

Ontologieentwicklung und -anwendung im dezentralen Wissensmanagement

Karlsruhe, 04.11.2005

Kolloquiumsvortrag

Christoph Tempich

Dezentrales Wissensmanagement

„Managing the processes of creating local knowledge within autonomous groups and exchanging knowledge across them is what we call distributed knowledge management (DKM)“
[Bonifacio et al. 2002]

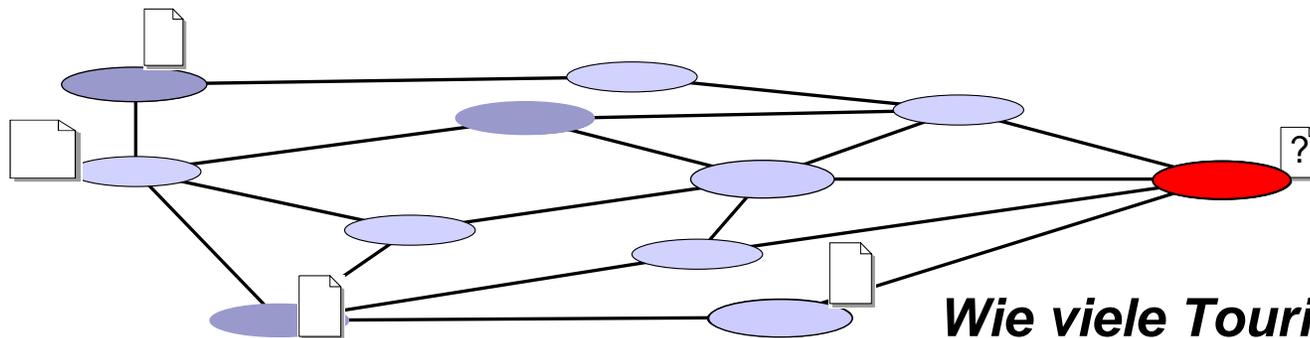


Tourismus Information austauschen

Berater

Hotel

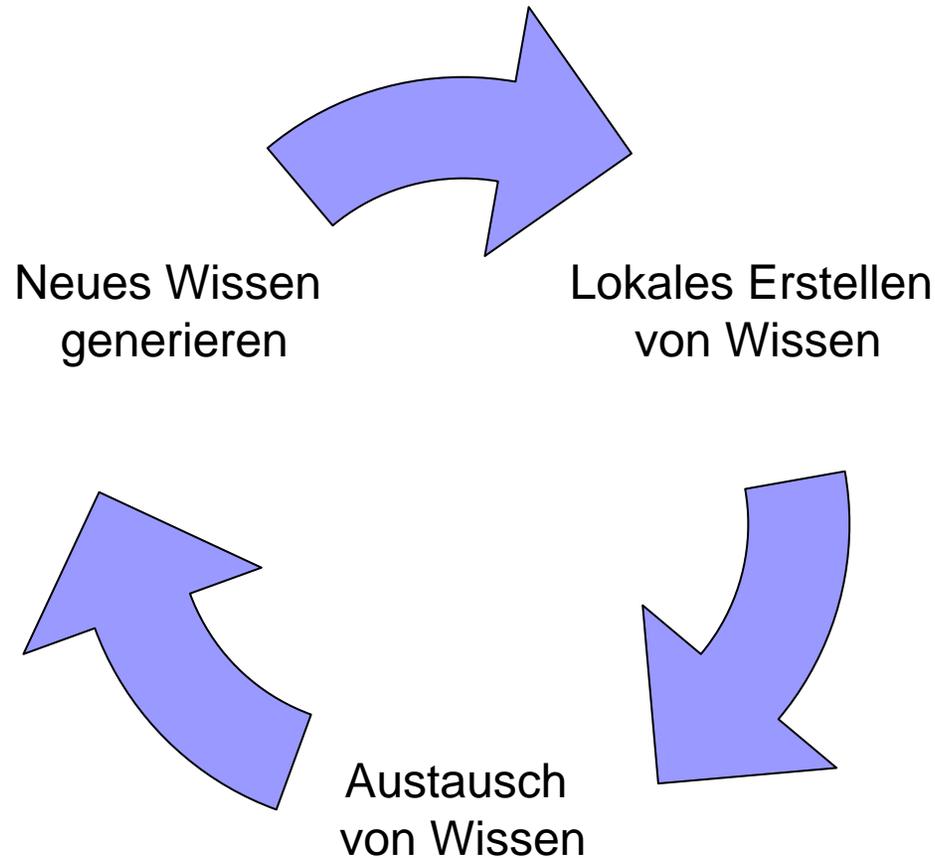
Öffentliche Einrichtung



**Wie viele Touristen
haben den Nord-Osten
von Mallorca im Jahr
2003 besucht?**

[PAKM 2004]

Prozess der Erstellung



Fokus meiner Arbeit

- Lokales Wissen erzeugen mit DILIGENT
 - Verständlich für den Menschen
 - Austauschbar mit Hilfe von Maschinen
- Lokales Wissen innerhalb von autonomen Gruppen austauschen mit REMINDIN
 - Kommunikation zwischen den Gruppen
 - Auffindbar innerhalb des Netzwerkes

Fokus meiner Arbeit

- Lokales Wissen erzeugen mit DILIGENT
 - Verständlich für den Menschen
 - Austauschbar mit Hilfe von Maschinen
- Lokales Wissen innerhalb von autonomen Gruppen austauschen mit REMINDIN
 - Kommunikation zwischen den Gruppen
 - Auffindbar innerhalb des Netzwerkes



■ DILIGENT Methodologie

Ontologie

An ontology is an explicit specification of a conceptualization.

[Gruber93]

An ontology is a shared understanding of some domain of interest.

[Uschold, Gruninger96]

■ Definiert

- Einen gemeinsamen Wortschatz
- Eine Beschreibung der Bedeutung von Begriffen
- Ein geteiltes Verständnis von Personen (und ggf. Maschinen)

Ontologieentwicklung

“the set of activities that concern the ontology development process, the ontology life cycle, and the methodologies, tools and languages for building ontologies.”

[Gomez-Perez 2004]

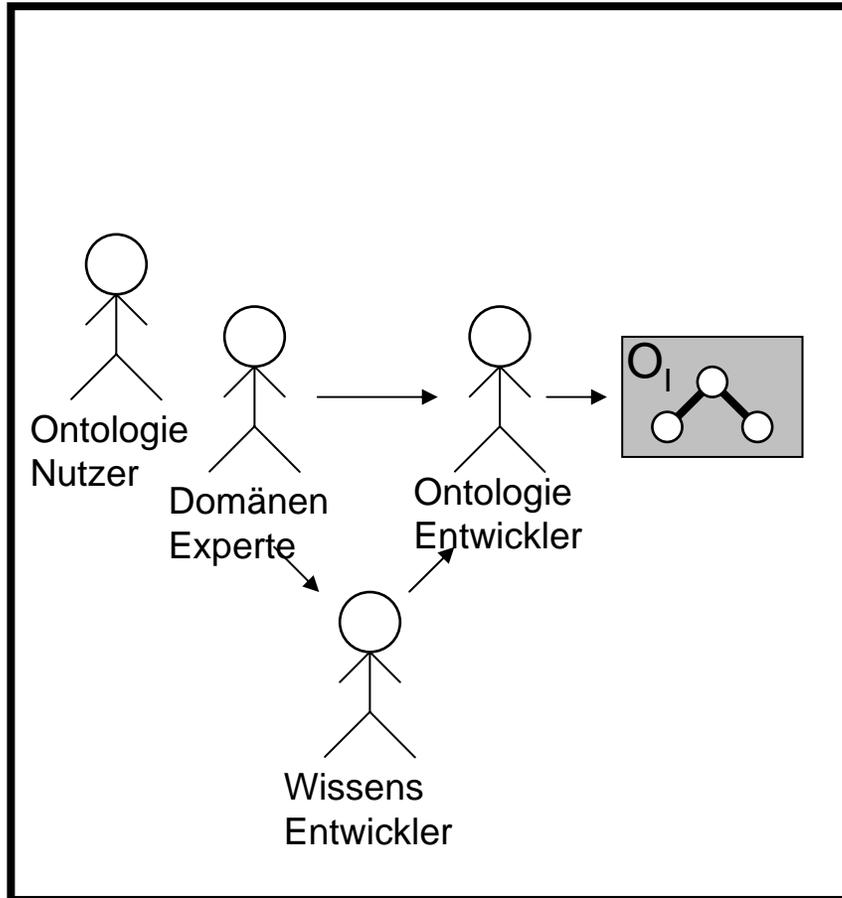
Verwandte Arbeiten

- METHONTOLOGY [Gomez-Perez97]
 - Ontologieerstellung für zentrale Anwendungen
- OTK Methodologie [Sure03]
 - Organisatorische Aspekte der Ontologieentwicklung für zentrale Anwendungen
- HCOME [Kotis05]
 - Ontologieerstellung in dezentralen Anwendungen, vages Prozessmodel

