wissen
- ▶ Wissen und Prozesse: Workflows / Wissensintensive Aufgaben
- ▶ Information Seeking & Searching
- ▶ Repositories, Datenbanken
- ▶ Digitale Bibliotheken und Collaboratories



# Methoden

- ▶ (Autorensysteme)
- ▶ Digitalisierung
- ▶ Informationsextraktion
- ▶ (Information Mining)
- ▶ Metadaten
- ▶ Ontologien
- ▶ Annotationen
- ▶ Digital Preservation (Langzeitarchivierung)
- ▶ (Information Retrieval)
- ▶ Recommendation
- ▶ Access Management

# Systeme

- ▶ Architekturen
- ▶ Repositories
- ▶ Web Server
- ▶ Content-Management-Systeme
- ▶ Wikis
- ▶ Digitale Bibliotheken

# Anwendungen

- ▶ DAFFODIL
- ▶ COLLATE
- ▶ Kollaborative Web 2.0-Systeme

# Information Engineering

Definition

Eigenschaften

# Was ist Information Engineering?

Eine erste Definition von **Information Engineering**:

*The application of an interlocking set of formal techniques for the planning, analysis, design and construction of information systems, applied on an enterprisewide basis or across a major sector of an enterprise.*

James Martin

# Was ist Information Engineering?

- ▶ Formale Methoden zur

- ▶ Planung
- ▶ Analyse
- ▶ Entwurf
- ▶ Realisierung

von **Informationssystemen**

- ▶ Methoden bauen aufeinander auf und sind voneinander abhängig
- ▶ Methoden werden zumeist unternehmensweit verwendet

## Was ist Information Engineering?

*Information Engineering is an integrated set of techniques, based on corporate strategic planning, which results in the analysis, design and development of systems which supports those plans exactly. Information Engineering is applied by managers and users with no knowledge of computers, but instead with an expert knowledge of their business - in conjunction with expert systems which provide rapid feedback to management for refinement of the strategic plans.*

Finkelstein

- ▶ Strategische Planung
- ▶ Kooperation und Partnerschaft von Managern und Nutzern mit Entwicklern

# Eigenschaften des IE

- ▶ Information Engineering beschreibt keine starre **Methodik**, sondern **Methoden**
- ▶ IE fördert Top-Down-Ansatz von Informationssystemen:
  1. Unternehmensweite Informationssystemplanung
  2. Analyse von Geschäftsbereichen
  3. Informationssystementwurf
  4. Informationssystemrealisierung
  5. Informationssystemeinführung



## Eigenschaften des IE

- ▶ IE unterstützt Aufbau eines sich laufend weiterentwickelnden Bestands an Wissen über ein Unternehmen
- ▶ IE unterstützt langfristige Evolution der Informationssysteme
- ▶ Orientierung an Unternehmensziele
- ▶ Benutzerbeteiligung
- ▶ Strategische Planung

# Eigenschaften des IE

Keine Beschränkung auf Unternehmen!

- ▶ Klassische Definition des IE zugeschnitten auf das "computerisierte Unternehmen"
- ▶ Web-Informationssysteme, Digitale Bibliotheken
- ▶ Offene **Communities** vs. geschlossene Unternehmen
- ▶ Vermittlung von Inhalt und Wissen an Communities
- ▶ Kollaborative Erstellung neuen Wissens und neuer Inhalte
- ▶ IE-Methoden auch auf diese Informationssysteme erweiterbar

# **Daten – Information – Wissen**

Die Begriffe Daten, Information und Wissen  
Sichtweisen auf Wissen

# Daten – Information – Wissen

- ▶ Im Zusammenhang mit Informationssystemen tauchen die Begriffe "Daten", "Wissen" und "Information" auf
- ▶ In der Informatik redet man auch von "Datenbanken" oder "Wissensmanagement"
- ▶ Was also ist der Unterschied zwischen Daten, Wissen und Information?

