

意味情報ネットワークに適した Context のモデル化

木下研究室

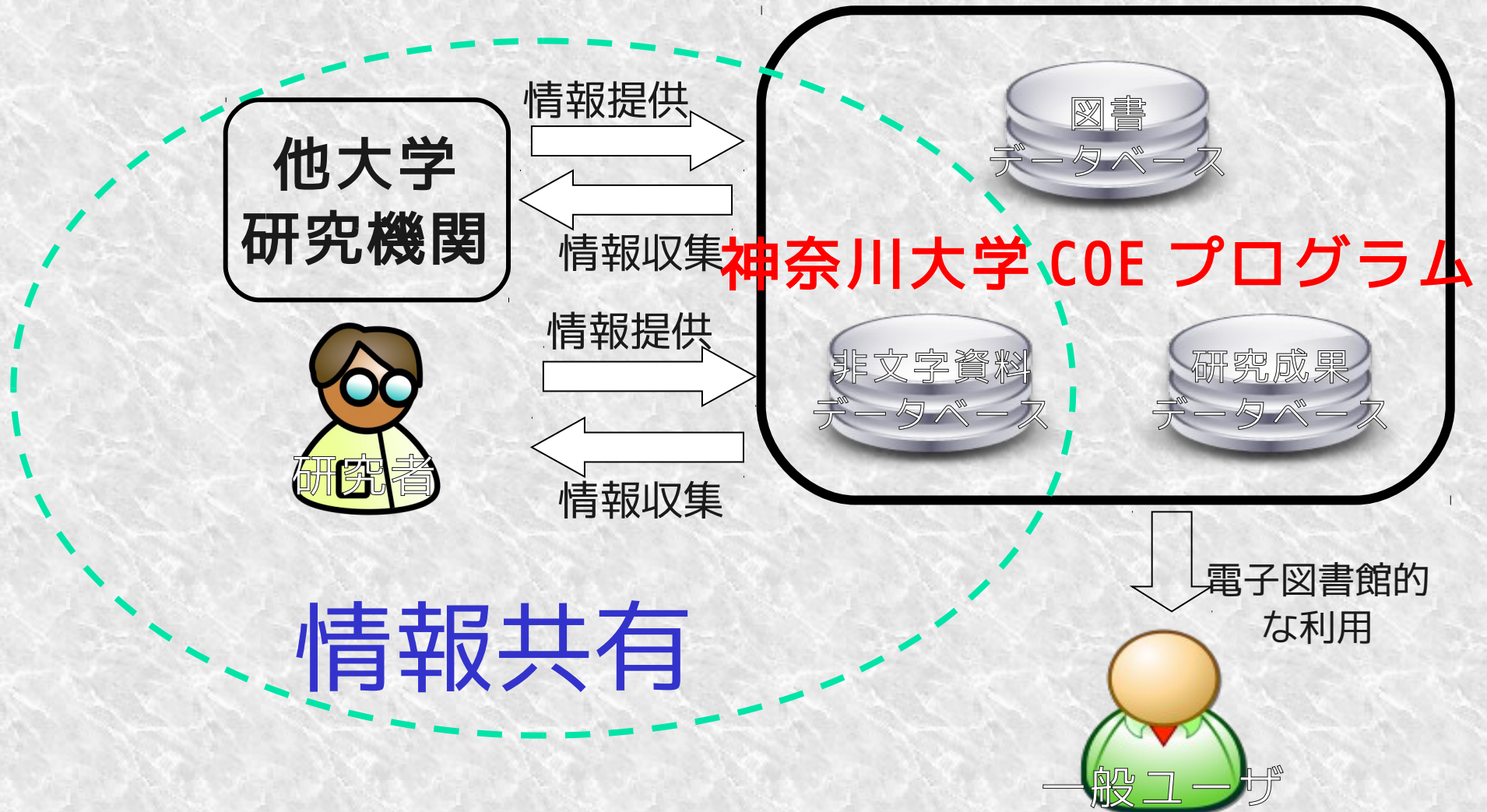
200770083

山下祐平

背景

- 非文字資料の**流通**（提供・収集）、**共有**
 - 既存のシステムでは、
 - データがドメイン固有となり、ドメイン間で情報の相互利用ができない
 - 人間が各ドメインの知識について理解し、モデル化しなければならない
- **新たな知見**や、**関連性**の発見支援

