Wie intelligent können unsere Daten denn noch werden?

Intelligence? What’s Intelligence in TC?

Prof. Dr. Wolfgang Ziegler

Hochschule Karlsruhe, Kommunikation und Medienmanagement
Institute for Information and Content Management I4ICM
1987-1997 Theoretical Physics (Ph.D.), Würzburg Univ., Germany

1997 Tech. Doc Services (CMS Consultant, Developer XML/XSLT)

2001 Central Documentation Services (LIEBHERR)

2003 Professor of Information & Content Management (Karlsruhe Univ. Appl. Sciences)

Independent CMS & CDP consultant
I4ICM Institute for Information and Content Management (I4ICM)
(Methodologies: REx, PI-Class®, PI-Mod, Content-Delivery, CoReAn)
Introduction

Various Aspects of Intelligent Systems
Intelligence helps!

- to systemize
- to automize
- to understand?
CMS principle: Controlled reuse of content modules (topics) in multiple documents or media by the use of metadata

CMS offer solutions for
- variant management
- version management
- translation management
- cross media publishing management

CDP principle: Dynamic content delivery for electronic (web) media supported by metadata driven search functionalities

©Prof. Dr. Ziegler
**Content Delivery Portals**

*Systems offering web based access to modular, aggregated content or other information for various user groups by related retrieval mechanisms.*

**Basic functionalities**

- Access or import content from relevant data sources and corresponding systems
- Manage and update content within the content lifecycle
- Retrieval functionalities including user interfaces for content searching by users.
- Web-based display of content on a modular or document based level
- Web services handling requests from other applications and events.
Mobile devices and corresponding apps used in private spheres, and that are consequently being demanded also in business use

Search driven information access from retrieval applications which includes facetted and filtered searching for consumer goods

Virtual reality and augmented reality enriching visual environments by additional information

Machine translation and its growing acceptance in private and business contexts

Artificial intelligence technologies which allow the automated extraction of decision-relevant information and knowledge from various data sources

Big data technologies for analyzing mass data generated for example from customer behavior in sales processes or from the tracking of machine states

IoT (Internet of Things) and Industry 4.0 initiatives which aim to control and react to the real-time behavior of products.
Industry 4.0 / IoT (Phase III) Content Delivery & Retrieval

Intelligent Product I4.0/IoT-ready

- Requirement for Standardized Exchange Format
- Supplier
- Additional Information & Sources
- On Site CDP
- Off Site Portal/Mobile CDP

User Information Service Information Machine State
- Machine state (errors, messages, operating conditions)
Intelligent Content
Methodologies, technologies, systems which allow to **automize** (or at least systemize) content processes. Processes cover content aggregation, content delivery, (retrieval, search, request) and quality assurance.

Content Intelligence
Methodologies, technologies which allow to **track** and **check** content and content related processes. Processes cover content creation and metadata use, content reuse in CMS and content use in CDP (overlap with artificial intelligence).
The Intelligence Cascade

- **Native Intelligence**
  Semantic content and semantic metadata for process automation

- **Augmented Intelligence**
  Additional relations between content objects described e.g. by ontologies

- **Artificial Intelligence**
  Automated extraction of metadata and knowledge by statistical methods
Native Intelligence
Native Intelligence

- Purpose of native intelligence is to automize processes within CMS or to access information stored in CDP
- Consists of
  - Semantic information modeling
    - Standards: DITA, DocBook, S1000D, ... PI-Mod
    - Custom Structures: CMS-dependant
    - Technologies: XML, (DTP)
  - Semantic metadata
    - Standards: Dublin Core, S1000D, iSPEC2200, RDS-PP, eClass, iiRDS
    - Technologies: RDF, Turtle, Notation 3, JSON, ...
    - Methodologies: PI-Classification, ...
Physical Objects (Product Components)

Content Objects (Modular Topics)

- Information Classes
- Operation
- Dismount
- Repair
The basic dimensions of (semantic) meta data for modular content topics

1. **Product-Class**
   - Component
   - Product
   - Extended Metadata (Variant Features/Properties & Functional)

2. **Information-Class**
   - Information Type
   - Information Product
Basic Dimensions of Module Classification

Adjusting the height

The height adjustment is infinitely variable.

1. Hold the functional unit 1 by the handle at the top.
2. With the other hand loosen the retaining screw 2 at the rear of the telescopic rod.
3. Pull the functional unit 1 to the desired height.
4. Tighten the retaining screw 2 by hand again hand tight.
5. Release the handle.

→ The device is now at the desired height.

Product-Class

Base/Telescopic Rod
X3B, X3-H1, X5-B, X5-D,...