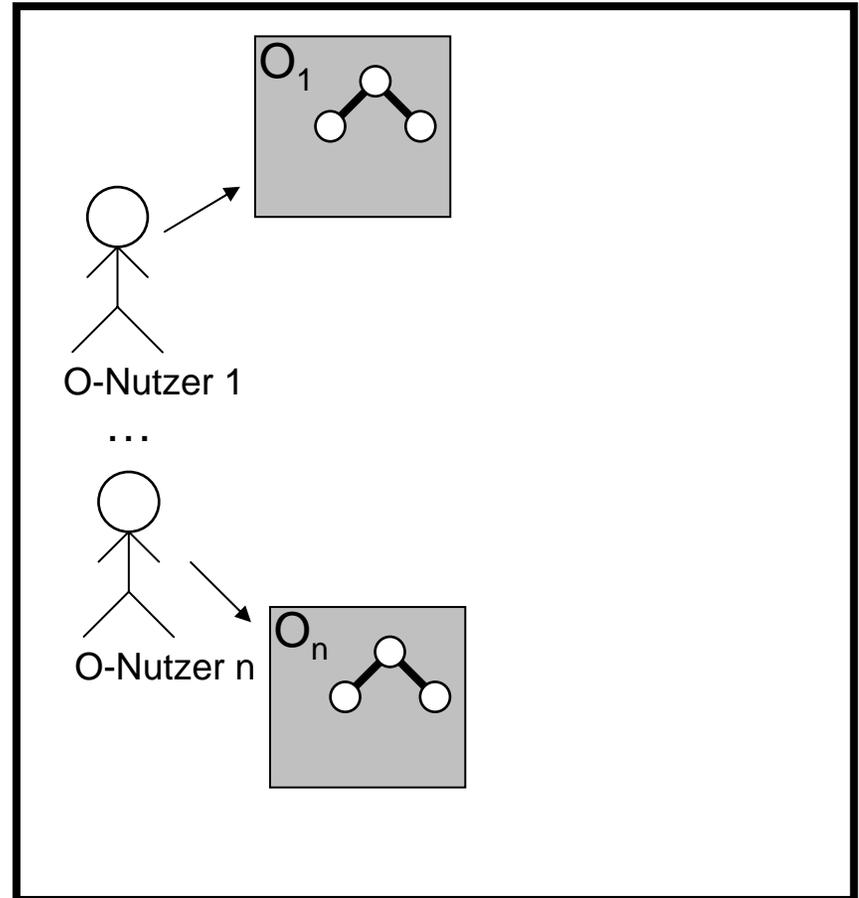
Was ist DILIGENT?

- Ist eine Ontologieentwicklungsmethodologie
- Für die verteilte Erstellung von Ontologien
- Innerhalb von autonomen Gruppen
- Für veränderbares Wissen

