Wie (gut) wird unser CMS genutzt?

Content Intelligence
auf Basis von REx-Kennzahlen

Prof. Dr. W. Ziegler - Hochschule Karlsruhe
Fak. Informationsmanagement & Medien

REX

Fragen an das Auditorium

- Wer nutzt CMS?
- Wer wurde mit Einführung eines CMS effizienter?

- Wer kennt die Modulzahl im CMS?
- Wer kennt den Wiederverwendungsgrad?
- Wer kennt die (modularen) Erstellungszeiten?

- Wer erhebt regelmäßig Kennzahlen?

REx

Agenda

- Methodische Grundlagen
- Das REx-Format und die CMS-Metrik
- Prototypische Daten und REx-Nutzung
- Ausblick und Entwicklungen
- Zusammenfassung und Diskussion

Methodische Grundlagen

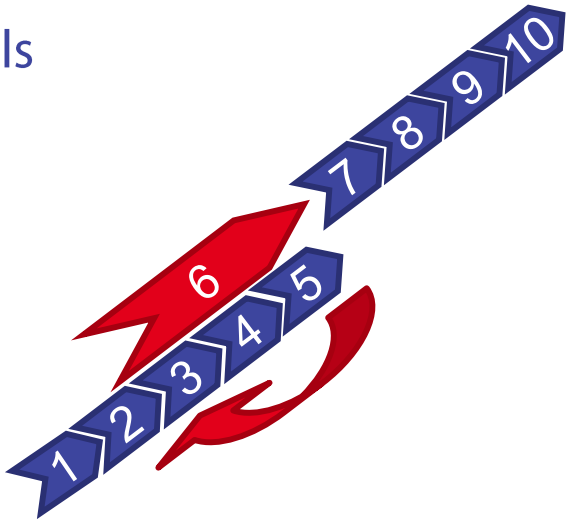
Content Management Methoden

- Modularisierung
- Klassifikation
- Variantenmanagement
- Aggregation
- Prozessmanagement (Content Lifecycle und Translation Management, Publishing, Delivery)

REx

Vorgehensmodell (Systemeinführung)

1. Analyse des IST-Zustandes
2. Darstellung des Optimierungspotenzials
3. Definition der SOLL-Zustandes
4. Evaluierung der Systeme und Auswahl
5. Erstellung eines Pflichtenheftes
6. Interne Vorbereitung
7. Systemanpassung
8. Installation des Systems
9. Schulungsmaßnahmen
10. Systemnutzung

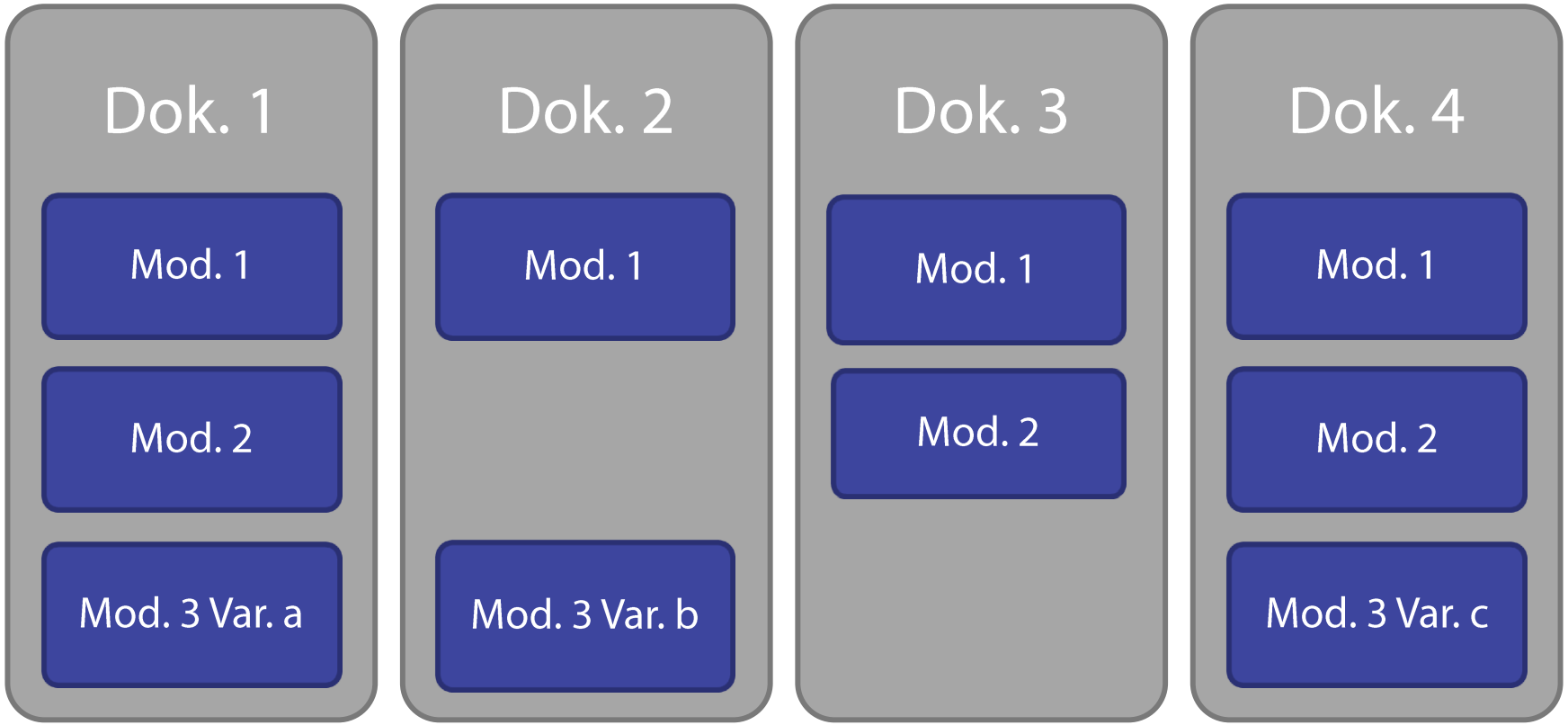


tekom CMS-Studie
2005/2008

REx

Optimierungs- und Wiederverwendungspotenzial

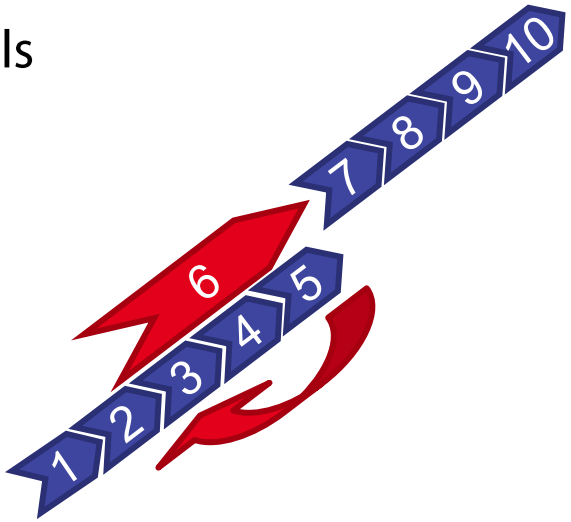
- Modulmatrix: Analyse und Modulplanung



REx

Vorgehensmodell (Systemeinführung)

1. Analyse des IST-Zustandes
2. Darstellung des Optimierungspotenzials
3. Definition der SOLL-Zustandes
4. Evaluierung der Systeme und Auswahl
5. Erstellung eines Pflichtenheftes
6. Interne Vorbereitung
7. Systemanpassung
8. Installation des Systems
9. Schulungsmaßnahmen
10. Systemnutzung