# Sichtweise des Wissensmanagement

- ▶ Unterscheidung zwischen Daten und Information auf der einen und Wissen der anderen Seite
- ▶ Kontinuum Daten – Information – Wissen
- ▶ Abgrenzung hinsichtlich der Fragestellung
  - ▶ Daten und Information: "wer – was – wo – und wann?"
  - ▶ Wissen: "wie?" und "warum?"
- ▶ Prozessorientierter Ansatz:
  - ▶ Wissen als Ressource, die zur Erledigung einer Aufgabe notwendig ist
  - ▶ Neues Wissen kann aus der Erledigung der Aufgabe entstehen
  - ▶ Integrierte Betrachtung von Daten, Information und Wissen

# Sicht der Informationswissenschaft

- ▶ Ursprünglich: Kontinuum Daten – Wissen – Information
- ▶ Gegensatz zur allgemein üblichen Begrifflichkeit
- ▶ Daher: Daten – Information – Wissen, mit Erläuterungen aus der Informationswissenschaft

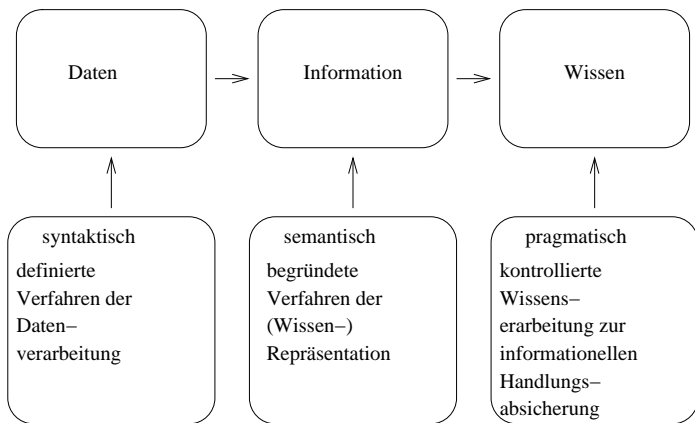
# Syntax, Semantik und Pragmatik

**Syntax:** Hierbei wird ein Objekt als Folge von Symbolen aufgefasst.

**Semantik** beschäftigt sich mit der Bedeutung eines Objektes.

**Pragmatik** orientiert sich an der Nutzung eines Objektes für einen bestimmten Zweck.

# Daten – Information – Wissen auf verschiedenen Ebenen





# Wissen

- ▶ Wissen ist die Teilmenge von Information, die von jemandem in einer konkreten Situation zur Lösung von Problemen benötigt wird
- ▶ Suche nach Wissen in externen Quellen
- ▶ Informationssysteme: Extraktion des benötigten Wissens aus der bereitgestellten Information
- ▶ Transformation von Information in Wissen als ein Mehrwert erzeugender Prozess
- ▶ Wissen ist Information in Aktion

## Beispiel: UNIX Manuals

- ▶ Beinhalten Information über Unix
- ▶ Beispiel: wie drucke ich ein Dokument?
- ▶ Benötigtes Wissen: Teilmenge der insgesamt in den Manuals verfügbaren, umfangreichen Information
- ▶ Mehrwert durch gezielte Bereitstellung (z.B. durch das Unix `man`-Kommando)

## Anderes Beispiel: Daten

T01 - T02 1:2

T03 - T04 2:0

T05 - T06 2:4

T07 - T08 1:4

T09 - T10 1:1

T11 - T12 1:0

T13 - T14 1:4

T15 - T16 1:3

T17 - T18 3:2

# Aus Daten wird Information

## Ergebnisse Fußball-Bundesliga 28. Spieltag

Hannover – Bayern	1:2
Schalke – Gladbach	2:0
Hamburg – Stuttgart	2:4
Aachen – Dortmund	1:4
Berlin – Bielefeld	1:1
Bremen – Nürnberg	1:0
Leverkusen – Bochum	1:4
Frankfurt – Cottbus	1:3
Wolfsburg – Mainz	3:2

# Wissen zur Entscheidungsfindung

**Problem: Soll ich heute mein BVB-Trikot auf der Arbeit anziehen?**



Ergebnisse Fußball-Bundesliga 28. Spieltag

Hannover – Bayern	1:2
Schalke – Gladbach	2:0
Hamburg – Stuttgart	2:4
<b>Aachen – Dortmund</b>	<b>1:4</b>
Berlin – Bielefeld	1:1
Bremen – Nürnberg	1:0
Leverkusen – Bochum	1:4
Frankfurt – Cottbus	1:3
Wolfsburg – Mainz	3:2