Information-Class

Operation/Height Adjustment
User Manual, Service Manual,...
Classifications of Components

<table>
<thead>
<tr>
<th>Product class 1</th>
<th>Product class 2</th>
<th>Product class 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Assembly/Functional group)</td>
<td>(Components/Functional unit)</td>
<td>Part</td>
</tr>
<tr>
<td>Complete device</td>
<td>Drive</td>
<td></td>
</tr>
<tr>
<td>Drive</td>
<td>Drive</td>
<td></td>
</tr>
<tr>
<td>Connection</td>
<td>Connection</td>
<td></td>
</tr>
<tr>
<td>Electric motor</td>
<td>Electric motor</td>
<td></td>
</tr>
<tr>
<td>Gear box</td>
<td>Gear box</td>
<td></td>
</tr>
<tr>
<td>Lighting</td>
<td>Cover</td>
<td></td>
</tr>
<tr>
<td>Cover</td>
<td>Cover</td>
<td></td>
</tr>
<tr>
<td>Light fitting</td>
<td>Light fitting</td>
<td></td>
</tr>
<tr>
<td>Heating</td>
<td>Heating element</td>
<td></td>
</tr>
<tr>
<td>Heating element</td>
<td>Heating element</td>
<td></td>
</tr>
<tr>
<td>Rotor</td>
<td>Blade</td>
<td></td>
</tr>
<tr>
<td>Blade</td>
<td>Blade</td>
<td></td>
</tr>
<tr>
<td>Blade mount</td>
<td>Blade mount</td>
<td></td>
</tr>
<tr>
<td>Impeller</td>
<td>Impeller</td>
<td></td>
</tr>
<tr>
<td>Protection</td>
<td>Safety grille</td>
<td></td>
</tr>
<tr>
<td>Safety grille</td>
<td>Safety grille</td>
<td></td>
</tr>
<tr>
<td>Display_operating element</td>
<td>Speed controller</td>
<td></td>
</tr>
<tr>
<td>Speed controller</td>
<td>Speed controller</td>
<td></td>
</tr>
<tr>
<td>Swivel control</td>
<td>Swivel control</td>
<td></td>
</tr>
<tr>
<td>Temperature probe</td>
<td>Temperature probe</td>
<td></td>
</tr>
<tr>
<td>Temperature control</td>
<td>Temperature control</td>
<td></td>
</tr>
<tr>
<td>Display</td>
<td>Display</td>
<td></td>
</tr>
<tr>
<td>Mounting bracket</td>
<td>Base</td>
<td></td>
</tr>
<tr>
<td>Base</td>
<td>Base</td>
<td></td>
</tr>
<tr>
<td>Base plate</td>
<td>Base plate</td>
<td></td>
</tr>
<tr>
<td>Telescopic rod</td>
<td>Telescopic rod</td>
<td></td>
</tr>
<tr>
<td>Ceiling mount</td>
<td>Ceiling mount</td>
<td></td>
</tr>
</tbody>
</table>

Analogous procedure of component-based decomposition and classification of software products:
- software components
- software classes.objects
- GUI components
- programming units
# Classification of Information Types

## Taxonomy of (intrinsic) Information Classes

<table>
<thead>
<tr>
<th>Information class 1</th>
<th>Information class 2</th>
<th>Information class 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Module type)</td>
<td>(Product life cycle)</td>
<td>(Detailed product life cycle)</td>
</tr>
<tr>
<td>Manual</td>
<td>Operation</td>
<td>Blower adjustment</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Height adjustment</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tilt adjustment</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Swivel activation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Temperature adjustment</td>
</tr>
<tr>
<td>Getting started</td>
<td></td>
<td>Check</td>
</tr>
<tr>
<td>Storage</td>
<td></td>
<td>Repair</td>
</tr>
<tr>
<td>Assembly</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maintenance</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Layout</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Disposal</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Function</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tech. data</td>
<td></td>
</tr>
<tr>
<td>Plan</td>
<td>Diagnostics</td>
<td>Error code</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Manual</td>
</tr>
<tr>
<td>Safety</td>
<td>General safety</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intended use</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Specific safety</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Foresee. misuse</td>
<td></td>
</tr>
</tbody>
</table>
Classification of Products

Taxonomy of (extrinsic) Product Classes

<table>
<thead>
<tr>
<th>Series</th>
<th>Model range</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>T series</td>
<td>T3 range</td>
<td>T3-B, T3-H1</td>
</tr>
<tr>
<td></td>
<td>T5 range</td>
<td>T5-B, T5-DH1, T5-DH2</td>
</tr>
<tr>
<td></td>
<td>TP range</td>
<td>TP-B, TP-DH1, TP-DH2</td>
</tr>
<tr>
<td>X series</td>
<td>X3 range</td>
<td>X3-B, X3-H1</td>
</tr>
<tr>
<td></td>
<td>X5 range</td>
<td>X5-B, X5-D, X5-DH1, X5-DH2</td>
</tr>
<tr>
<td></td>
<td>XP range</td>
<td>XP-B, XP-D, XP-DH1, XP-DH2</td>
</tr>
</tbody>
</table>
CMS Taxonomies from Topic Classification