神奈川大学 21 世紀 COE プログラム 人類文化研究のための非文字資料の体系化




非文字資料（民具資料）の例

生摩明良山棟用良木枕用具化アサシ物 H12.1.17

民俗資料調査カード

重要

資料名	(地方名) 4ユウナ	(標準名)	写真・形状・寸法等
借入日	昭和・平成 年 月 日	通番号	1916
寄贈者	(住所) 只見町大字 字 番地	分類番号	28117-7
所有者	(氏名)		21122-9
材質	鉄 雑木		
使用年代	昭和 年頃から 平成 年頃まで・現在も使用中		
使用目的	屋根の削りに使った		
収蔵場所	旧五十嵐家		
備考	ちゅうな		
調査年月日	平成 2 年 10 月 26 日		
調査員	目黒鶴吉		



只見町教育委員会

例 1) チュウナ

只見町民俗資料調査カード

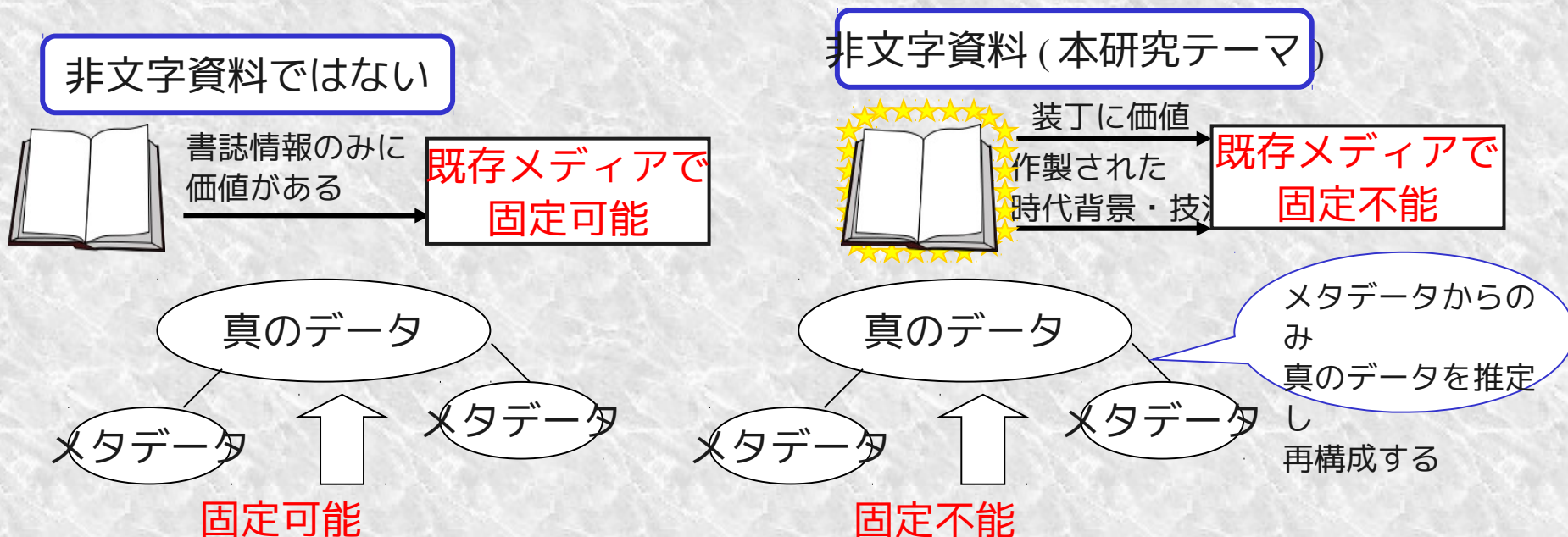
第1次分類番号	通番号	分類番号	第2次分類番号	分類	番号
	1560			9	237