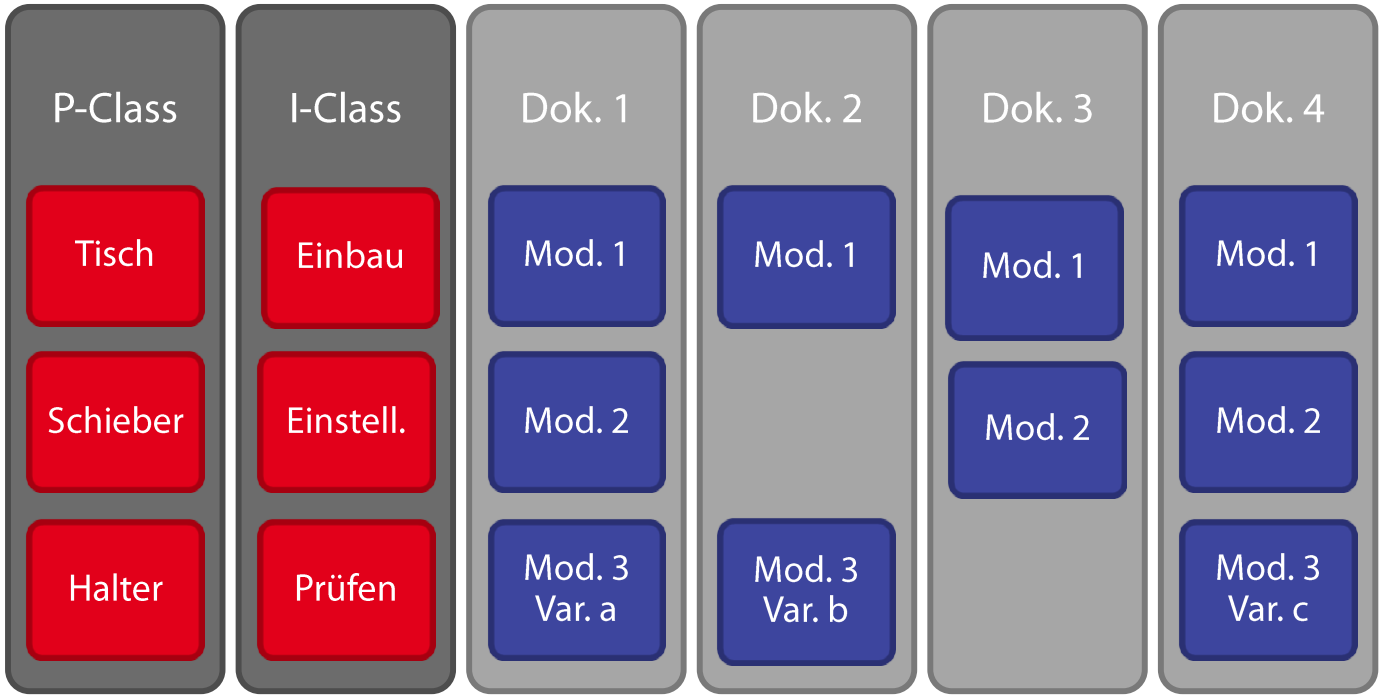


tekom CMS-Studie
2005/2008

REx

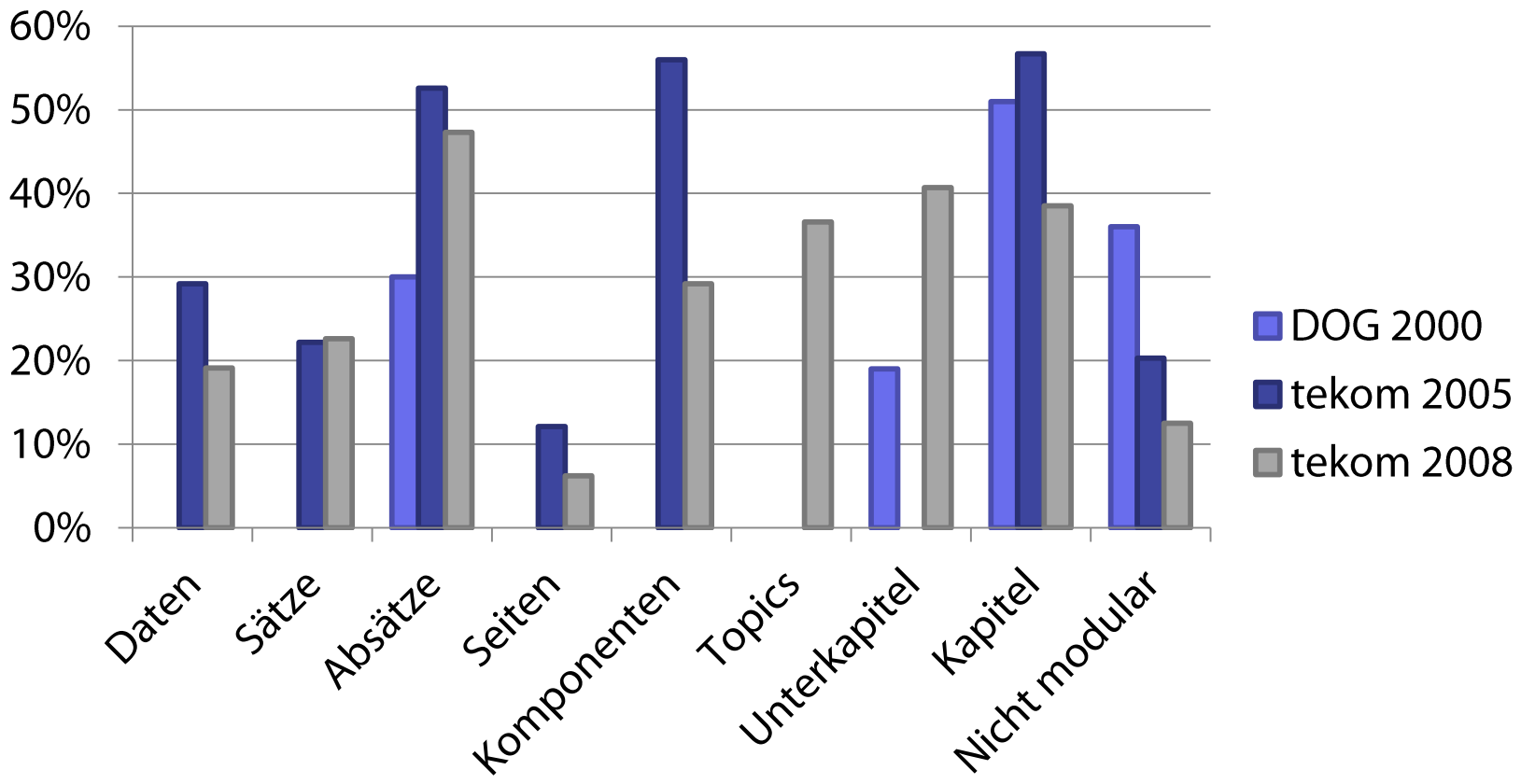
Klassifikation von Modulen

- Modulmatrix: Metadatenplanung für CMS-Einsatz, z. B. PI-Klassifikation



REx

Modularisierung

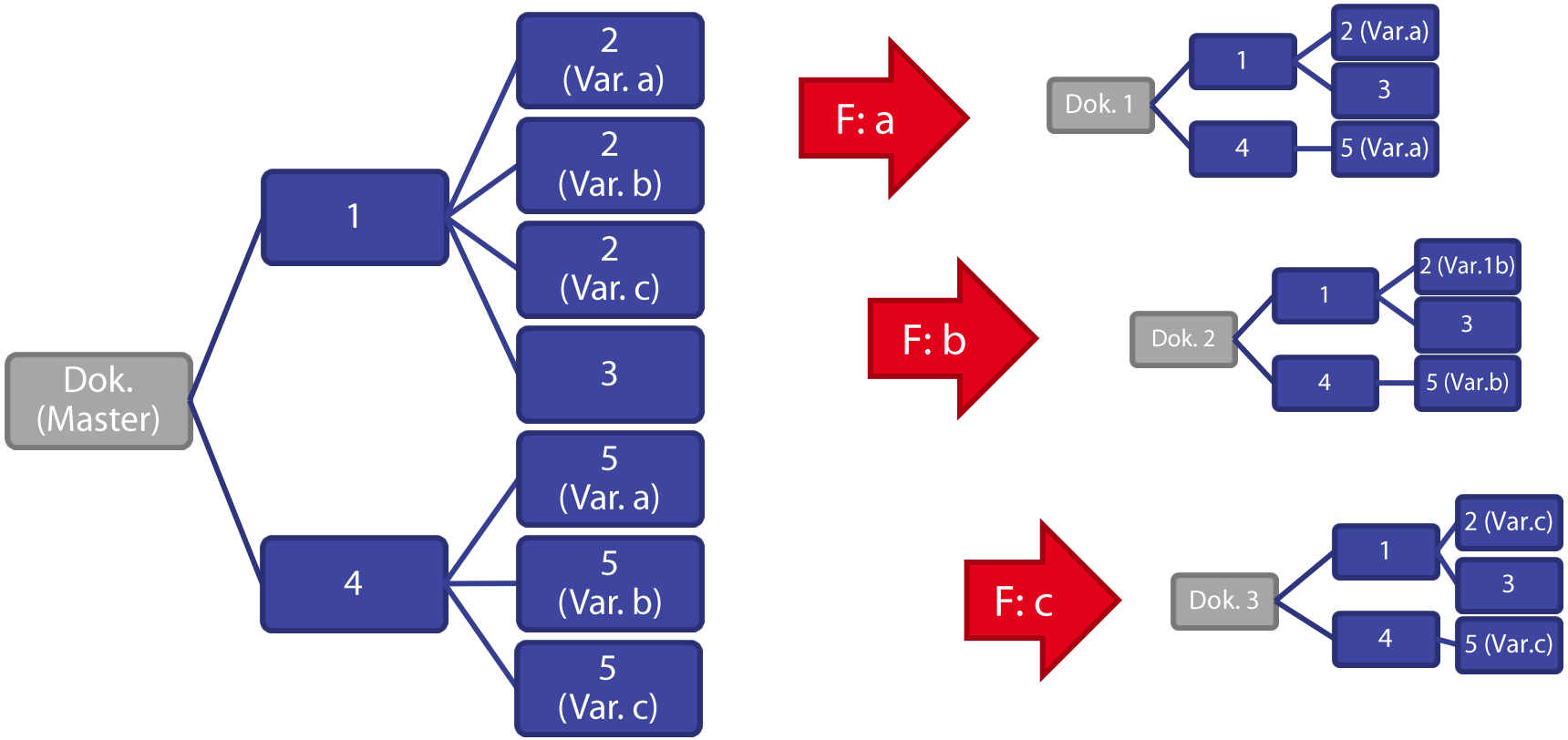


- tekom CMS-Studie 2005/2008;
- W. Ziegler: „Die richtige Größe finden“, Produkt Global 2007