DILIGENT Prozess: Rollen

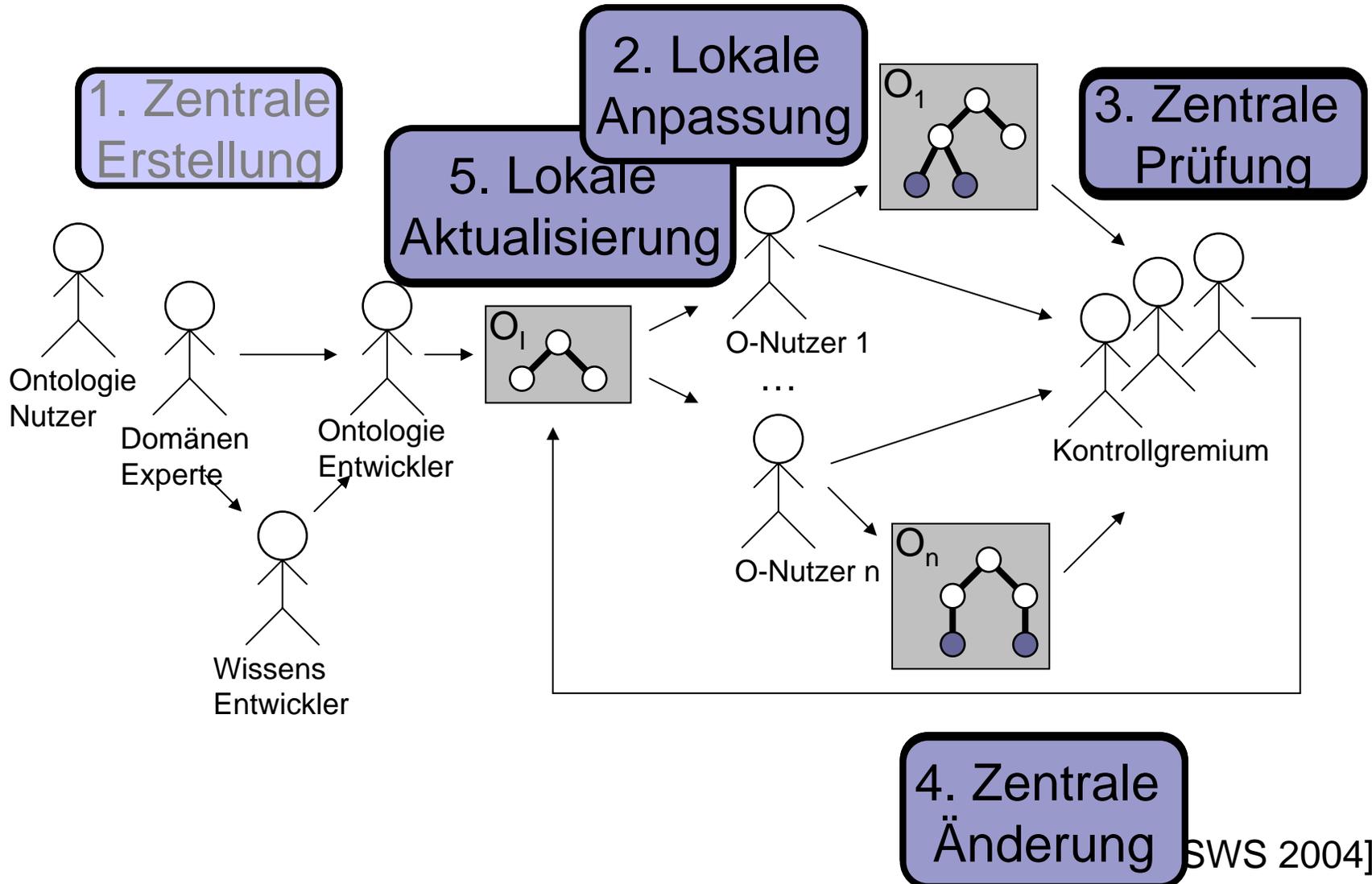


Kontrollgremium

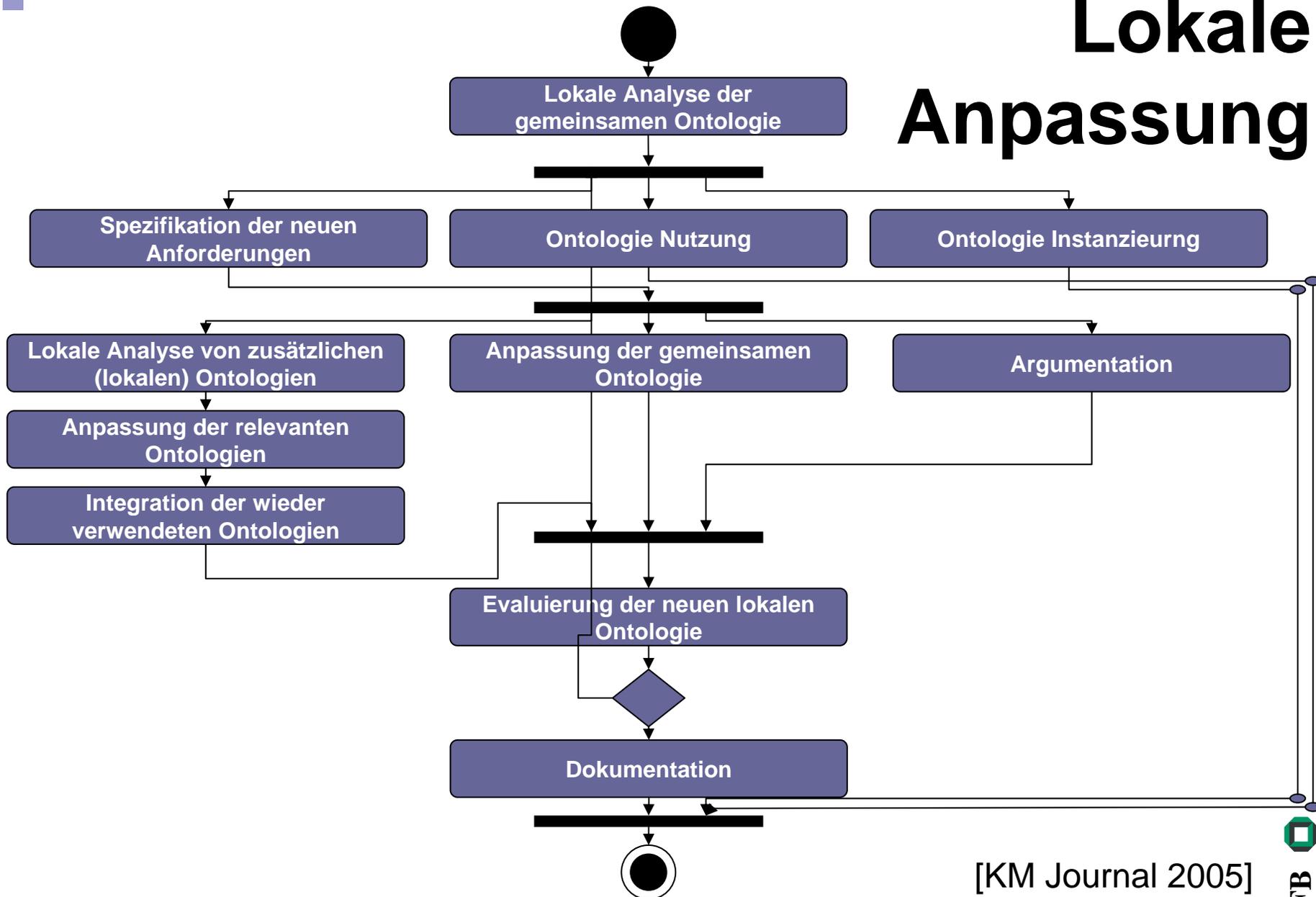


Nutzer

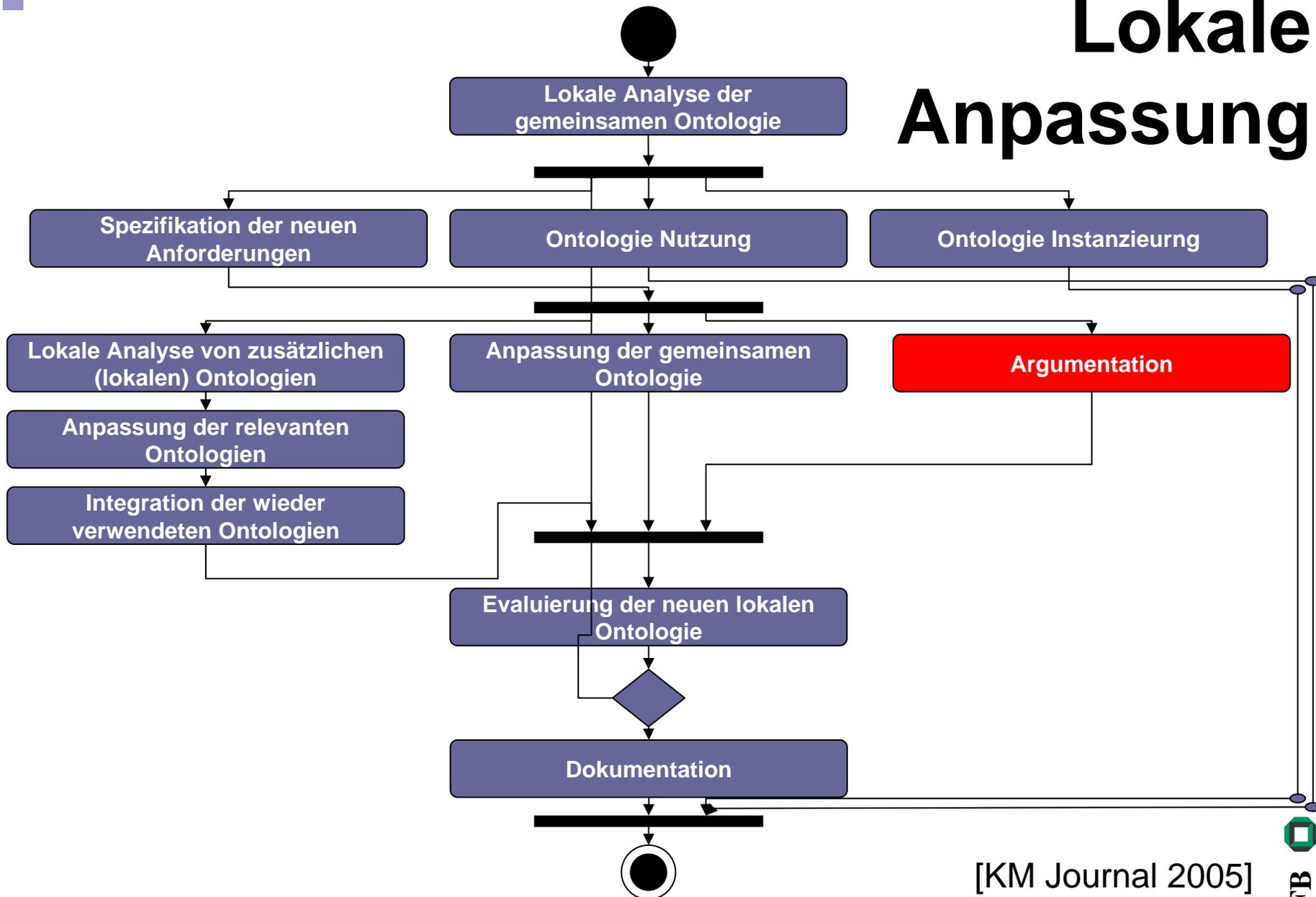
DILIGENT Prozess: Phasen



Lokale Anpassung



Lokale Anpassung





- Argumentation während der Ontologieentwicklung

Konsensbildung problematisch

- Unterschiedliche Interessen führen zu gegensätzlichen Konzeptualisierungswünschen
- Konzeptualisierungsdiskussion ist unstrukturiert
 - Argumentation mit Hilfe von eMails oder mündlich
 - Keine Richtlinien vorhanden
- Einigungsprozess nicht nachvollziehbar

Beispiel: Modellierung Geographischer Informationen

- Fragestellung:
Ist die Beziehung „istTeilVon“ transitiv?
- Beispiel:
 - Mallorca istTeilVon Spanien
Palma di Mallorca istTeilVon Mallorca
 - Andalusien istTeilVon Spanien
Gibraltar istTeilVon Andalusien
- Antwort:
In diesem Fall abhängig davon, ob die Beziehung politisch oder geographisch definiert ist



■ Argumentationsontologie

10 Anforderungen an die Argumentationsunterstützung

1. Benutzung einer etablierten Terminologie
(IBIS; Compendium, Rethorical Structure Theory)
2. Ontologiebezug
(Konzeptuelles Level vs. Formales Level)
3. Relevante Argumente
(z.B. Beispiel, Evaluierung; Alternative, Gegenbeispiel)
4. Unterstützung des Gesamtprozesses
(Entscheidungen; Positionen; ...)

10. ...