標準名	カリアガリ ユッコギ (ヤ) 別名	<input type="checkbox"/> 写真・感度・寸法・重量など	<input type="checkbox"/> 紀年名・使用・墨書など
採集年月日	昭和・平成 年 月 日	採集者	只見町教育委員会
採集地	只見町大字 採取 字		
所有者	(住所) 只見町大字 採取 字 (氏名)		
入手経路	(田所有者) 自家製・購入・譲渡・もらう・交換・その他 () 代価 (只見町) 寄贈・購入・製作・その他 () 代価		
材料名	入手方法 採取・栽培・購入		
製作地	製作者		
製作年代	明・大・昭和 年頃まで製作 現在も製作中		
採取・処理・製作方法	<ul style="list-style-type: none"> ・ 加茂織・青木織などを購入し、自家製。 ・ 呼び名は各地で、又は家庭ごとにもそれぞれ違った呼び名をしている。 ・ カリアゲユッコギ ・ ホソユッコギ ・ サルツバカマ ・ カリアガリユッコギ ・ ホソユッコギ ・ モンベ (腰にゴム紐を使ったもの) ・ 男物は小便マチがあり女物にはない。 		
使用地	使用者		
使用年代	明・大・昭和 年頃まで使用 現在も使用中 使用時期		
使用器具	保存・管理場所		
使用方法	<ul style="list-style-type: none"> ・ 山仕事、田畑、狩猟等屋外の仕事に着用した。 ・ また、細くて軽く、動き易いため、屋内屋外を問わず仕事着として使用するに便利である。 		
他の使用方法・由来・俗信・禁忌使用地域など			
話者名	(住所) 只見町大字 字 (生年月日) 明・大・昭和 年生		
話者名	(住所) 只見町大字 字 (生年月日) 明・大・昭和 年生		
ネカ番号	15-9-2	調査年月日	平成 2 年 7 月 4 日
		調査員	馬場 尊



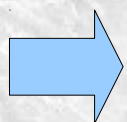
例 2) カリアゲユッコギ

非文字資料とメタデータ



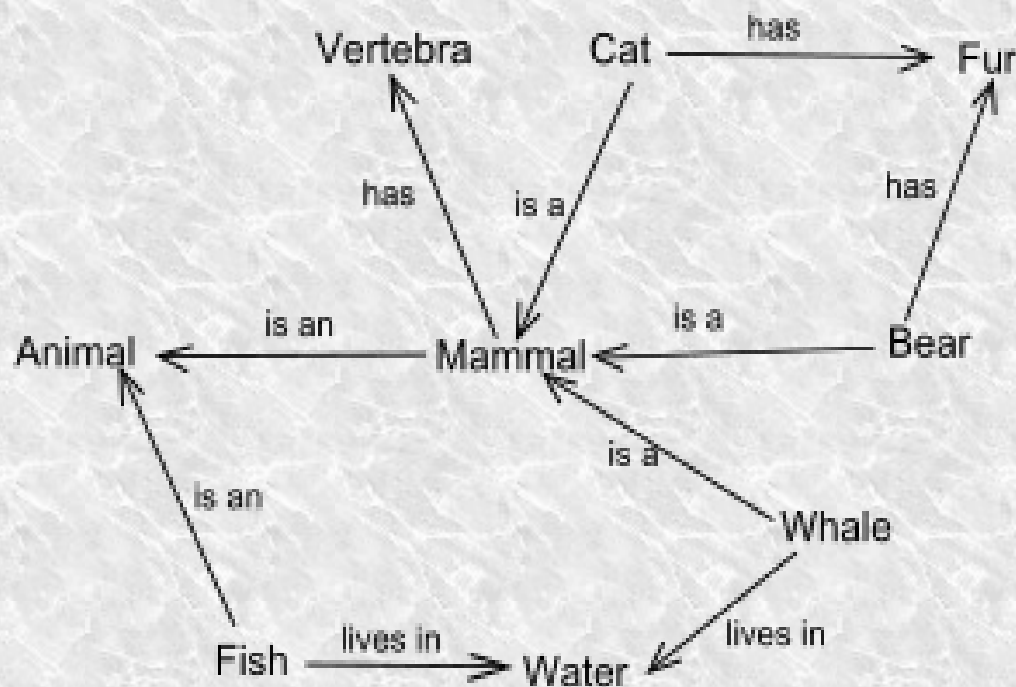
メタデータの観点から見た非文字資料

詳細なメタデータ記述 = 真のデータの再構成



メタデータを付与し、その意味を
Ontology によって与える

Ontology を用いた メタデータの概念の定義



メタデータ + Ontology によって
意味ネットワークが構築される

- 概念と概念間の関係を記述する
- 概念では記述可能なメタデータを定義
 - e.g. Animal では体長、体重など
- 継承関係などが現れる場合がある
 - e.g. 体長は