REx

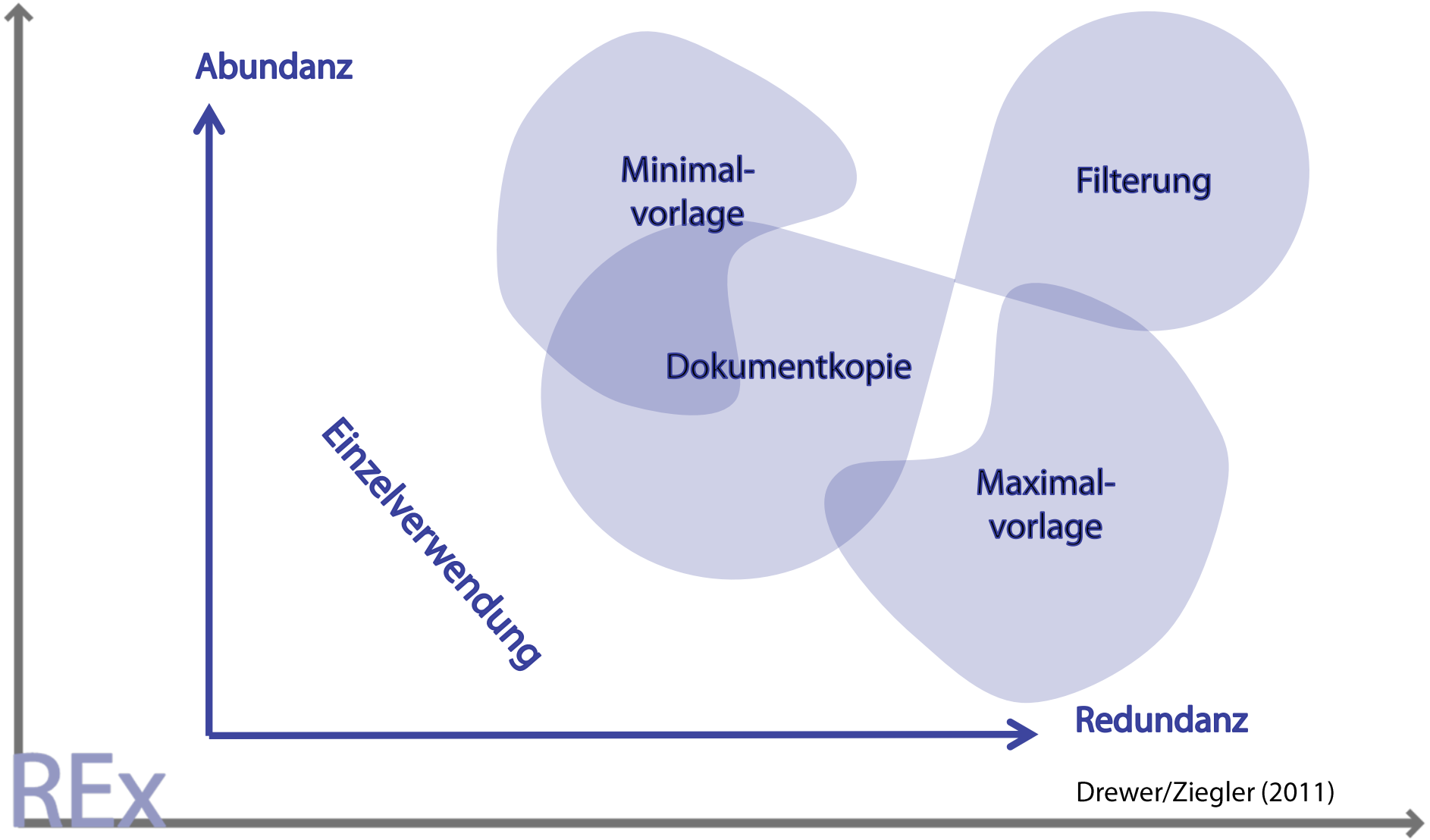
Variantenmanagement (Dok.ebene)

- Filterung von (Master-)Dokumenten



REx

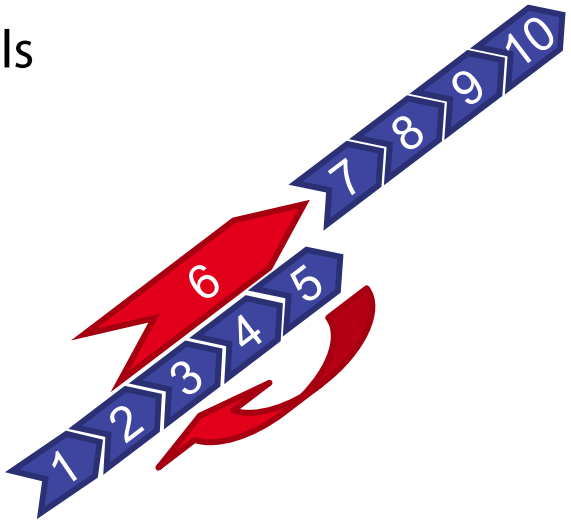
Erstellungsmethoden



Drewer/Ziegler (2011)

Vorgehensmodell (Systemeinführung)

1. Analyse des IST-Zustandes
2. Darstellung des Optimierungspotenzials
3. Definition der SOLL-Zustandes
4. Evaluierung der Systeme und Auswahl
5. Erstellung eines Pflichtenheftes
6. Interne Vorbereitung
7. Systemanpassung
8. Installation des Systems
9. Schulungsmaßnahmen
10. Systemnutzung