Content Topics

- Heating
- Display
- Rotor
- X3B
- T3B
- Multidimensional Information Space

- Safety
- Functional Description
- Repair

- User Manual
- Service Manual

©Prof. Dr. Ziegler
Implementation of PI-Methodology
CMS „Taxonomies“ from Topic Classification

Intrinsic Taxonomies

Heating
Display
Rotor

Extrinsic Hierarchies

X3B
T3B

Variant properties

Multidimensional Information Space

Content Topic

Safety
Functional Description
Repair

User Manual
Service Manual

Hierarchies, Taxonomies, List, ...

Variants

Intrinsic Taxonomies

Extrinsic Hierarchies

Intrinsic Taxonomies

Extrinsic Hierarchies

Multidimensional Information Space

Functional Metadata
CMS „Taxonomies“ from Topic Classification

Multidimensional Information Space

Content Topic

Variant Properties/Features

Location

Geometry

Parts No

Material

Features

Functional Metadata (Collections)
Product features included or in addition to (extrinsic) product classes / names

Example PI-Fan: Type TP-DH2 (Combination type)
- Table Fan (T)
- Continuous Switch (P)
- Display (D)
- Heating (H)
- 2-Level Heating (2)

Goal: Facilitate planning of new variants/configurations and metadata handling in CMS
Einstellen der Gebläsestärke


- Drehen Sie den Drehregler (Abb. 23) bis die gewünschte Stärke erreicht ist.

Die Stärke des Gebläses kann ebenfalls an der Fernsteuerung eingestellt werden (s. Abschnitt 5.4 Fernsteuerung).

- Drehen Sie den Drehregler (Abb. 23) bis die gewünschte Stärke erreicht ist.

Die Stärke des Gebläses kann ebenfalls an der Fernsteuerung eingestellt werden (s. Abschnitt 5.4 Fernsteuerung).
CMS – CDP-Kopplung
Content Delivery Portal (PI-Fan)

Structured Search
Facets
Navigation

Direct Search

Rotor

Cleaning the rotor

Mounting the rotor

Schwenkbereich freiräumen

Docufy
Topic Pilot  [www.pi-fan.de]
Bedienungsanleitung T3-B

Schaema
Content Delivery Server
[www.pi-fan.de]
Schema

Teleskopstange und Standplatte montieren

1. Befestigen Sie die Justierschraube 1 an der Teleskopstange 2, wie in Abbildung 1 zu sehen und ziehen Sie die Justierschraube handfest an.
2. Stecken Sie die Teleskopstange 2 in die runde Aussparung an der Oberseite des Bodenständers 3.
3. Befestigen Sie die Teleskopstange 2 am Bodenständer 3, indem Sie die Befestigungsschraube 4 durch den Bodenständer stecken, und handfest anziehen.

Die Teleskopstange ist nun am Bodenständer befestigt.

[www.pi-fan.de]
CDP and Semantic Search

Suchen wie man spricht

Topics in der Trefferliste

Empolis
Content Express
[www.pi-fan.de]
CDP as Content Service

Practice innovation
IDS c-rex.net

[www.pi-fan.de]
Limitations of Classical Metadata

- Taxonomies and hierarchies are two-dimensional descriptions of a three-dimensional world
More Complexity (and Dimensions)

Intrinsic Taxonomies

Extrinsic Hierarchies

Mutidimensional Information Space

Content Topic

- Safety
- Repair

- User Manual
- Service Manual

Variant Features/Properties

Content Topic

- X3B
- T3B

Fotor

Display

Heating

Hierarchies, Taxonomies, List, ...

Functional Metadata
Typical challenges

- Multi occurrences of product components at different locations (in taxonomy)
- Relations between product components; Dependencies of topics on combinations of components
- Dependencies of additional variant properties on product components
- Dependencies of information types on other taxonomic values
Augmented Intelligence
Augmented Intelligence

- Purpose of Augmented Intelligence is to model the complexity of real world products and information
- Overcome typical shortcomings of the taxonomic modelling of metadata
- Introduce model of objects, their properties and (conditional) relations between each other
More Complexity (and Dimensions)

Intrinsic Taxonomies
- Heating
- Display

Extrinsic Hierarchies
- X3B
- T3B

Mutidimensional Information Space

Content Topic

Safety
- Functional Description
- Repair

User Manual

Service Manual

Functional Metadata

Variant Features/Properties

Hierarchies, Taxonomies, List, ...
Usual CMS / CDP Data Exchange
Augmenting CMS / CDP by Ontologies
Augmenting CMS by Ontologies

CMS

CDP

RDF, HTML, XML..., ZIP...