セマンティックな検索の問題点

- グラフ探索なので探索コストがグラフの規模の増加に伴い急激に増加してしまう
 - グラフを Context に分割し探索範囲を限定
- ただし、適切な分割が不可欠
- その上で、分割された Context 間を接続する紐を定義する
 - 共通語彙など

Context

- Context は文脈
 - 前後関係、背景あるいは暗黙の条件といった意味合いを持つ
 - Context の例
 - 意味グラフの Context
 - 情報リソースの意味グラフが持つ Context
 - クエリの Context
 - クエリに含まれる前提条件
 - ユーザの Context
 - 言語、年齢、性別、所属などのユーザが持つ属性

非文字資料における Context の例

民具における Context の例



■ Context 内の Ontology

推論や探索に利用しやすいように
Context を小さくする



Context が小さいほど Context 間の
Ontology 構築が容易となる

■ Context 間の Ontology

Context 間の違いが
従来の Ontology では記述出来ない



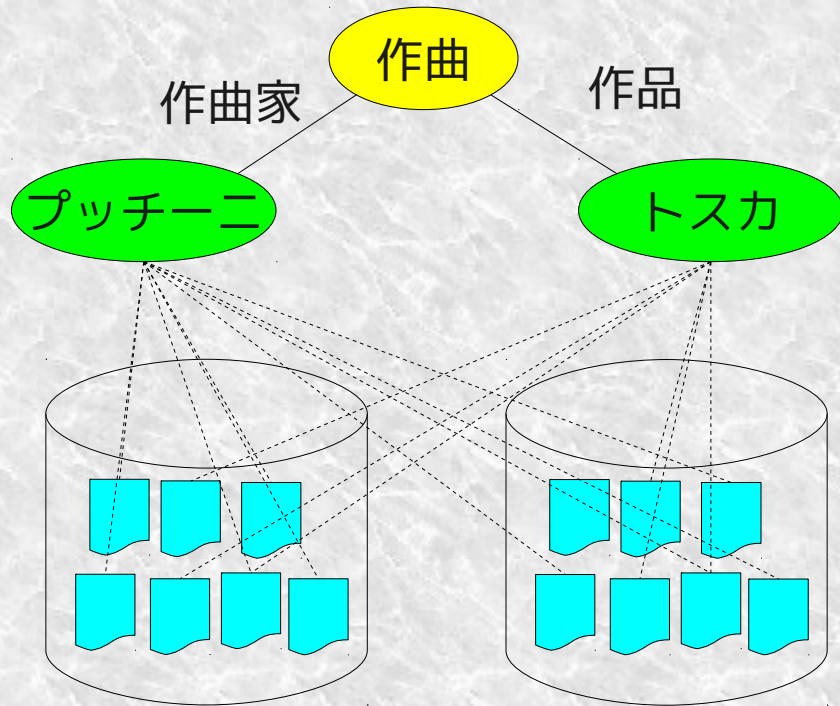
Context 間に着目する事で . . .

異なる Context 間にまたがった同一の
Entity の関連性を表す事が出来る

目的

- 意味情報に基づいた非文字資料の体系化、デジタルアーカイブの構築のためのモデルの提案
 - Context モデルの有効性
 - 適切なクエリの生成、推論などの人間が行っていた部分をマシンが半自動的に担う