## Wer schönere Farben bevorzugt...;-)

**Problem: Soll ich heute mein Schalke-Trikot auf der Arbeit anziehen?**

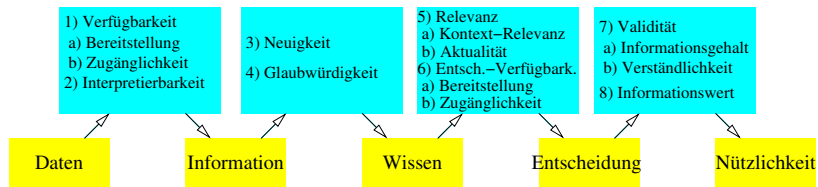


### Ergebnisse Fußball-Bundesliga 28. Spieltag

Hannover – Bayern	1:2
<b>Schalke – Gladbach</b>	<b>2:0</b>
Hamburg – Stuttgart	2:4
Aachen – Dortmund	1:4
Berlin – Bielefeld	1:1
Bremen – Nürnberg	1:0
Leverkusen – Bochum	1:4
Frankfurt – Cottbus	1:3
Wolfsburg – Mainz	3:2

# Rolle des Wissens zur Entscheidungsunterstützung

Wissen dient zur "informationellen Handlungsabsicherung"



# Sichtweisen auf Wissen

- ▶ Enzyklopädisches Wissen
  - ▶ Alltagswissen
  - ▶ Lexikalisches Wissen
  - ▶ Fachwissen
  - ▶ Erfahrungswissen
- ▶ Handlungswissen
- ▶ Räumliches Wissen
- ▶ Episodisches Wissen



# Alltagswissen

Dies sind 4 Äpfel



Unternehmen sollen Gewinne erwirtschaften



# Lexikalisches Wissen

**Schimmel** bezeichnet weißes Pferd



**Abbildung:** Weißes Pferd, Jerzy Panek

**Fohlen** bezeichnet junges Pferd... und das Team von Borussia Mönchengladbach ;-)

Gewinn ist Umsatz minus Kosten

$$G = U - K$$

Komplexe Zahlen besitzen einen Real- und Imaginärteil

$$\left( \begin{array}{c} \text{Realteil} \\ \underbrace{1.631} \\ \text{Imaginärteil} \\ \underbrace{4.73i} \end{array} \right)$$

# Erfahrungswissen

Unreife Früchte können Durchfall verursachen



**Abbildung:** Unreife Frucht der Red Savina, schärfster Chili der Welt

# Handlungswissen

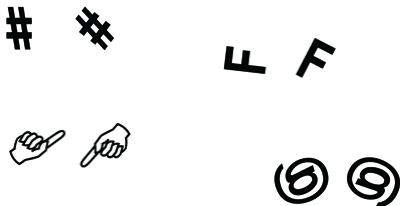
Kenntnis von Prozeduren und Problemlösungen



# Räumliches Wissen

Analoge Repräsentation physischer Objekte durch Repräsentation räumlicher Information und Verhältnisse

Beispiel: Mentale Rotation analoger Objekte





# Episodisches Wissen

Autobiografisches Wissen, das aus erinnerbaren Erlebnissen besteht

Beispiel: George spielt Golf in Yale



AP    
George W. Bush  
(als Student in Yale):  
Auf dem Campus  
chancenlos

# Implizites vs. explizites Wissen

## Implizites Wissen:

*In mentalen Modellen, kognitiven Bildern Konstrukten und Karten sowie routinierten Verhaltens- und Vorgehensweisen eingelagertes Wissen bezeichnet man als **implizites Wissen** (tacit knowledge)*

## Explizites Wissen (Polanyi)

*Spezifisches, reproduzierbares Wissen, als Know-how oder als Fingerspitzengefühl beschrieben*



# Implizites vs. explizites Wissen

Aufgabe von Informationssystemen:

- ▶ Implizites Wissen in Explizites überführen
- ▶ Explizites Wissen bei Anwendung evtl. wieder implizit

# Wissensmanagement

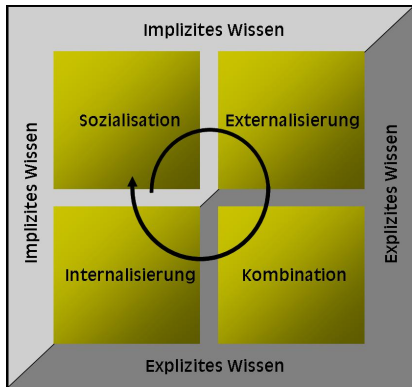
SECI-Modell

# Wissensmanagement

”Wissensmanagement beschäftigt sich mit den Möglichkeiten, auf die Wissensbasis eines Unternehmens Einfluss zu nehmen. Unter der Wissensbasis eines Unternehmens werden alle Daten und Informationen, alles Wissen und alle Fähigkeiten verstanden, die diese Organisation zur Lösung ihrer vielfältigen Aufgaben benötigt.”  
(Wikipedia)