Argumentationsontologie

- Terminologie übernommen von IBIS
 - Issue, Idea, Argument

DILIGENT

Argumentationsontologie

- Terminologie übernommen von IBIS
 - Issue, Idea, Argument

IBIS: Issue Based Information Systems

Standard Methode zur Beschreibung von beliebigen Argumenten

Argumentationsontologie

- Terminologie übernommen von IBIS
 - Issue, Idea, Argument
- Separate Diskussion von konzeptionellem Modell und Umsetzung
 - Issue: Konzeptionelle Anforderung
 - Idea: Umsetzung

Argumentationsontologie

- Terminologie übernommen von IBIS
 - Issue, Idea, Argument
- Separate Diskussion von Konzeptionellem Modell und Umsetzung
 - Issue: Konzeptionelle Anforderung
 - Idea: Umsetzung
- Effektive Auswahl von Argumenten
 - Analyse anhand der Rethorical Structure Theory
 - Example, Counter Example, Alternative, Evaluation

Argumentationsontologie

- Terminologie übernommen von IBIS

- Issue, Idea, Argument

- Separate Diskussion
Umsetzung

- Issue: Konzeptionelle
- Idea: Umsetzung

- Effektive Auswahl von

- Analyse anhand der
- Example, Counter Ex

RTS: Rethorical Structure Theory

Standard Methode zur Beschreibung
von Argumenten innerhalb eines
Textes

25 verschiedene
Argumentationstypen

Argumentationsontologie

- Terminologie übernommen von IBIS
 - Issue, Idea, Argument
- Separate Diskussion von Konzeptionellem Modell und Umsetzung
 - Issue: Konzeptionelle Anforderung
 - Idea: Umsetzung
- Effektive Auswahl von Argumenten
 - Analyse anhand der Rethorical Structure Theory
 - Example, Counter Example, Alternative, Evaluation
- Entscheidungsunterstützende Konzepte
 - Z.B. Agree, Disagree, Decision



■ Evaluierung

Evaluierung

- Prozessmodell
 - Fallstudie auf Mallorca
 - Anwendung im SEKT Projekt
- Argumentation
 - Experimente am Institut
 - Anwendung bei der Erstellung einer Rechtsontologie

Anwendungen

SWAP: P2P System XAROP

The screenshot shows the XAROP P2P system interface. On the left, there is a 'Peer Folders' pane showing a local peer named 'estevé' with folders for 'Competitor Destinations', 'Incoming Markets', 'Passengers', and 'Tourist Profile'. Below it is an 'Ontologies' pane with a search box containing 'accommodation' and a 'Concept Search' tree showing a hierarchy of concepts like 'New Technologies and Innovation', 'Quality and Hospitality Management', etc. The main area is a 'Cluster Map' showing a network of nodes representing different concepts, such as 'estevé: Tourist Profile (7)', 'estevé: Competitor Destinations (13)', 'estevé: "accommodation" (10)', 'estevé: Indicator (3)', 'estevé: Passengers (14)', 'estevé: Incoming Markets (14)', and 'estevé: "mallorca" (10)'. At the bottom, a 'Details' pane shows a table of peers and a tree view of the selected peer's structure, including folders like 'start2_data', 'AcroMed', 'Literatur', 's12', 'a_files', and 'WordNet'. A context menu is open over the 'communication' folder, showing options like 'Insert Concept', 'Remove Concept', 'Insert Relation', 'Edit concept', 'Sort Alphabetically', 'Sort Frequency', 'Sort Adaption Rate', 'ShowSwabbi', 'Generate Query', and 'select to align with'.

Sekt: Digital Library

The screenshot shows the Sekt Digital Library website. The browser address bar displays 'http://www.w3.org/TR/rdf-primer/'. The page content includes a search result for 'Automation (14992)' with a description: 'Describes how a bank can use technology to automate its business'. The page also shows a navigation menu with items like 'home', 'events', 'news', 'project', 'partners', 'resources', 'r&d', 'internal pages', and 'my folder'.

Sekt: Rechtsfallstudie

The screenshot shows the Sekt website, which is a 'Semantically-Enabled Knowledge Technologies' platform. The page is titled 'ActoDeTercero' and is part of a 'legal case study' series. The page content includes a navigation menu, a 'Wiki contents' section with a tree view of the site structure, and a main text area with 'Pro: Justification' and 'Contra: Justification' sections. The 'Pro: Justification' section states: 'Se trata de una ca... Normalmente se referirá a los testim... proceso o la función judicial que no...'. The 'Contra: Justification' section states: 'No se trata de... que los actos de tercería y los actos... encontrarse una forma más clara d...'. The page also features a search bar and a 'you are here' breadcrumb trail: 'home > internal pages > sektwikis > diligent > legal case study > rootconcept'.

04.11.2005

DILIGENT: Mein Beitrag

- Definition des DILIGENT Prozessmodells
 - Unterstützt Erstellung von Ontologien
 - Für verteilte, dynamische und offene Systeme
- DILIGENT Argumentationsmodell
 - Anpassung für die Ontologieerstellung
 - Argumente sind nachvollziehbar
 - Diskussion ist strukturiert und effizient
 - Einigungsprozess ist reproduzierbar
- Evaluierung in Fallstudien

Fokus meiner Arbeit

- Lokales Wissen erzeugen mit DILIGENT
 - Verständlich für den Menschen
 - Austauschbar mit Hilfe von Maschinen
- Lokales Wissen innerhalb von autonomen Gruppen austauschen mit REMINDIN
 - Kommunikation zwischen den Gruppen
 - Auffindbar innerhalb des Netzwerkes

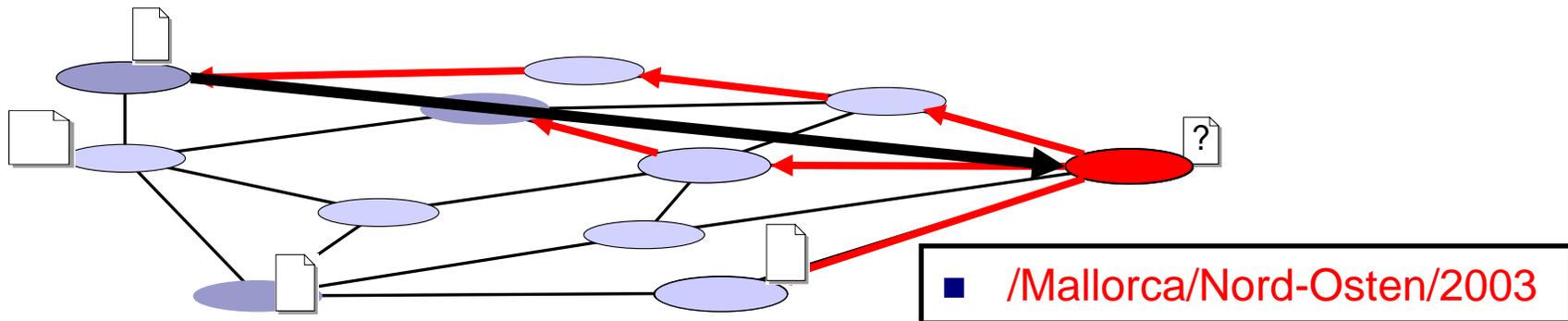


■ Routingalgorithmus



Problemstellung

- Jeder Peer ist eindeutig identifizierbar
- Peers tauschen Nachrichten aus
- Peers sind autonom
- Volatilität der Peers
- Information durch Ontologie beschrieben

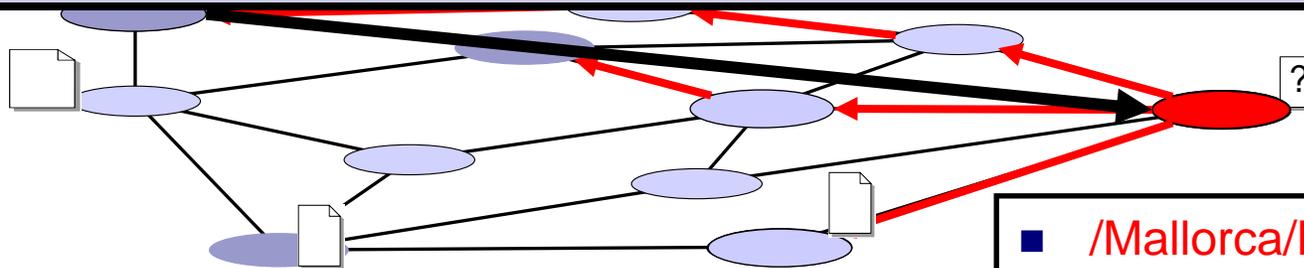




Problemstellung

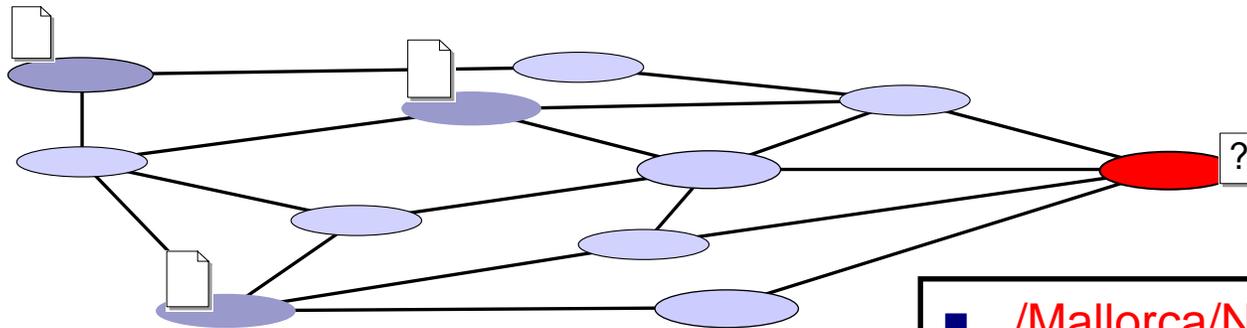
- Jeder Peer ist eindeutig identifizierbar
- Peers tauschen Nachrichten aus
- Peers sind autonom

Wie finde ich effizient die richtigen Knoten im Netzwerk ?