トピックマップ

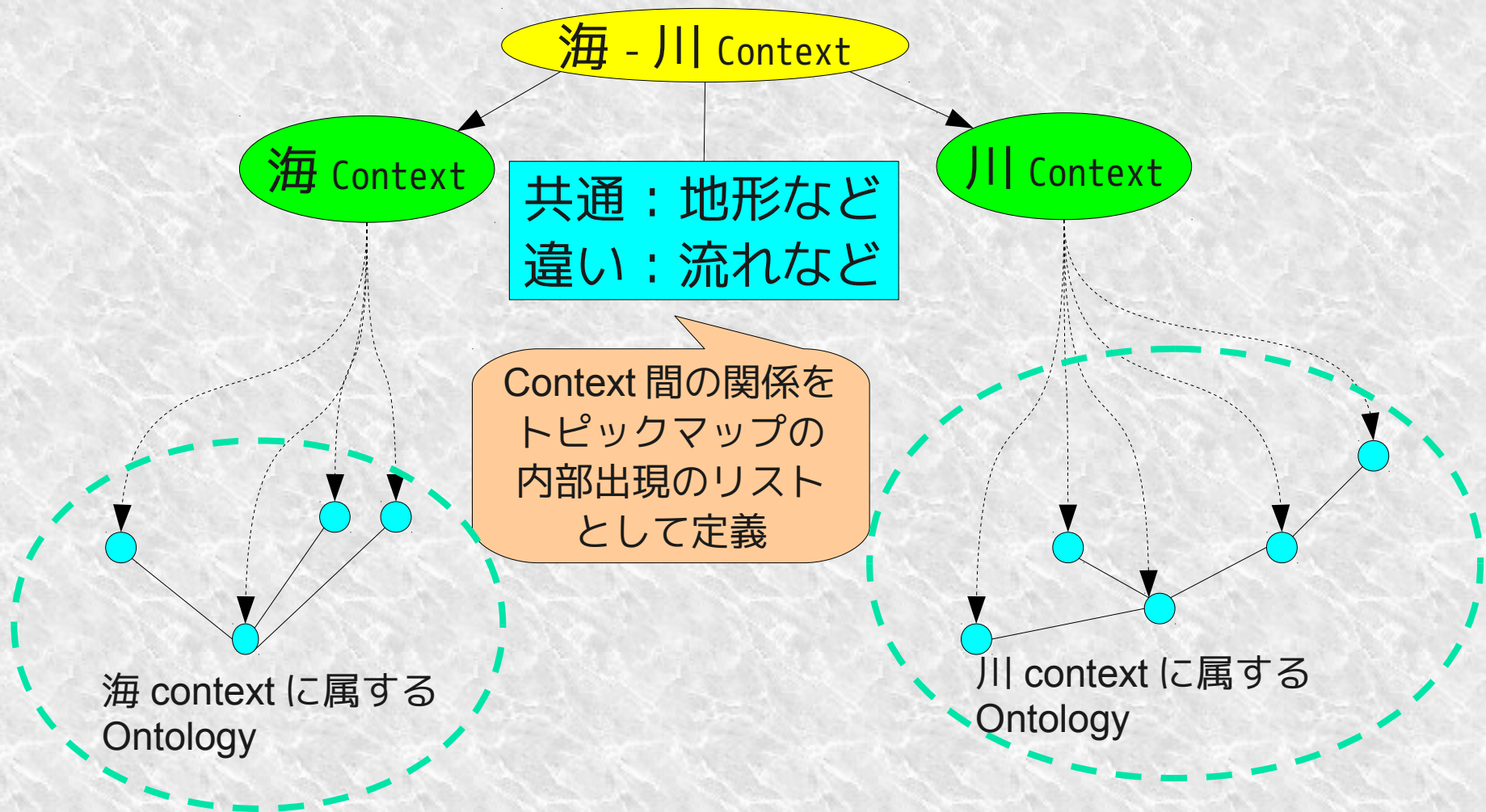


- トピック (Topics)
 - 問題領域での**主題**
- 関連 (Associations)
 - **主題間の関係**を表現
 - 関連の接続先のトピックの役割はロールとして記述される
- 出現 (Occurrences)
 - 主題に関連する**情報**

トピックマップのまとめ

- トピック、関連を自由に定義でき、情報リソースとの関連付けができる
 - 意味情報ネットワークと情報リソースが分離
 - シソーラス、Ontology、ファセット分類などのデータ構造を簡単に定義できる。
 - 既存の Ontology や RDF などの資産をトピックマップにマッピングすることができる。
- 主題が公開可能で、情報の共有流通が容易
 - ➡ 非文字資料に必要な Context 間の関係を記述する方法に適している

トピックマップを用いた提案データ構造



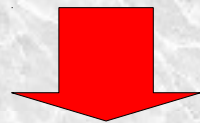
Context の関係性

- Context を利用するためには、それらの関係性が記述されている必要がある
 - e.g. 江戸の食文化と明治の食文化では、時代の Context が異なっている
- たくさんの Context を定義できることが望ましいが、それらの関係性は無数にあり当初の目的である計算量の削減に反している