Hochschule Karlsruhe Technik und Wirtschaft
UNIVERSITY OF APPLIED SCIENCES
Ontology modelling of PI-Fan (I)

Source: Ontolis

©Prof. Dr. Ziegler
Publishing from Ontology modelling of PI-Fan (I)

3 Montage
3.1 Rotor montieren

VORSICHT

Verletzungshaft durch Vorschläuche von Kleinteilen
- Halten Sie Kleinkinder und Haustiere vom Ort der Montage des Geräts fern, um das Vorschläuche von Kleinteilen wie Schrauben zu vermeidern!

- Vorgehen
1. Schieben Sie die Schutzgitter-Rückseite 1 auf den Antrieb 2, wie in Abbildung 1 zu sehen.
2. Befestigen Sie nun die Schutzgitter-Rückseite 1 mit der Befestigungsmutter 3, indem Sie sie mit einem Kreuzschlitz-Schraubenzieher handfest anziehen.
   - Die Schutzgitter-Rückseite ist nun montiert.
3. Schieben Sie die Befestigungsschraube 4 durch den Rotor 5.
4. Schieben Sie nun die Befestigungsschraube 4 durch die Befestigungsschraube 3 und ziehen Sie sie mit einem Kreuzschlitz-Schraubenzieher handfest an.

Source: Ontolis
Augmenting CMS or CDP by Ontologies
Multilingual Terminology & Ontology

- image editing
- photo editing
- picture editing

BROADER & NARROWER

Source: Coreon
Ontology modelling of Product/Terminology

Source: Coreon
Ontology modelling of Product/Terminology (II)

Source: Coreon
Augmenting CDP by Ontologies
Ontology modelling of PI-Fan (III)

Source: Intelligent views / K-infinity
Standardizing Exchange by Ontologies
Artificial Intelligence (in TC)
Purpose of AI is the extraction of knowledge from (large/big) data and content sources

→ Assigning content automatically to a given ontology or taxonomy

Scenarios (in TC and within context of CMS/CDP)
- Migration of legacy data (e.g., between CMS)
- Structured access to unstructured content
- Quality Control (classification, duplicates)
- ...
Automated Text Classification

- Training data (pre-classified topics)
- Test data/topics (unclassified)
- Algorithm (Weighting)
- Model
- Prediction/Classification

Confidence Measurement
(by cosine similarity)


\[ p = \frac{s_1 - s_2}{s_1 - s_n} \]
Document Reconstruction of Information Classes

Content Intelligence
Types of Content Intelligence

- Content Control
  - Controlled language checker CLC
  - Terminology checking
  - Similarity analysis (content duplicates) in CMS

- Reuse tracking in CMS: Report Exchange (REx) Metrics

- Use tracking in CDP: Content Relevance Analytics (CoReAn)
Similarity Measurements of Topics

J. Oevermann; fastclass.de
Content Intelligence
REx & Content Relevance Analytics

Business Intelligence (REx)

→ Metrics:
  • Reuse Rates (Abundancy)
  • Redundancy
  • Document Sharing factor
  • Variant management
  • Correlations; Distributions

... Artificial Intelligence

→ Quality assurance:
  • Similarity analysis
  • Classification quality

Web Analytics (CoReAn)

Indirect feedback

→ Metrics:
  • visiting time,
  • Visit frequency
  • search behaviour
  • search terms
  • ...

Direct feedback

→ Rating
→ Satisfaction

→ Improve:
  • Product
  • Information
  • Terminology (Harvesting)
Content is managed in CMS by **Native Intelligence**, and in many other data sources. **Artificial Intelligence** and Content Intelligence (REx, ...) can ensure content quality, variant management and reuse efficiency in CMS.

**Augmented Intelligence** will/might drive the interaction between product and information development.

Industry 4.0 / IoT concepts rely on precise identification of parts and corresponding information delivered by CDP:

- **Native Intelligence** of data is required to ensure retrieval precision.
- **Artificial Intelligence** can ensure data/metadata quality.

Manual search & retrieval processes in CDP rely on most complete network of information.

**Augmented intelligence** reveals network and related content.

**Artificial Intelligence** can assign structured and unstructured content (documents and parts of them) automatically to the information network to reduce manual effort.
How intelligent will our content be in the future?

- Depends on the use case and level of digitization

- The Digitization Cascade → The Intelligence Cascade
  - Do → Native Intelligence
  - Let Guide → Augmented Intelligence
  - Let Do → Artificial Intelligence

- Smart Examples
  LifeSciences, Navigation, Consumer, ..., Documentation/Information)
In: tekom Schriftenreihe Band 22

www.i4icm.de