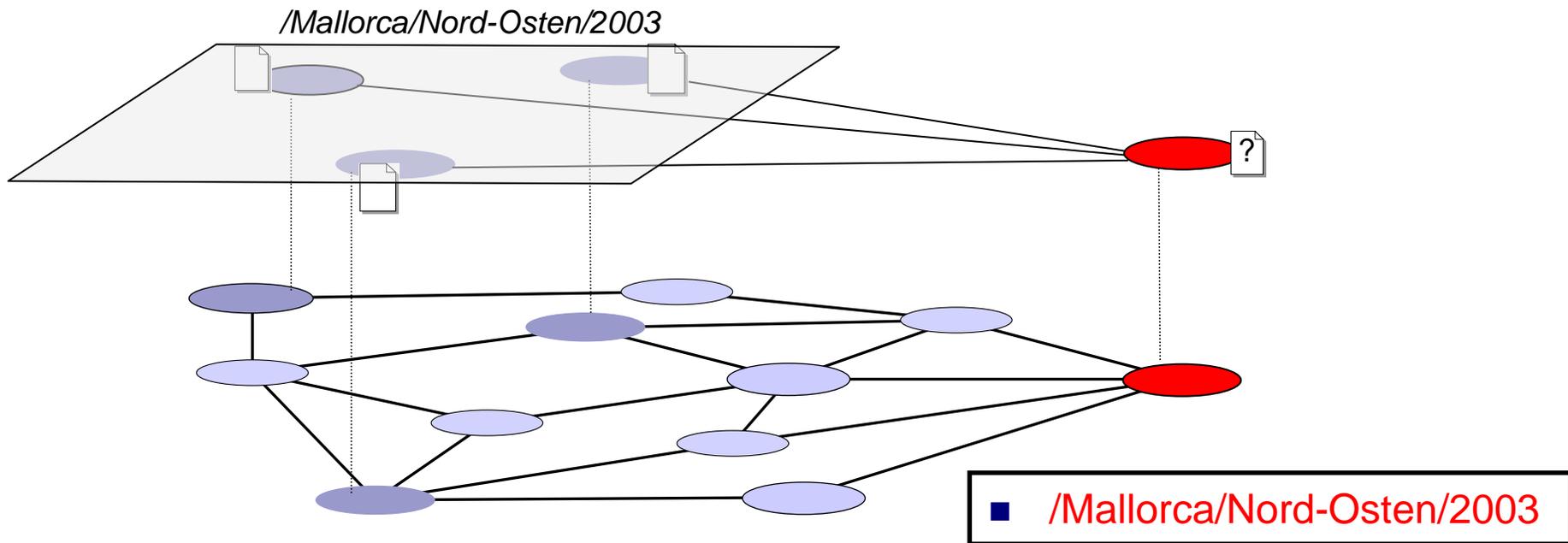
Logische Overlay Struktur



■ /Mallorca/Nord-Osten/2003



Logische Overlay Struktur



Klassifikation von Overlay Netzwerken

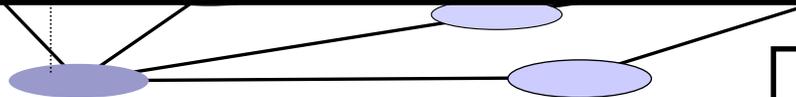
	Grad der Zentralisierung		
	Ohne	Super Peer	Hybrid
Unstrukturiert	Gnutella v0.4 FreeHeaven	Gnutella v0.6 Kazaa	Napster SnowCap
Semi-strukturiert	ForeSeer, ... Remindin'	Gia	
Strukturiert (Peers)	Structella	Edutella	
Strukturiert (Peers und Dokumente)	CHORD CAN PASTRY P-Grid	TOPICS	

Klassifikation von Overlay Netzwerken

	Grad der Zentralisierung		
	Ohne	Super Peer	Hybrid
Unstrukturiert	Gnutella v0.4 FreeHeaven	Gnutella v0.6 Kazaa	Napster SnowCap
Semi-strukturiert	ForeSeer, ... Remindin'	Gia	
Strukturiert (Peers)	Stru		
Strukturiert (Peers und Dokumente)	CH CA PA P-C	<div style="border: 2px solid red; padding: 10px; text-align: center;"> <p>.... and Sripranidkulchai, ...</p> </div>	

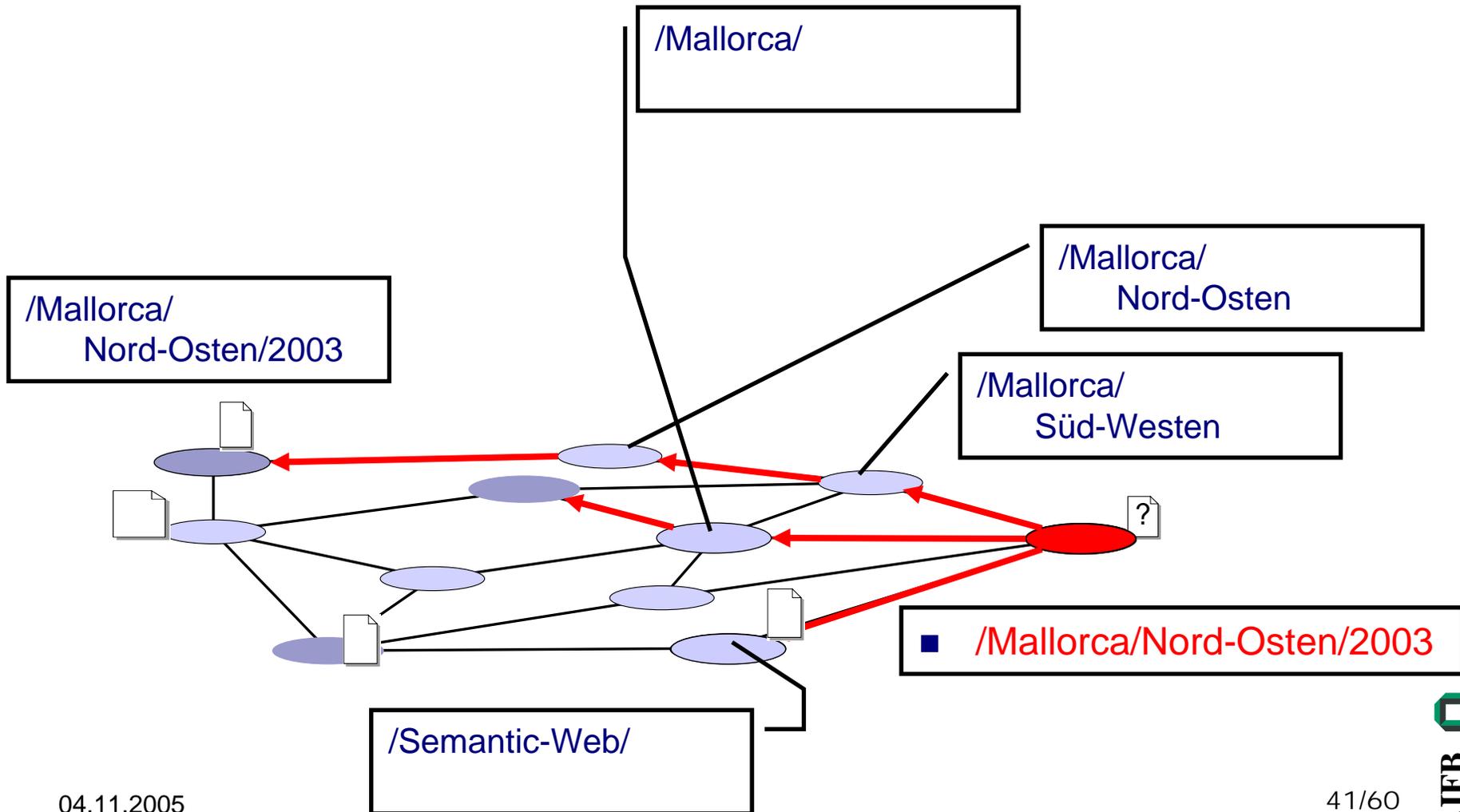
Logische Overlay Struktur

- 1. Wie können im Vergleich zu anderen Routing Verfahren die Anzahl der versendeten Nachrichten reduziert oder der Recall erhöht werden?**
- 2. Wie kann ein solches Overlay Netzwerk effektiv organisiert werden?**