- ▶ Wissen als Produktionsfaktor
- ▶ Informationssysteme zur Verwaltung des Wissens
- ▶ Chief Information Officers (CIO) für Informationsmanagement
- ▶ Kritik: Wissensmanagement verwendet undifferenzierten Wissensbegriff

# SECI-Modell

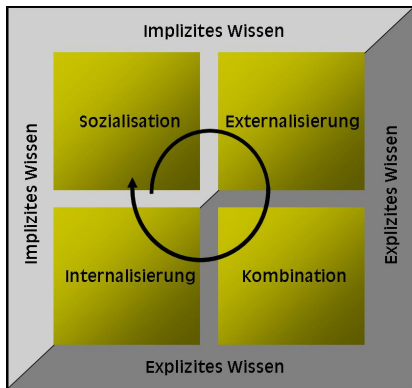


[Nonaka & Takeuchi 95]

## Modell der Wissenserzeugung

- ▶ Sozialisation
- ▶ Externalization
- ▶ Kombination
- ▶ Internalisierung

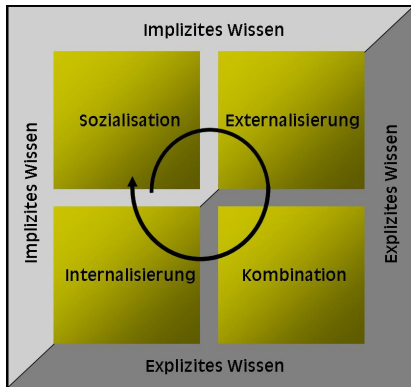
# SECI-Modell



## Sozialisation

- ▶ Teilen von Erfahrungen
- ▶ Implizites Wissen erzeugt neues implizites Wissen
- ▶ z.B. Sprache, Imitation, Beobachtung, Übung

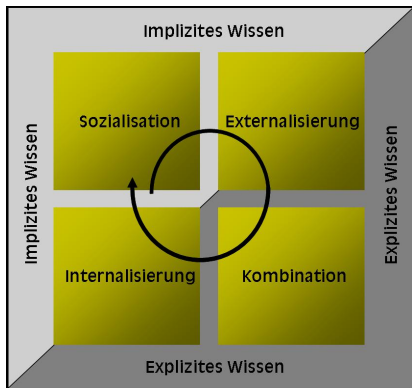
# SECI-Modell



## Externalisierung

- ▶ Implizites Wissen wird artikuliert
- ▶ z.B. Modellbildung
- ▶ Schlüsselprozess der Wissensumwandlung
- ▶ Erschaffung expliziten Wissens aus Implizitem

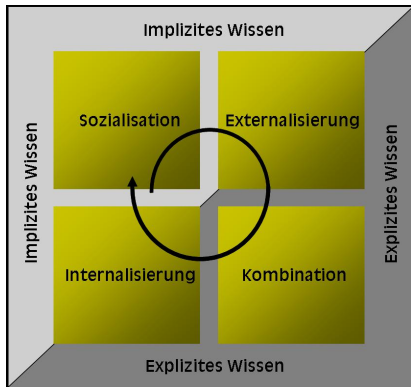
# SECI-Modell



## Kombination

- ▶ Konzepte werden in ein Wissenssystem eingebunden
- ▶ Isolierte Teile zu einem gemeinsamen Ganzen

# SECI-Modell



## Internalisierung

- ▶ Explizites Wissen wird (zu Implizitem) verinnerlicht
- ▶ “learning by doing”