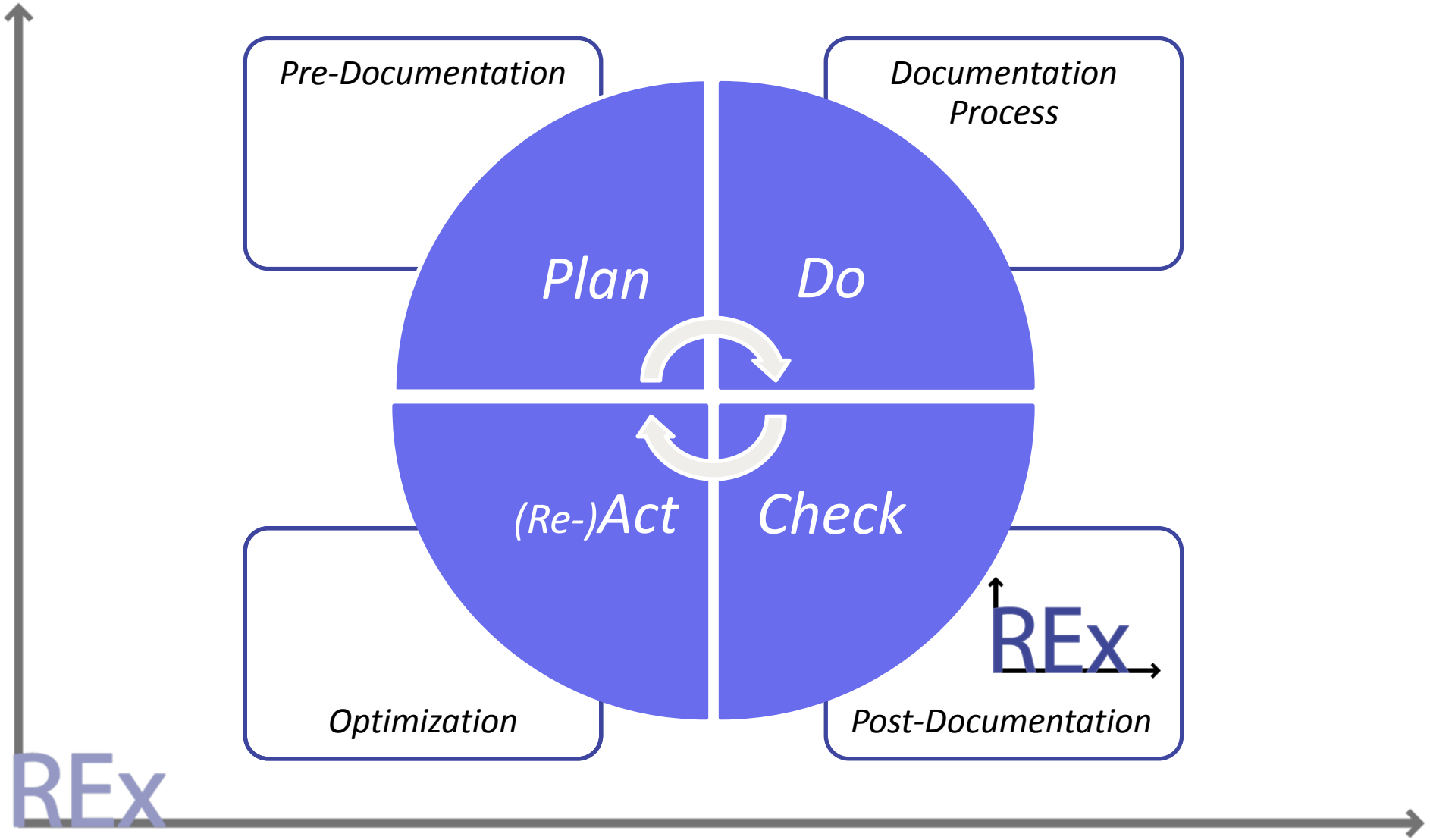


tekom CMS-Studie
2005/2008

REx

Systemnutzung und Prozessmanagement

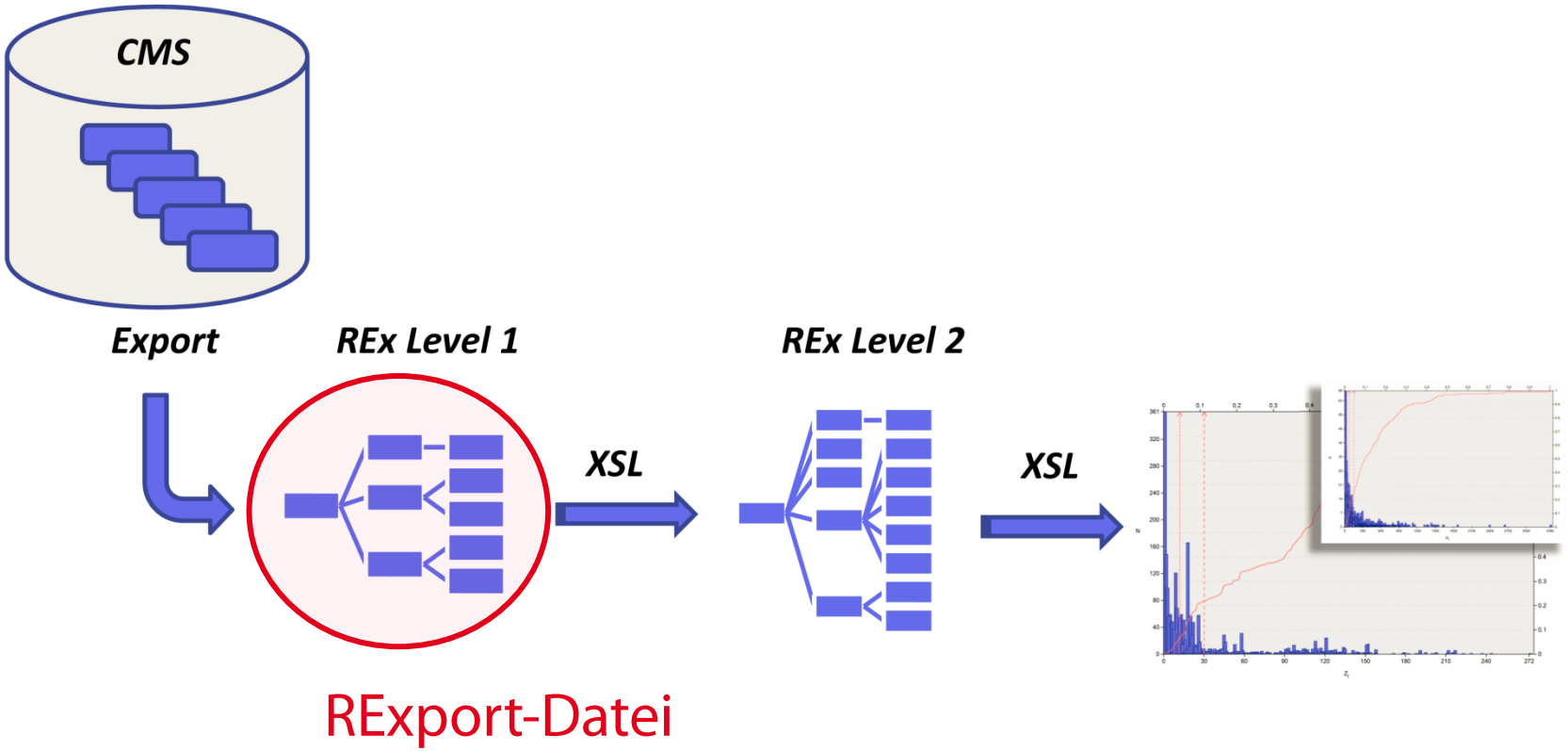
basierend auf
PDCA-Zyklus
nach Deming



Das REx-Format und die CMS Metrik

REx

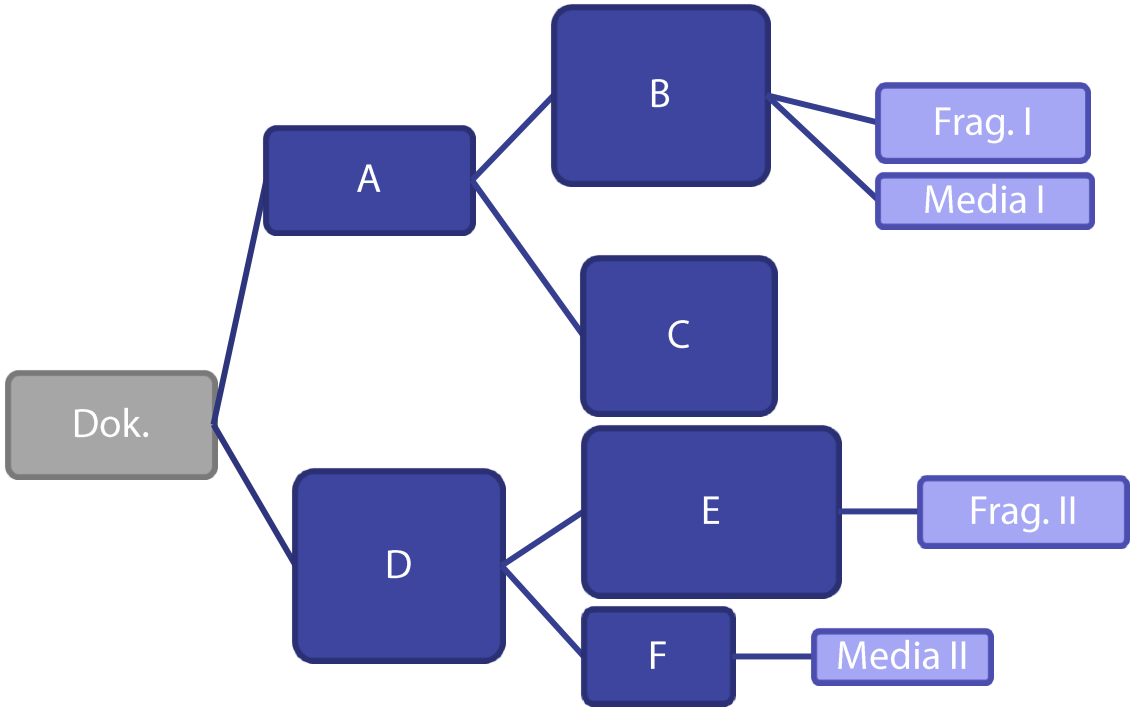
REx und CMS



REx

Report-Exchange-(REx) Format

- CMS exportiert Informationen über Objekte und hierarchische Aggregation (Verwendungsnachweis)



REx

Report-Exchange-(REx) Format

Beispiel

```
<?xml version="1.0" encoding="UTF-8"?>
<REx1 xmlns:xs="http://www.w3.org/2001/XMLSchema"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xmlns:fn="http://www.w3.org/2005/xpath-functions" rexVersion="1.1" rexLevel="1"
    xsi:noNamespaceSchemaLocation="..\xsd\REx_1.1-Level_1.xsd">
    <meta>
        <sysOwner>none</sysOwner>
        <sysType>none</sysType>
        <sysTime>2012-09-29T16:05:08Z</sysTime>
        <sysVersion>none</sysVersion>
        <analysisType>
            <sysGlobal mediaChannel="all" lang="all" docType="all"/>
        </analysisType>
        <default_segmentType>word</default_segmentType>
    </meta>
    <raw>
        <node node_Type="doc" id="PSL-BA-BA-120_0-120_V001-2011.04.18-de.xml" lang="de" version="0"
            mediaChannel="none" docType="PSL" dateFinal="2011-04-18T14:54:12Z"
            dateInitial="2011-04-18T14:44:29Z">
            <size>
                <none>1</none>
            </size>
            <reuse id="47126-de.xml" lang="de" version="3"/>
            <reuse id="1212-de.xml" lang="de" version="29"/>
            <reuse id="1210-de.xml" lang="de" version="10"/>
            <reuse id="1215-de.xml" lang="de" version="10"/>
            <reuse id="108985-de.xml" lang="de" version="0"/>
            <reuse id="91712-de.xml" lang="de" version="1"/>
```

REx

Report-Exchange-(REx) Format

Beispiel

```
95500        <node node_Type="module" id="E20705-de.xml" lang="de" version="3" docType="XMLJDoc"
95501            dateFinal="2007-04-26T08:23:53Z" dateInitial="2007-04-26T08:23:53Z">
95502            <size>
95503                <word>200</word>
95504                <char>1662</char>
95505                <element>98</element>
95506            </size>
95507            <reuse id="12877-de.xml" lang="de" version="4"/>
95508            <reuse id="12877-de.xml" lang="de" version="4"/>
95509            <reuse id="12877-de.xml" lang="de" version="4"/>
95510            <reuse id="12877-de.xml" lang="de" version="4"/>
95511            <reuse id="12877-de.xml" lang="de" version="4"/>
95512            <reuse id="12877-de.xml" lang="de" version="4"/>
95513        </node>
```

```
195421       <node node_Type="graphic" id="G171096.ai" lang="none" version="1" docType="application/ai"
195422           dateFinal="2011-05-24T07:41:42Z" dateInitial="2011-05-24T07:41:42Z">
195423           <size>
195424               <none>1</none>
195425           </size>
195426       </node>
```

REx

Metrik und (statische) Kennzahlen

- www.home.hs-karlsruhe.de/~ziwo0001/PUB/CMS-Metrik_Ziegler.pdf

Metrische Untersuchung der Wiederverwendung

im Content Management

Statische Kennzahlen in der Technischen Redaktion

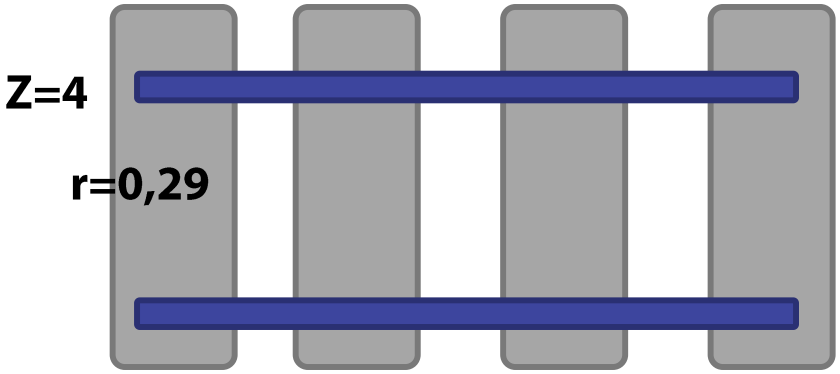
W. Ziegler[1] [Stand: 12. September 2008]

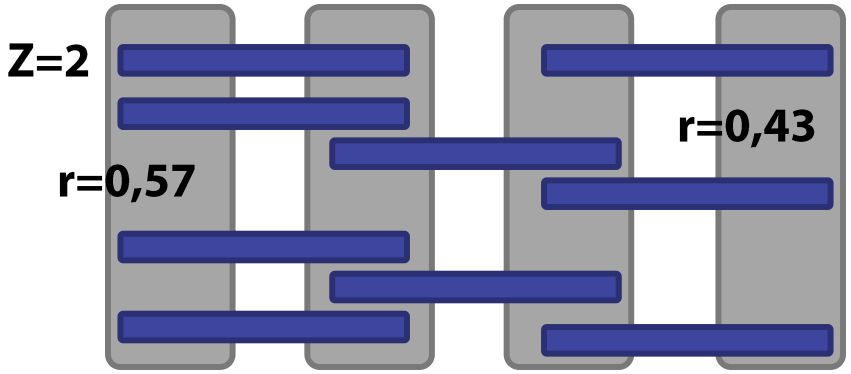
Die Einführung von Content Management (CM) Methoden und Systemen (CMS) in der Technischen Redaktion wird mit der effizienten Nutzung von wiederverwendbaren Inhalten verbunden. Allerdings fehlen bisher systematische Messmethoden, um den Wiederverwendungsgrad der Inhalte zu ermitteln und darzustellen. Der vorliegende Artikel führt entsprechende Messgrößen grundlegend ein