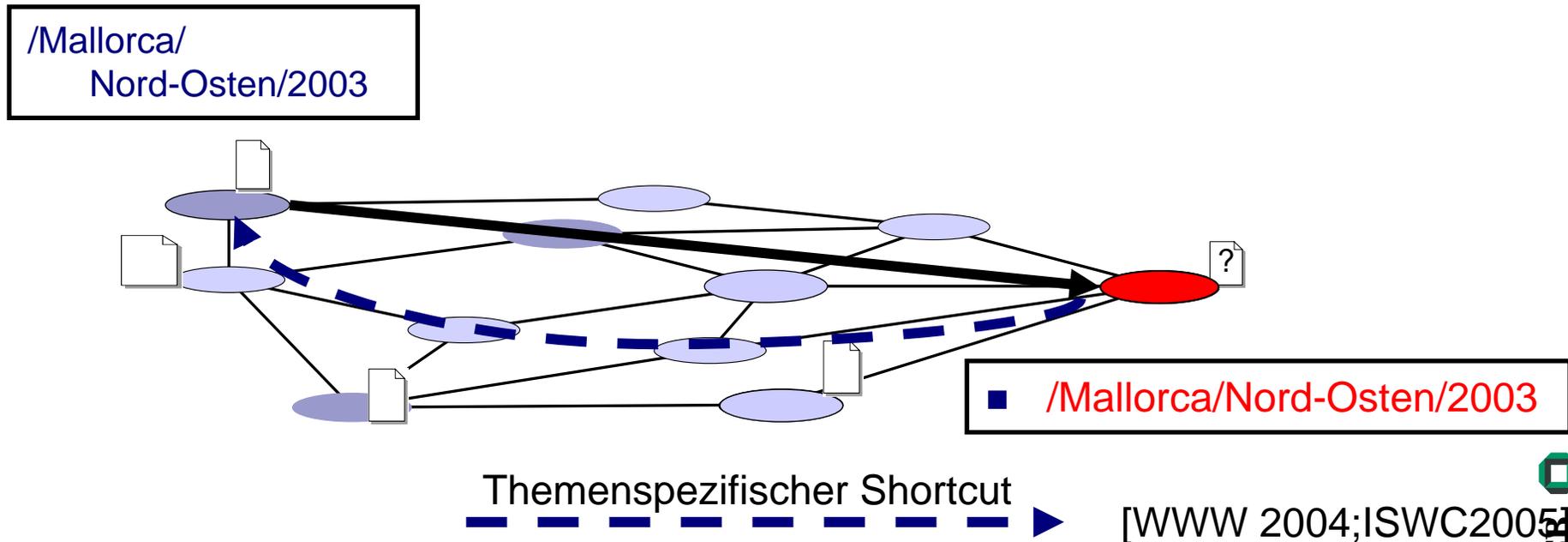


■ /Mallorca/Nord-Osten/2003

Auswahl basierend auf thematischer Nachbarschaft



Auswahl basierend auf thematischer Nachbarschaft



Peerauswahl mit Remindin'

■ Indexaufbau

Informationen über fremde Peers basieren nur auf deren Antworten und Anfragen

■ Indexpflege

Der Shortcut Index ist begrenzt, irrelevante Shortcuts werden gelöscht

■ Shortcutauswahl & Anfrage Weiterleitung

Lokale Auswahl von Peers basierend auf dem lokalen Index und
Anfrageweiterleitung

■ Evaluation

Vergleich mit anderen Algorithmen die den Stand der Forschung widerspiegeln



■ Indexaufbau

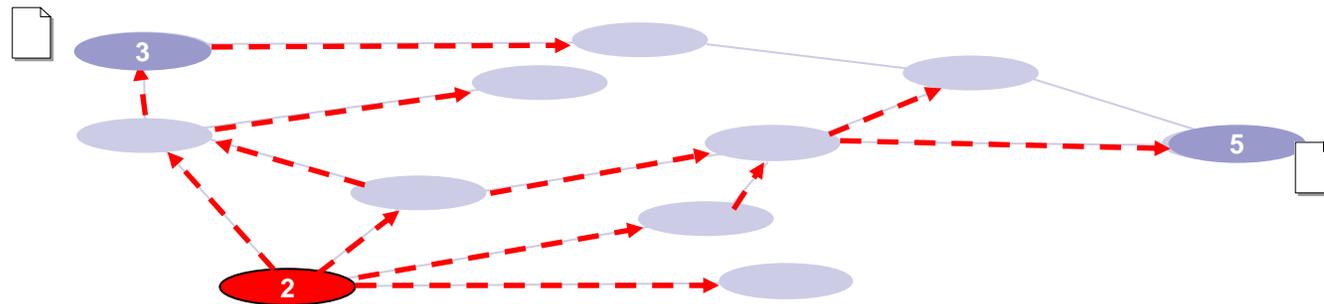
Shortcutterstellung basierend auf sozialen Metaphern

Man fragt eine Person

- Wenn sie Fragen in der Vergangenheit erfolgreich **beantwortet hat**
- Wenn sie diese oder ähnliche Fragen in der Vergangenheit **gestellt hat**
- Wenn sie gut mit anderen Personen **vernetzt** ist

Ansonsten nach dem Zufallsprinzip

Beispiel: Shortcut Erstellung zu Antwortenden

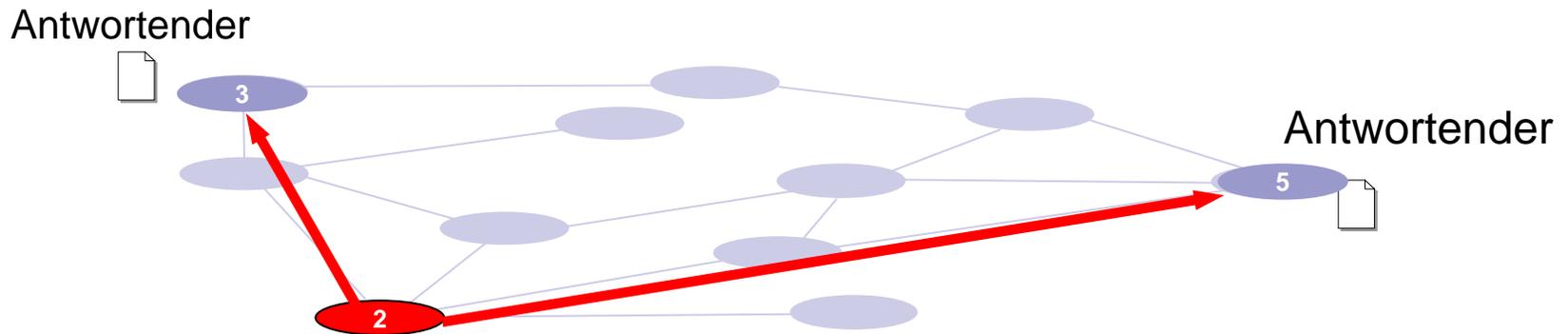


■ /Mallorca/Nord-Osten/2003

Flooding (radius = 3 hops)



Beispiel: Shortcutterstellung zu Antwortenden



■ [/Mallorca/Nord-Osten/2003](#)

Erstelle Antwort Shortcut

<i>Shortcuts Peer 2</i>				
Query Q'	PID	Query hits	SC Type	Last Time Used
/Mallorca/Nord-Osten/2003	5	100	Antwort	12:54:30
/Mallorca/Nord-Osten/2003	3	93	Antwort	12:54:30