# Information Lifecycle

Der Information Lifecycle im Information Engineering

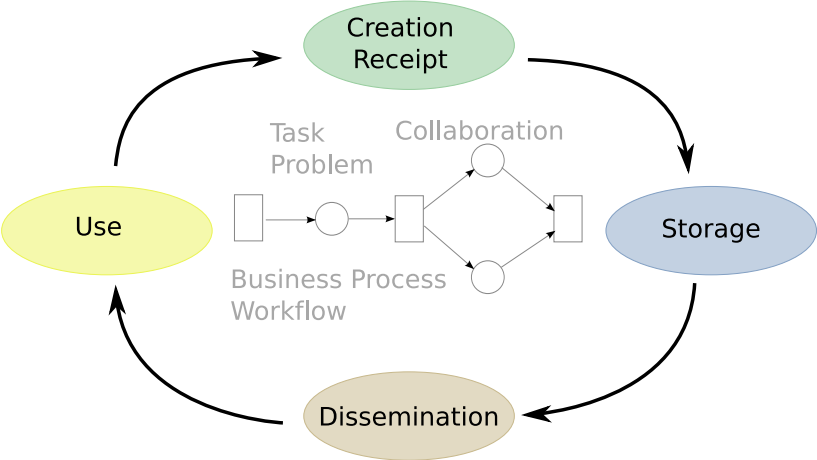
Phasen des Information Lifecycle

IE-Methoden im Information Lifecycle

## Information Lifecycle im IE

- ▶ In dieser VL: Konzentration auf Methoden und Werkzeuge des IE
- ▶ Grundlage der weiteren Betrachtungen: **Information Lifecycle**
- ▶ Betrachtung aller Lebensphasen von Information (Erstellung bis Anwendung)
- ▶ Ideal: Informationssystem unterstützt alle Phasen
- ▶ Einsatz von Methoden hängt vom Informationssystementwurf ab (nach Planung und Analyse)
- ▶ Implementierung und Integration der Methoden und Werkzeuge Teil der Realisierung des Informationssystems

# Information Lifecycle



# Information Lifecycle

- ▶ Information dient der Erfüllung einer Aufgabe
- ▶ Formulierung als Prozess
- ▶ Teilaufgaben als untergeordnete Prozesse
- ▶ **Wissensintensive Aufgaben**

# Creation/Receipt

- ▶ Autorensysteme (Authoring)
  - ▶ Erstellung einfacher Texte
  - ▶ Hypertext- und Multimediadokumente, Hypermedia
- ▶ Metadaten
  - ▶ Daten über Dokumente
  - ▶ z.B. bibliographische Daten
- ▶ Annotationen
  - ▶ Spezielle Art von Metadaten
  - ▶ Verschlagwortung, "tagging" bzw. semantische Annotation
  - ▶ Textuelle Kommentare, Diskussionen
- ▶ Markup

# Creation/Receipt

- ▶ Digitalisierung
  - ▶ Erfassung und Umsetzung von nicht-digitalen Quellen in digitale
  - ▶ Textuelle Dokumente (Scanning, OCR)
  - ▶ Nicht-textuelle Daten (Schall, Messdaten, Video, Bilder) → Sampling
- ▶ Modellierung
  - ▶ Formale Darstellung von Prozessabläufen, Sequenzen, Anwendungsfällen, Zuständen und Beziehungen

# Storage

- ▶ Kurz- bis langfristige Speicherung der Information
- ▶ Flüchtige Information im Hauptspeicher
- ▶ Ablage in Dateien
- ▶ Größere Datenmengen, (semi-)strukturierte Daten: Datenbanken
- ▶ Dokumente: Digitale Bibliotheken/Repositories
- ▶ Langzeitarchivierung (digital preservation)
- ▶ Wissensrepräsentation und Organisation, Ontologien
- ▶ Aufbereitung für die Suche: Indexierung

# Dissemination

- ▶ Information Retrieval
- ▶ Recommendersysteme
- ▶ Zugangsmanagement (access management)



# Use

- ▶ Informationssysteme dienen häufig der Erfüllung einer bestimmten Aufgabe oder eines bestimmten Zwecks
- ▶ Informationssystem kann Anwendungen gezielt unterstützen
- ▶ Beispiel Geisteswissenschaften: Interpretation des Materials mittels annotationsbasierter Diskussion (damit auch Generierung neuer Information)
- ▶ Beispiel Informationssystem für Multimediadokumente: Abspielfunktionen für Videos oder Musik

# Zusammenfassung

# Zusammenfassung

- ▶ Begriff des Information Engineering erläutert
  - ▶ Keine starre Methodik, sondern Methoden und Werkzeuge
  - ▶ Konzentration auf Realisierung von Informationssystemen
- ▶ Daten – Information – Wissen
- ▶ Wissensmanagement
- ▶ Information Lifecycle
  - ▶ 4 Phasen
  - ▶ Mögliche Methoden für jede Phase