REx

Charakterisierende Größen

- **Redundanz**: Wiederverwendungsrate **r** von Dokumenten (oder Modulen)
- **Abundanz**: Wiederverwendungszahl **Z** von Modulen (oder Fragmenten/Medien)





Statische Systemkennzahlen (Auszug)

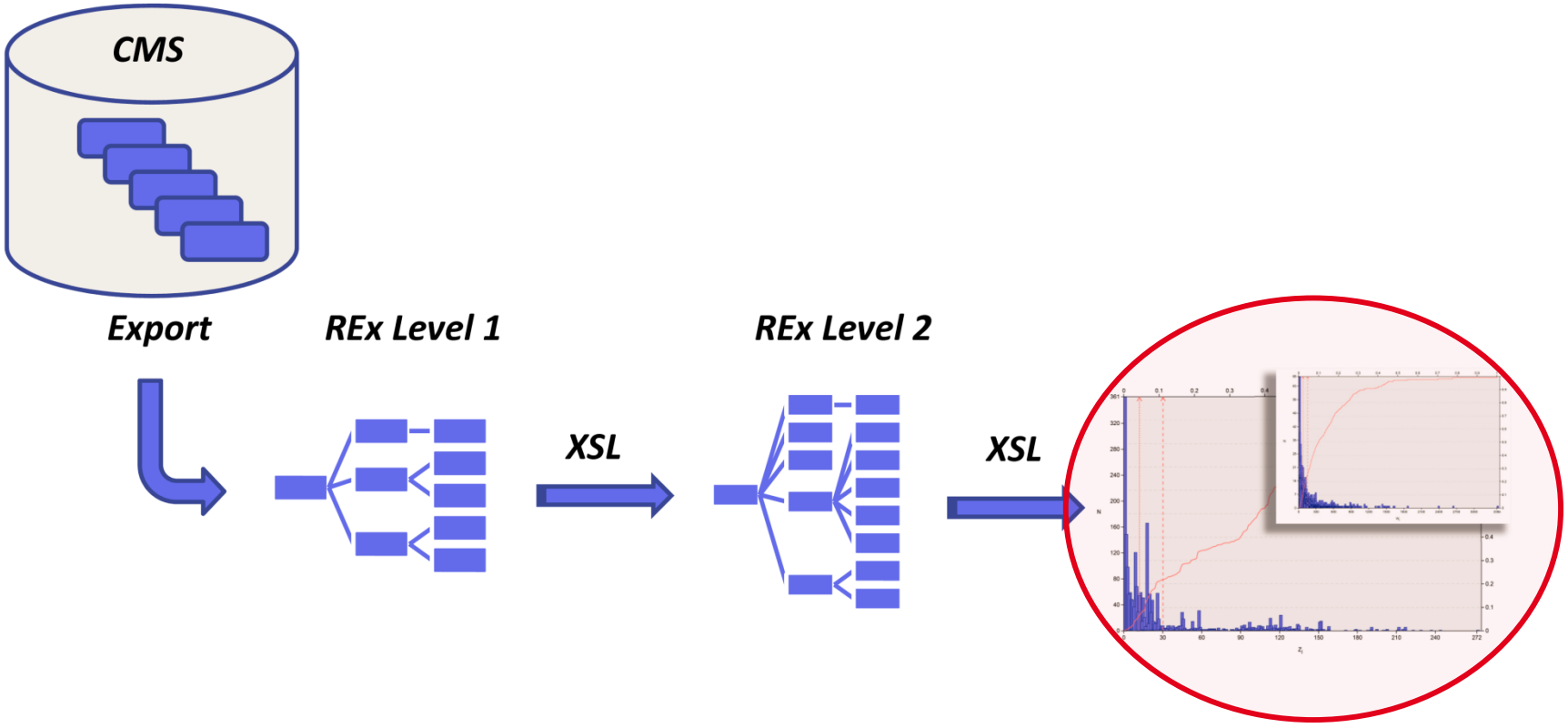
- Dokumentzahlen D
- Modulzahlen (genutzt/nicht genutzt) N
- Mittlere Dokumentgröße $\bar{N}_i$
- Modulgrößen W
- Mittlere Wv.rate (Mittlerer Redundanzgrad) $\bar{r}_c$
- Mittlere Wv.häufigkeit (Abundanz) $\bar{Z}$
- Versionierungen

- Statistische Verteilungen und Korrelationen für Analyse

Prototypische Daten und REx-Nutzung

REx

REx-Auswertung (Report)



Übersicht über Systemgrößen

Analyse | **Drucken**

Zusammenfassung		
Dokumente	D	492
Module	N	2275
Verwendete Module	N_{used}	1914
Mittelwert: Anzahl Module pro Dokument	$\bar{N}_i$	150.73
Mittelwert: Inhaltsbezogene Wiederverwendungsrate von Dokumenten	$\bar{r}^c_i$	0.9904
Mittelwert: Modulbezogene Wiederverwendungsrate von Dokumenten	$\bar{r}_i$	0.9927
Mittelwert: Segmente pro Modul	$\bar{W}_j$	155.26
Mittelwert: Gewichtete Verwendungszahl von Modulen	$\bar{Z}^c_j$	33.94
Mittelwert: Verwendungszahl von Modulen	$\bar{Z}_j$	30.19
Gesamtinhalt in Segmenten	$\bar{W}_j * N$	353000
Wiederverwendeter Inhalt in Segmenten	$Z^{>1} * W_j$	1.17E7
Wiederverwendeter Inhalt in Modulen	$Z^{>1}$	66771

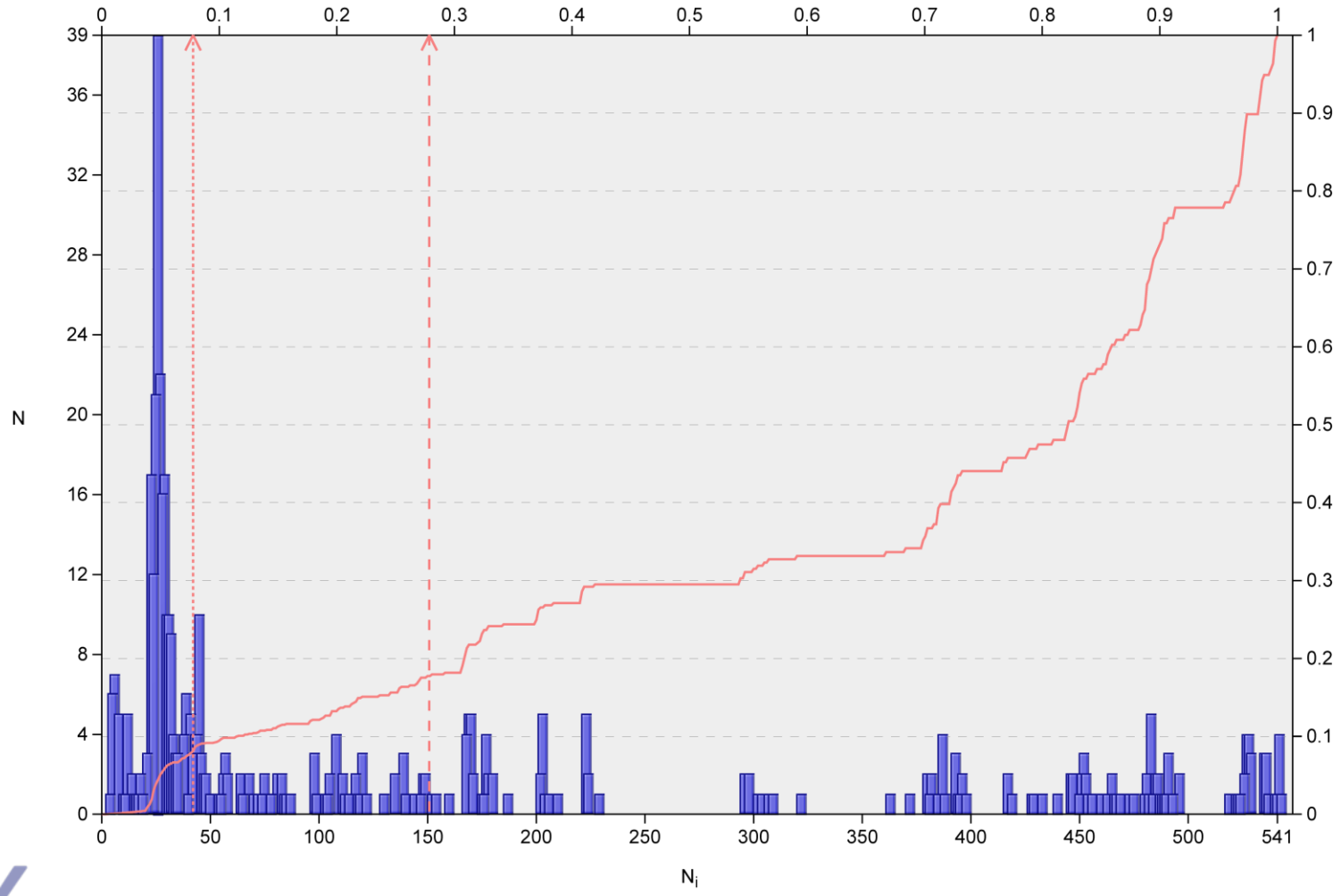
Angaben zum CMS	
Unternehmen	none
CMS	none
Zeitpunkt	none
Anwendungszweck	none

Angaben zur Analyse	
Typ der Analyse	Kanalauswertung
Sprachen	de
Dokumenttypen	all
Ausgabemedien	all
Nicht verwendete Objekte	enthalten
Ausgewertete Einheit für die Segmentgrößen	word

Angaben zu den Segmentgrößen	
Einheit "word"	gesetzt für
Dokumente	0%
Module	100%
Fragmente	-