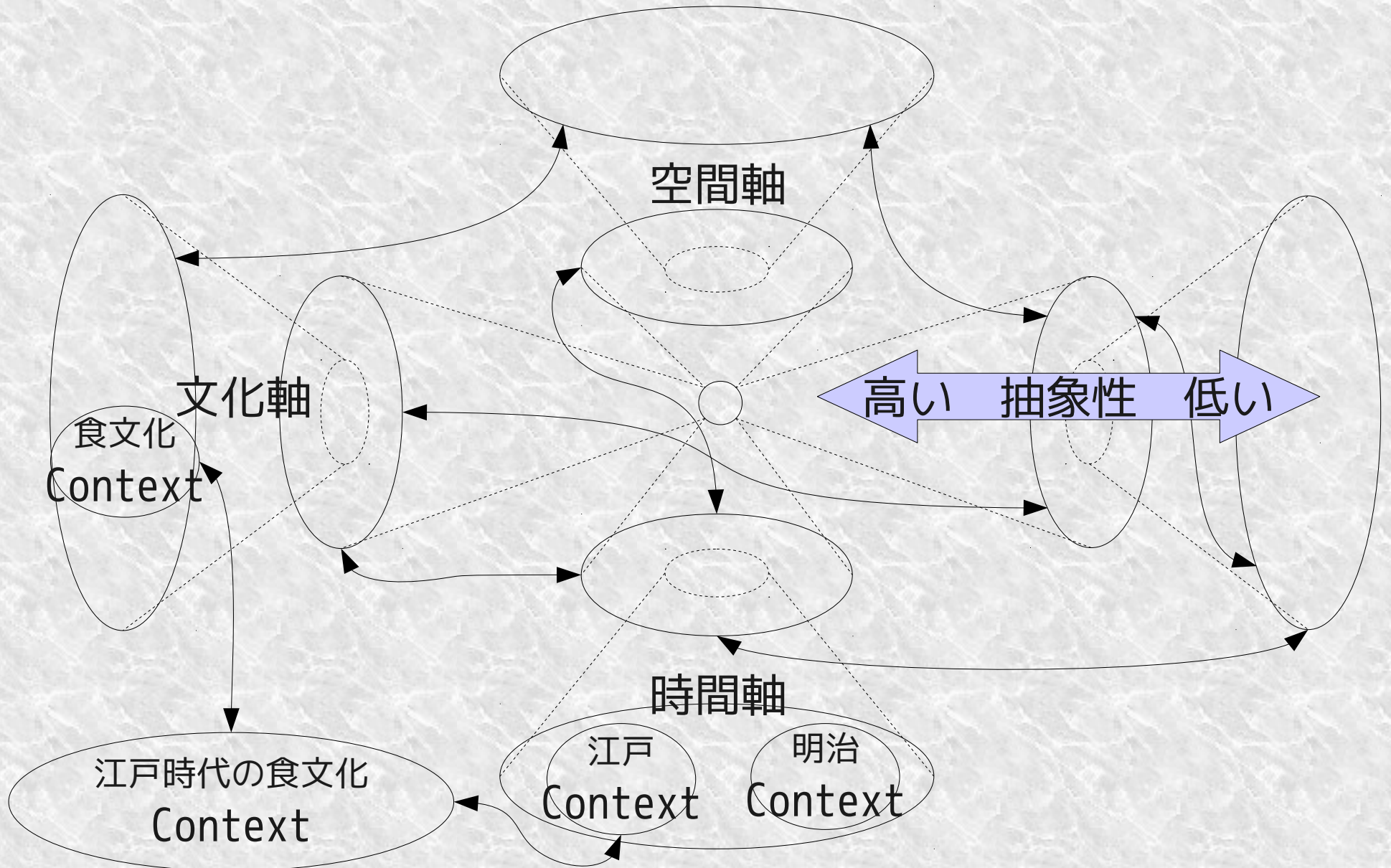
Context の抽象度と組み合わせ

- 制約を加えないまま Context を増やすと Context 間の関係が無数に現れるのでそれらをすべて定義するのは難しくなる。

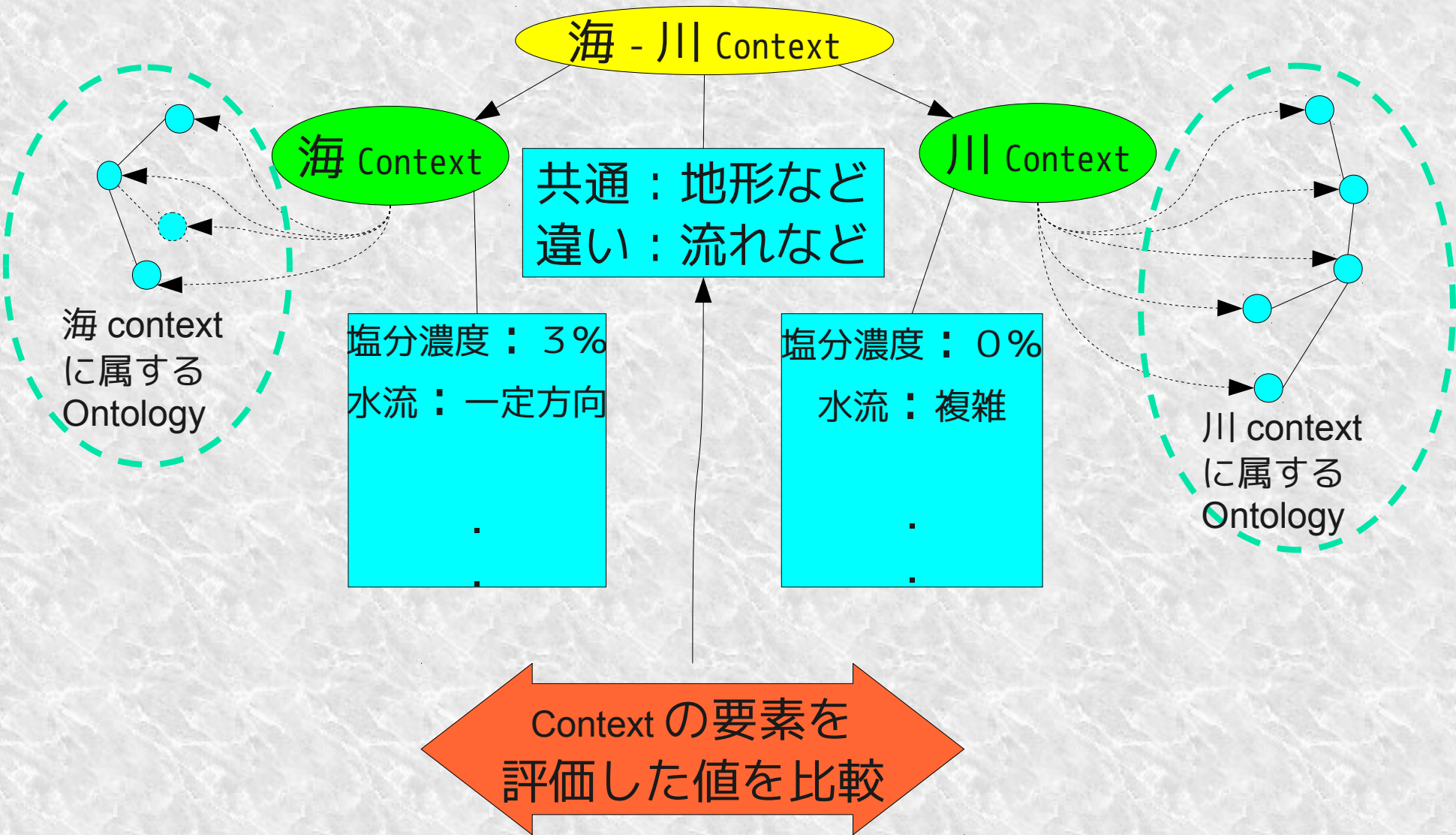


- 多次元の軸に分離し、抽象度を定義する。
それらの組み合わせで任意の Context を表現する。
- それぞれの軸での特徴から Context の関係を導く

Context の抽象度と組み合わせ



提案モデルのデータ構造



提案モデルの特徴を活用した検索

- Context に分ける意味
 - 探索コスト削減以外の魅力が少ない
 - Context に分離する別の目的を決めることで、どのように Context 分けるのがよいのかがわかる
- 推論もしくは探索経路に活用
 - Context 同士を比較し、それらの特徴から新しいノードを推論する
 - 重要なパスを先に探索するなど

まとめ

- 意味グラフを適切な Context に分割すると
 - 計算量の削減だけでなく探索、推論の優先度にご利用可能
- Context 同士を紐で関係付けるには、トピックマップが適している
- Context を抽象度に基づいた木構造とすると
 - Context の組み合わせであるとする事で、従来検索の利点を活用できる

今後の課題

- 人間による主観的な評価方法の決定

発表リスト

- トピックマップを用いた非文字資料における Context の表現
 - 技術と社会・倫理研究会
- Context 表現のための手法の考察
 - 技術と社会・倫理研究会