Weitere Shortcut Typen

- Antwortende
 - Liefern Antworten auf eine bestimmte Anfrage
- Empfehlung
 - Schicken mir eine Anfrage
 - Schicken meine Anfrage zu einem Antwortenden
- Bootstrapper
 - Kennen viele Peers und sind bekannt



■ Indexpflege

Shortcut Index Management

Lösche Shortcuts mit niedrigem Rang

- *Gewichteter Durchschnitt*
 - Ähnlichkeit zu eigenem Inhalt
 - Benutzung (LRU)
 - Shortcut Typ

- Gewichte sind von Applikation abhängig

Shortcuts Peer 2 (lokal: /Mallorca/Nord-Osten/; Aktuelle Zeit: 14:00)

Query Q'	PID	Query hits	SC Type	Last Time Used	Ähnlichkeit	Rang
/Mallorca/Nord-Osten/	5	100	Antwort	12:54:30	1,0	1
/Mallorca/Süd-Westen	3	1	Empfehlung	12:54:30	0,5	2

löschen →



- Indexnutzung
(Anfrage Weiterleitung)

Peerauswahl basiert auf lokaler Rangliste

1. **Ähnlichkeitsbasiert:**
Rang abhängig von der Ähnlichkeit des Shortcuts und der Anfrage
2. **Gradbasiert:**
Rang der Bootstrapping Shortcuts abhängig von Ein-/Ausgangsgrad des fremden Peers
3. **Netzwerkbasieret:**
Auswahl von Peers aus der Netzwerkschicht
4. **Zufallsbasiert:**
Austausch von ausgewählten Peers mit zufällig ausgewählten



■ Evaluierung

Rundenbasierte Simulation

Datensatz: Open Directory Projekt

- Verteilung des Inhalts:
 - Ein Editor = Ein Peer
 - 1024 Peers, 1646 Topics
 - Zipf verteilte Topics
 - Wenige oft vorkommende Themen
- Anfrage Verteilung:
 - Schlechtester Fall: Gleichverteilte Anfragepopularität
 - 30000 Anfragen, Jeder Peer stellt im Durchschnitt 30 Anfragen
- Netzwerk Topologie:
 - Power Law
 - $k=2$, hops $h=6$
- Volatilität
 - Nach 15 Anfragen stellt jeder Peer Anfragen zu neuen Themen
 - Hier: Peers immer Online
 - On- und Offline Zeiten nach Gnutella

[S. Saroiu, et.al. 2003]

Rundenbasierte Simulation

Datensatz: Open Directory Projekt

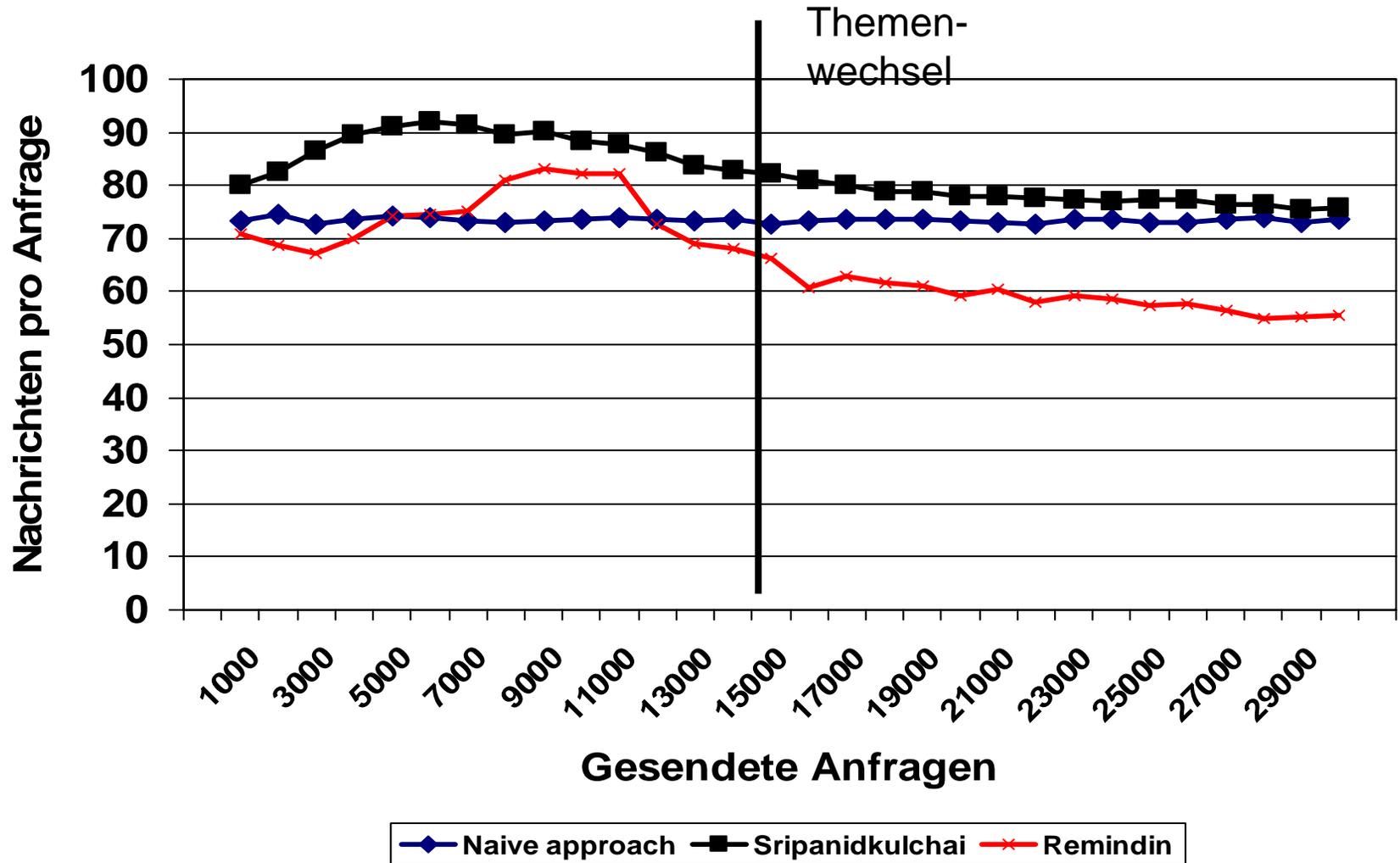
- Verteilung des Inhalts:

- Ein Editor = Ein Peer
- 1024 Peers, 1646 Topics

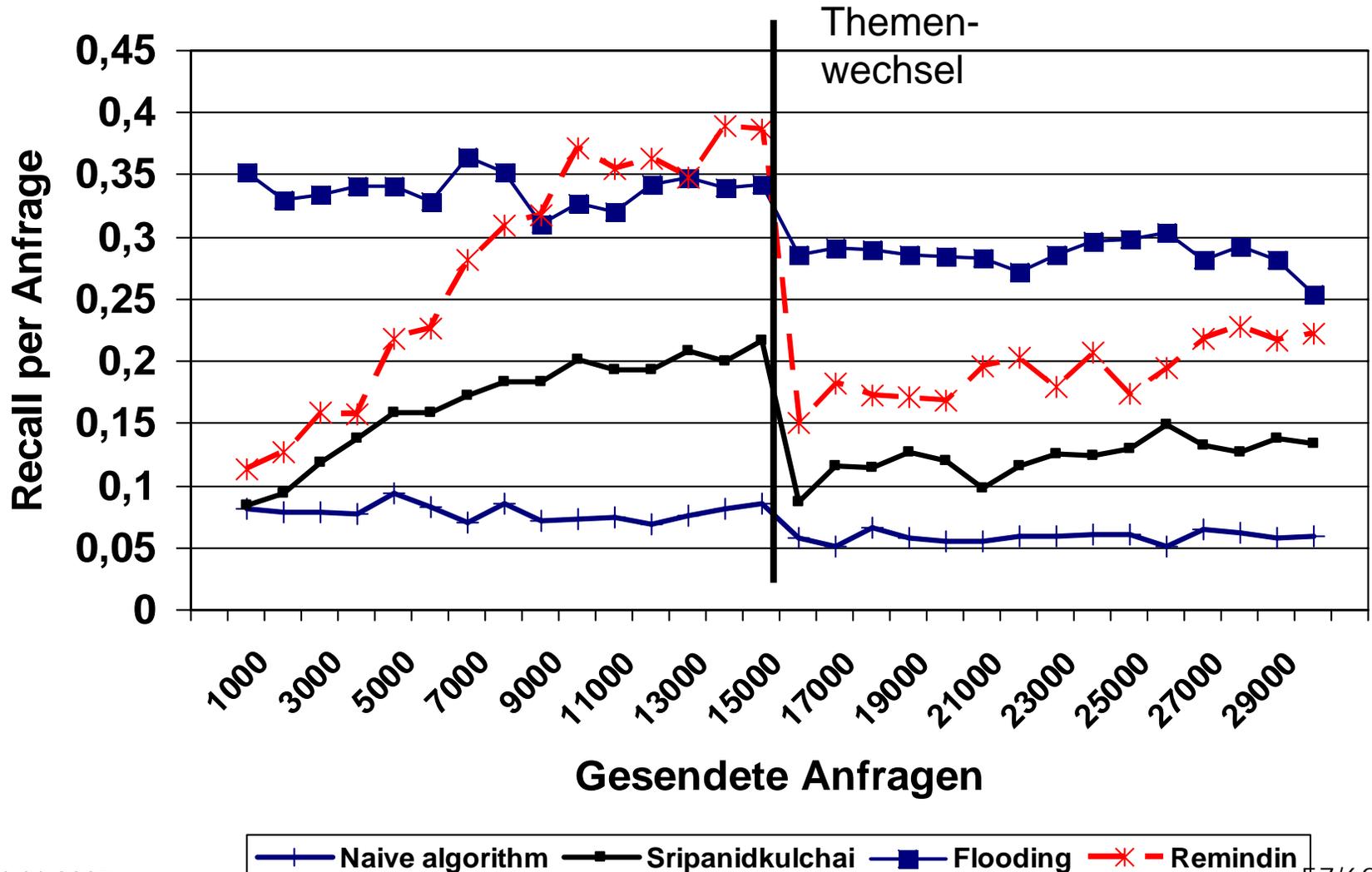
- Vergleich mit verwandten Arbeiten:

- • IBL LRU –40 (Content Provider Shortcuts + LRU)
[Kunwadee Sripanidkulchai 2003]
- • Naive (Gnutella Style Network)

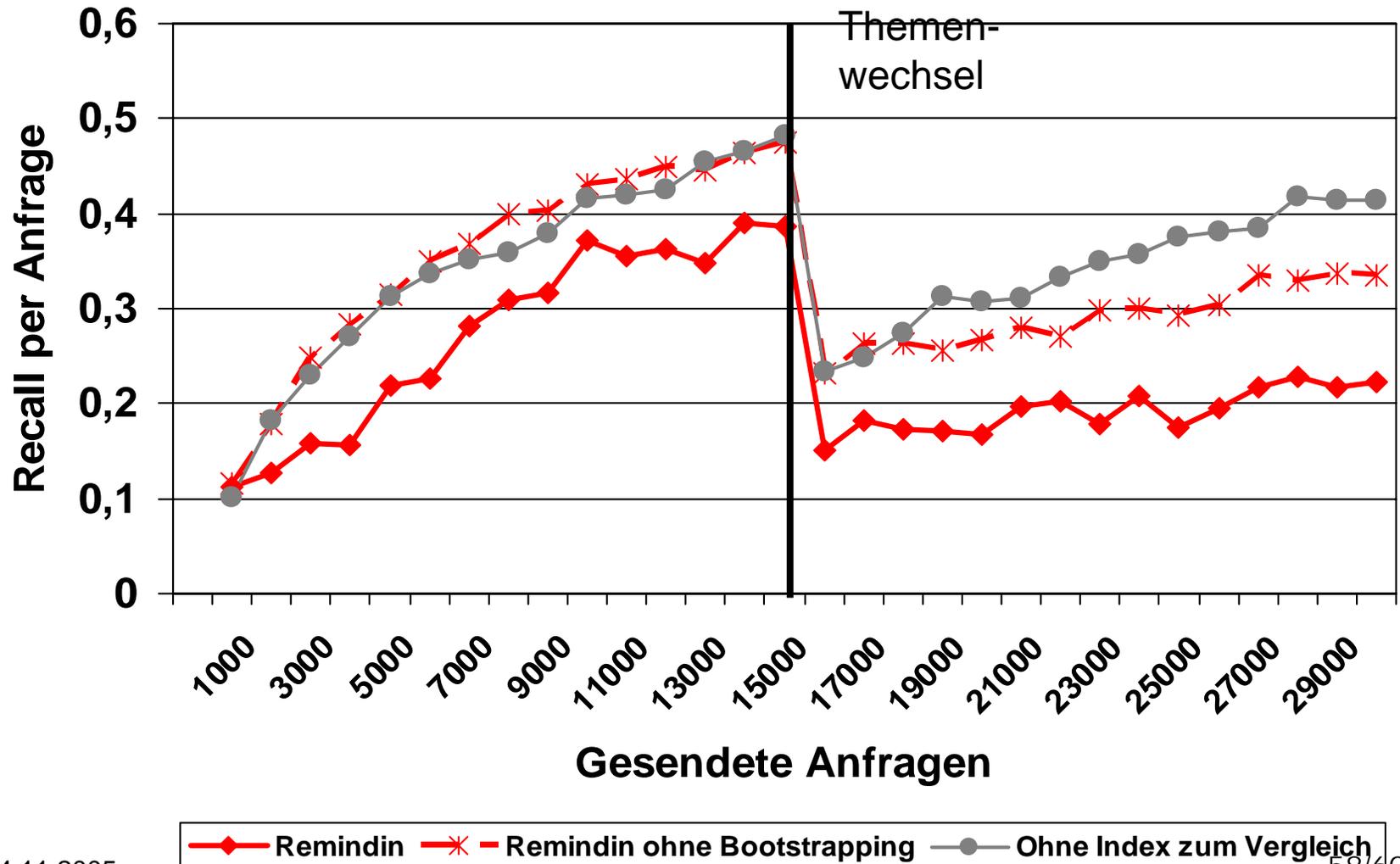
Vergleich der Routing Algorithmen (Nachrichten)



Vergleich der Routing Algorithmen (Recall)



Vergleich der Routing Algorithmen (Recall)



Remindin: Mein Beitrag

- Definition eines Routingalgorithmus für Unstrukturierte Peer-to-Peer Systeme
- Ausnutzung der Ontologie als Wissensbasis
- Finden von Informationen effizient möglich

Weitere Arbeiten

■ DILIGENT

- Methodik zur Unterstützung der automatischen Ontologierstellung
- Kostenmodelle für die Ontologierstellung

■ REMINDIN'

- Weitere Datensätze
- Verknüpfte Anfragen
- Volatile Netzwerke
- Netzwerkanalyse

Zusammenfassung

- Erste Methodologie zur Unterstützung der Ontologieerstellung in verteilten Systemen
- Modell zur Formalisierung nachvollziehbarer und effizienter Argumentation
- Effizienter Algorithmus zum Finden von Peers mit ähnlichen Interessen

Publikationen

- [ESWS2004] H. S. Pinto, S. Staab, Y. Sure, C. Tempich: *OntoEdit Empowering SWAP: a Case Study in Supporting Distributed, Loosely-Controlled and evolving Engineering of oNTologies (DILIGENT)*. **ESWS2004**
- [ESWC2005] C. Tempich, H. S. Pinto, Y. Sure, S. Staab: *An Argumentation Ontology for or Distributed, Loosely-controlled and evolving Engineering processes of oNTologies (DILIGENT)*. **ESWC2005**
- [ECAI2004] H. S. Pinto, S. Staab, C. Tempich: *DILIGENT: Towards a fine-grained methodology for Distributed, Loosely-controlled and evolving Engineering of oNTologies*. **ECAI 2004**
- [PAKM2004] C. Tempich, M. Ehrig, C. Fluit, P. Haase, E. L. Martí, M. Plechawski, S. Staab: *XAROP: A Midterm Report in Introducing a Decentralized Semantics-based Knowledge Sharing Application*. **PAKM 2004**
- [WWW2004] C. Tempich, S. Staab, A. Wranik: *REMINDIN': Semantic Query Routing in Peer-to-Peer Networks Based on Social Metaphors*. **WWW 2004**
- [ISWC2005] A. Löser, C. Tempich, B. Quilitz, W.-T. Balke, S. Staab, W. Nejdl. *Searching Dynamic Communities with Personal Indexes*. **ISWC2005**

Dankeschön

Christoph Tempich

Karlsruhe, 4.11.2005