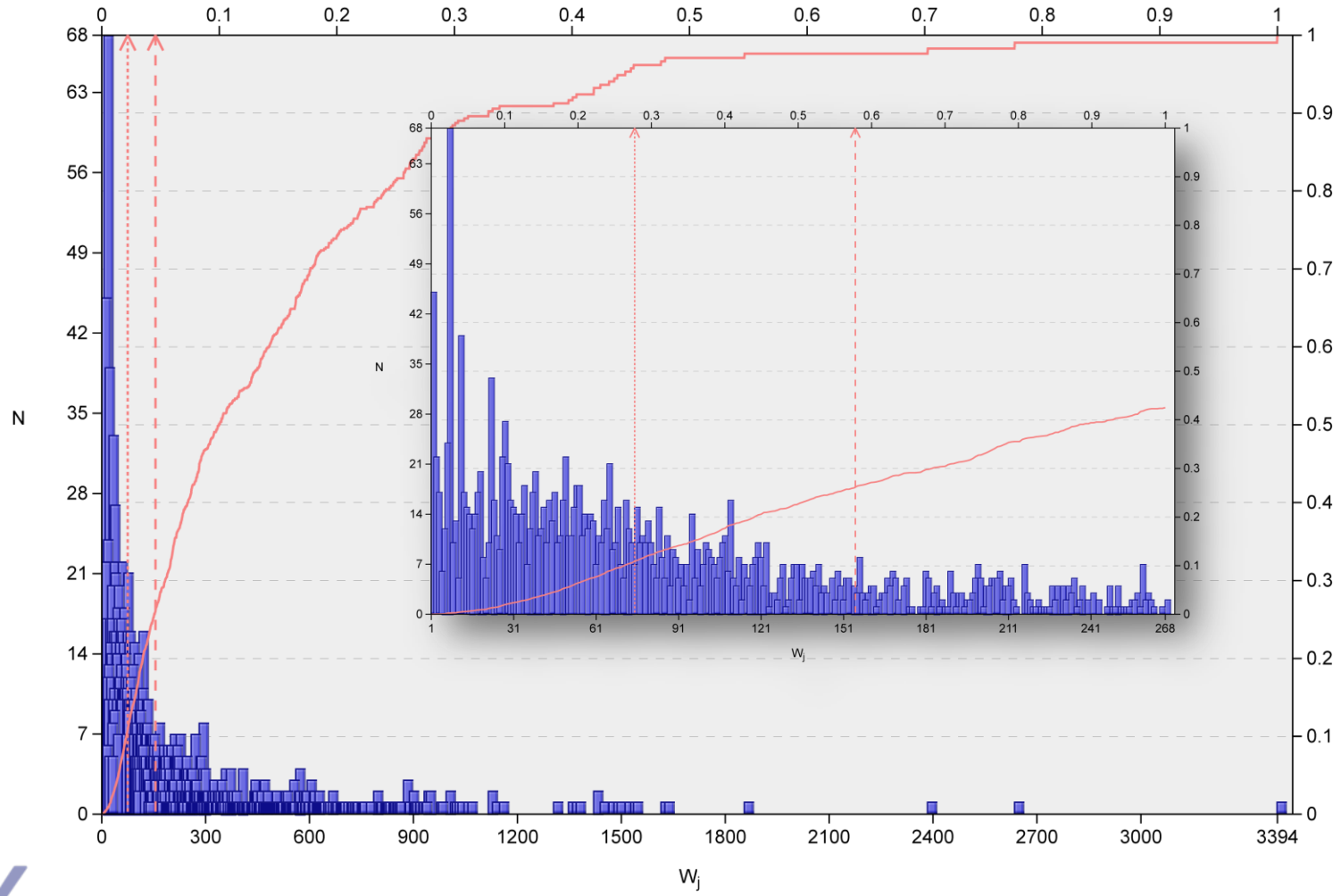
REx

Dokumentgrößen (Verteilung)



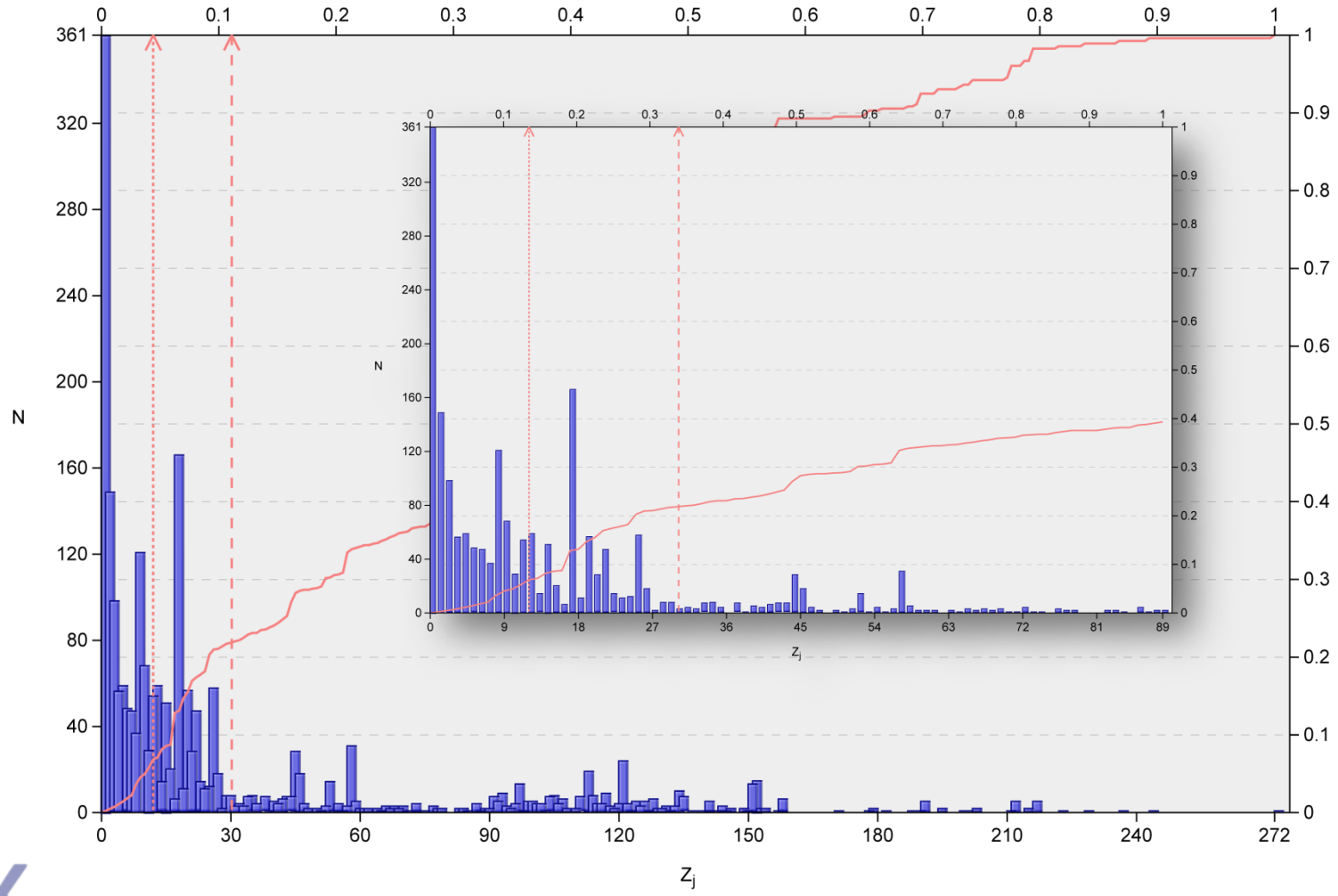
REx

Modulgrößen (Verteilung)



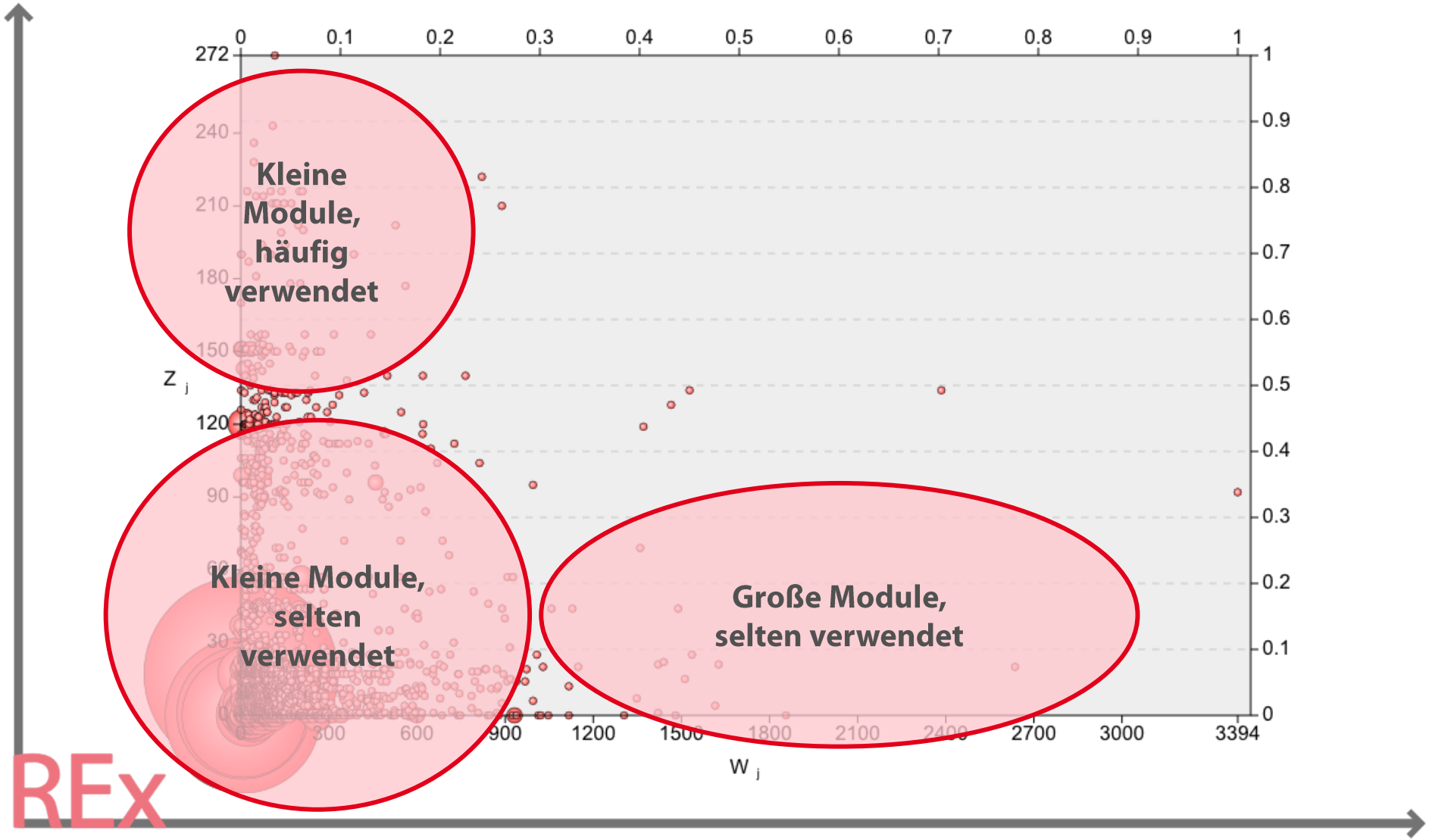
REx

Verwendungszahlen v. Mod. (Verteilung)



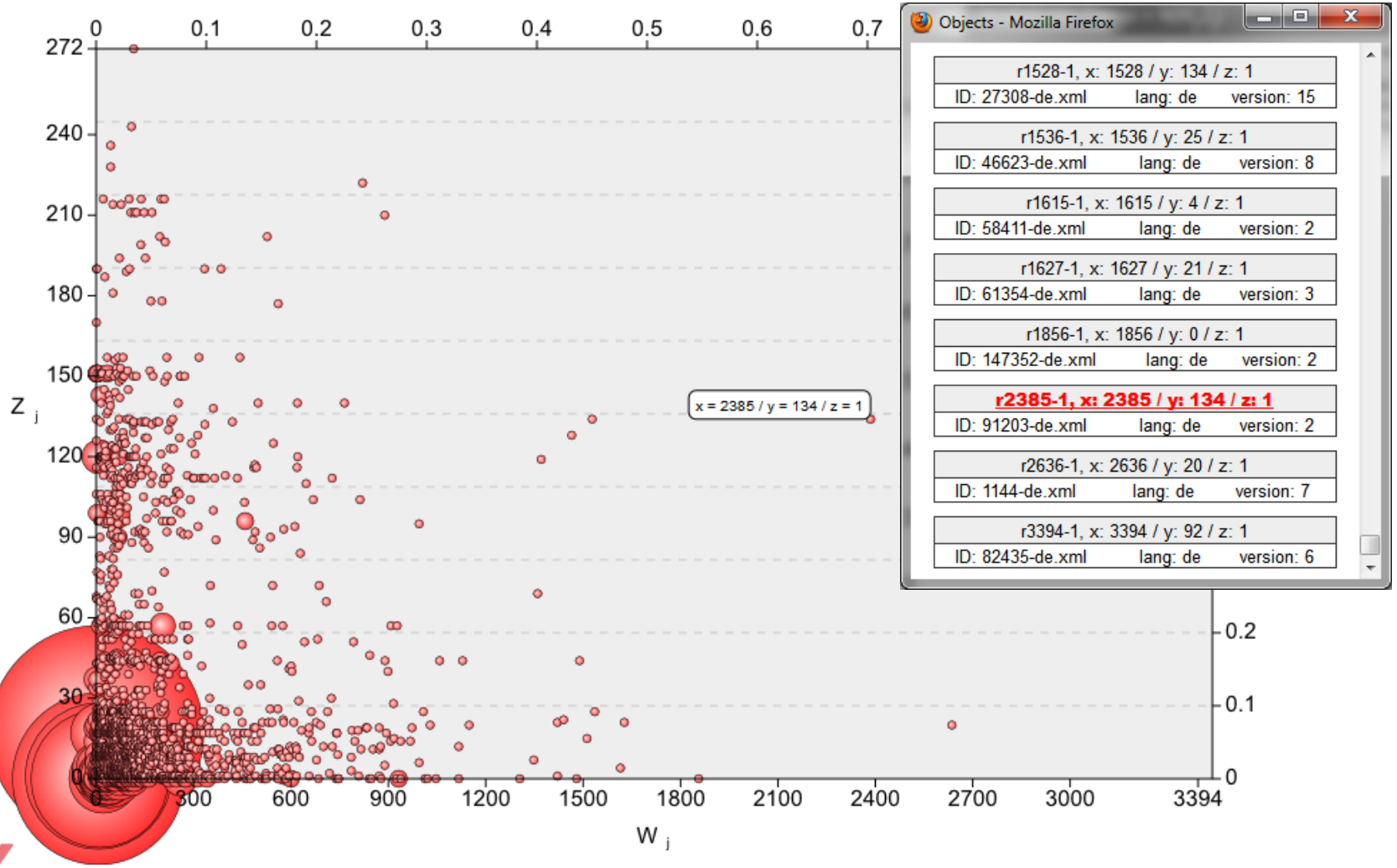
REx

Korrelationen (Größe W vs. Verwendungszahl Z)



REx

Korrelationen (Größe W vs. Verwendungszahl Z)



REx

Vorläufige Trends

- Viele kleine Module
- Hoher Redundanzgrad
- Geringe Abundanz (Reichweite)
- Software stark unterschiedlich zu Maschinenbau
- Clusterbildung analysierbar
- Modulare Nachverfolgung v. Interesse
- Modulleichen und Individualmodule auffindbar
- Metadatenanalyse notwendig

Kooperationspartner

Erstimplementierung durch CMS Anbieter + Kunden

- DOCUFY GmbH
- Expert Communication Systems GmbH
- Fischer Computertechnik GmbH
- gds AG – global document solutions
- Ovidius GmbH
- SCHEMA GmbH
- Siemens AG (HyBrix)

REx

Ausblick und Entwicklungen

REx

Grundlagenentwicklung

- REx 1.2
  - Metadaten vollständig exportierbar und analysierbar
  - Weitere Objekttypen (Gruppen, Master)
- Dynamische Kennzahlen
- (Projekt-)Verlaufskennzahlen
- Änderungshäufigkeiten und Versionierung
- Übersetzungsverläufe
- …
- Weitere CMS mit RExport-Schnittstelle

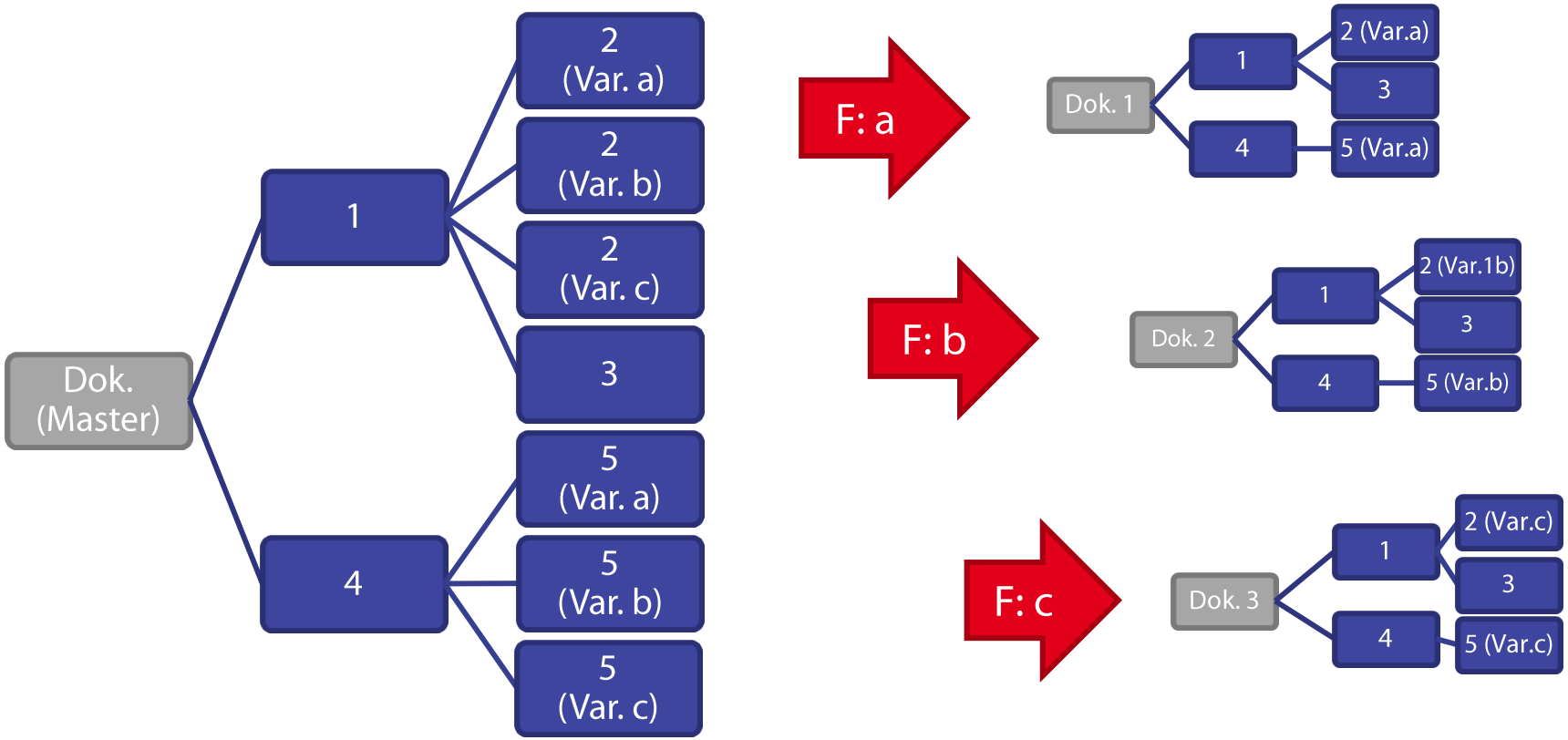
REx

Problem der Individualisierung

- Unterschiedliche Systemlogiken
- Anpassung der CMS in Kundenprojekten
- Unternehmensspezifische Prozessgestaltung und Systemnutzung
- Technische Auswertung und Interpretation des Variantenmanagements

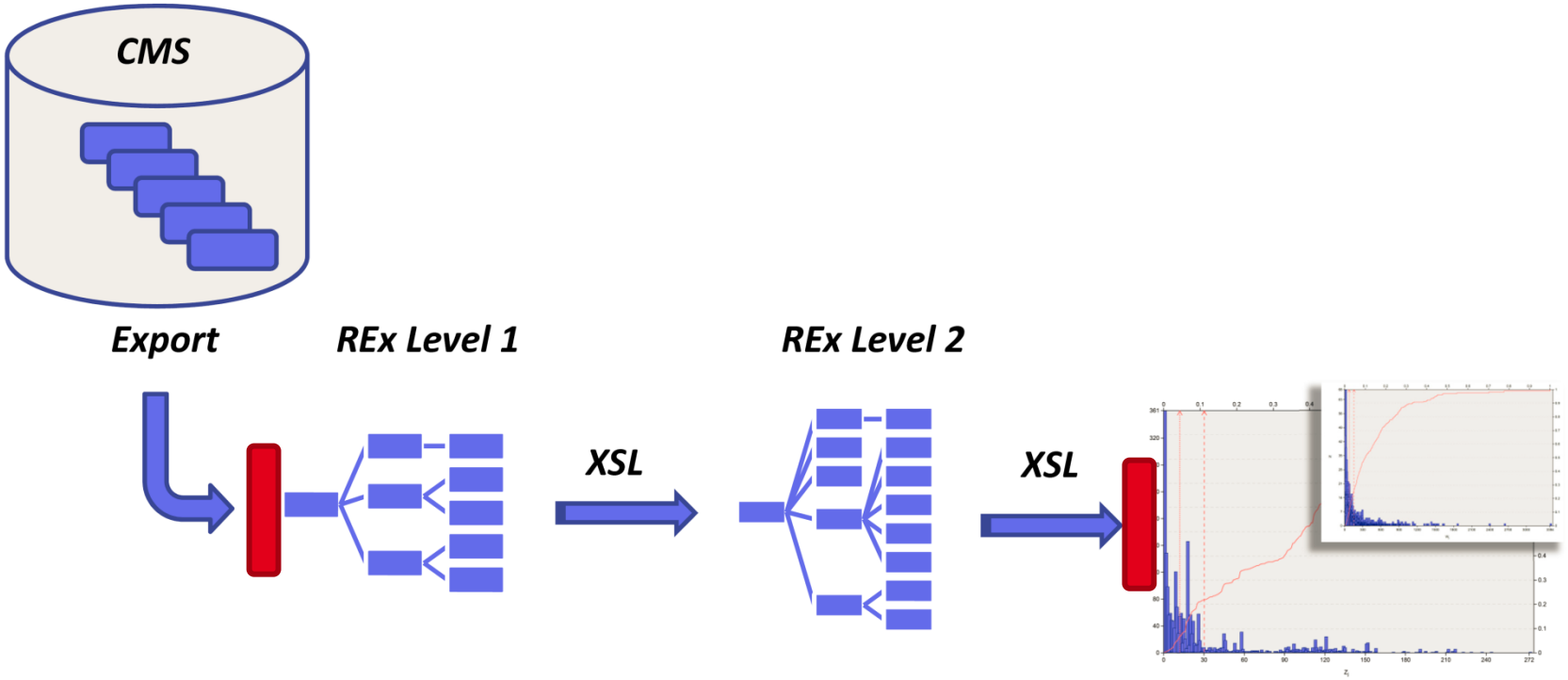
Variantenmanagement (Dok.ebene)

- Filterung von (Master-)Dokumenten



REx

REx und spezielle CMS-Adapter



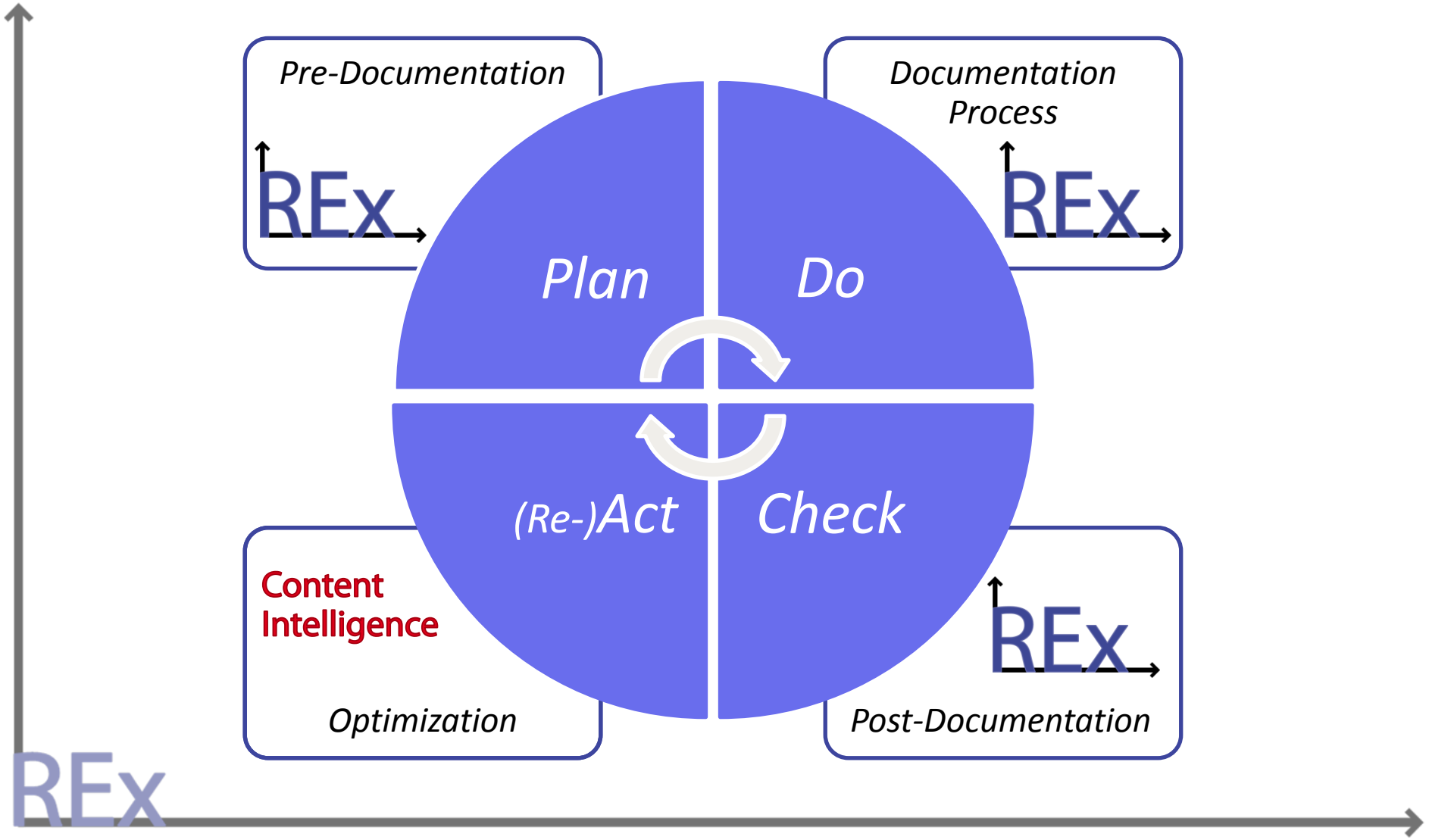
Technische Umsetzung

- Derzeit anonymisierter RExport extern/manuell ausgewertet und Kooperationspartnern zur Verfügung gestellt
- Zukünftige Möglichkeiten könnten sein
  - Web-Dienst/Upload-Server
  - (lizenzierte) CMS-Komponente
- Austauschformat für verschiedene DB-/CMS-Applikationen

REx

Zusammenfassung

Prozessmanagement



Definition von Content Intelligence

- **Methoden und Technologien zur Messung, Darstellung und Optimierung der CM-Prozesse**
- Methoden: statistisch und linguistisch
- Prozesse: Erstellung, Änderung, Verarbeitung und Nutzung von Content auf Basis der modularen Wiederverwendung

Ziele der REx-Methode (und v. CI)

- Monitoring und Optimierung der unternehmensinternen CM-Prozesse auf Basis einheitlicher Metrik
- Wissenschaftliche Studien zur Nutzung von CM-Implementierungen
- Unternehmens-/branchenübergreifende Benchmarks und Best Practices zur CMS-Nutzung und der CM-Methoden
- Einführung von „Content Intelligence“-Methoden
  - Monitoring und (REx-)Reports
  - Dashboards
  - Controlled Reuse
  - Reuse Memory (Autorenunterstützung)

REx

Literatur/Quellen

- P. Drewer, W. Ziegler „Technische Dokumentation“, Vogel Verlag (2011)
- C. Oberle, W. Ziegler „Content Intelligence für Redaktionssysteme“, technische kommunikation 6/12
- www.home.hs-karlsruhe.de/~ziwo0001/rex/intro.html

Kontakt für REx-Interessenten

- Prof. Dr. Wolfgang Ziegler
  Hochschule Karlsruhe
  Studiengang Kommunikation u. Medienmanagement
  Fak. Informationsmanagement u. Medien
  wolfgang.ziegler@hs-karlsruhe.de

  claudia.oberle@hs-karlsruhe.de
  (Diss.vorhaben Universität Mainz/HS Karlsruhe)

- Umsetzung:
  Institut für Informations- und Content-Management (I4ICM)
  Steinbeis GmbH an der HS Karlsruhe