

# Wikipediaカテゴリオントロジー

- Wikipediaカテゴリの知識工学的観点  
からの活用を目指して -

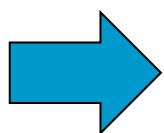
北海道大学大学院情報科学研究院

知識ベース研究室

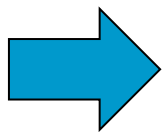
吉岡真治

# 「Wikipediaカテゴリ」とは

- インターネット上の百科事典Wikipediaにおいてページを分類するための情報
- Wikipediaにおけるカテゴリの定義
  - カテゴリは第一義として、「分類」を示すものです。
  - 上記に加えて、ウィキペディアのカテゴリとしては「関連が深いキーワード」を示すことができます。
  - また、カテゴリはウィキペディアの骨組みの意味を持ちます。



Wikipediaカテゴリには、概念階層などの知識工学に有用な情報が存在



YAGOや日本語Wikipediaオントロジーの提案

# Wikipediaのカテゴリ(英語版の説明)

## ■ 2種類のカテゴリの存在

– トピックカテゴリ (Topic categories)

- トピック名 (多くの場合、Wikipediaのページ名と同じ)

例: Category:France

– セットカテゴリ (Set categories)

- クラス名 (多くの場合、複数形).
- 例: Category:Cities

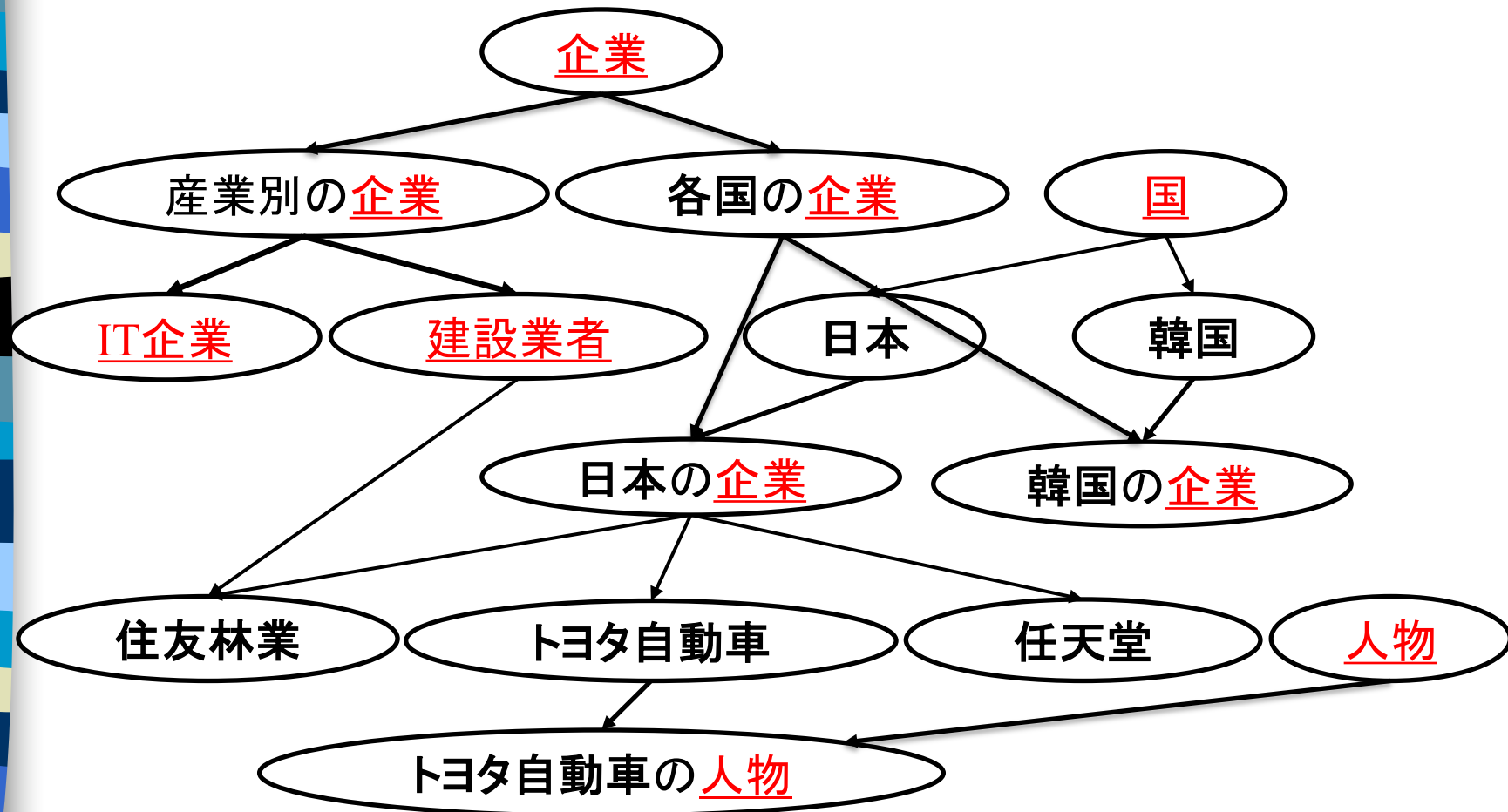
## ■ これらのクラスが組み合わされることが多い。

– Category:Cities of France

– 日本語の分割のためのカテゴリの多くは、この2種類のカテゴリの組み合わせ

# Wikipediaカテゴリ階層構造の問題

- 親子関係を推移的な包含関係と捉えると問題が生じる

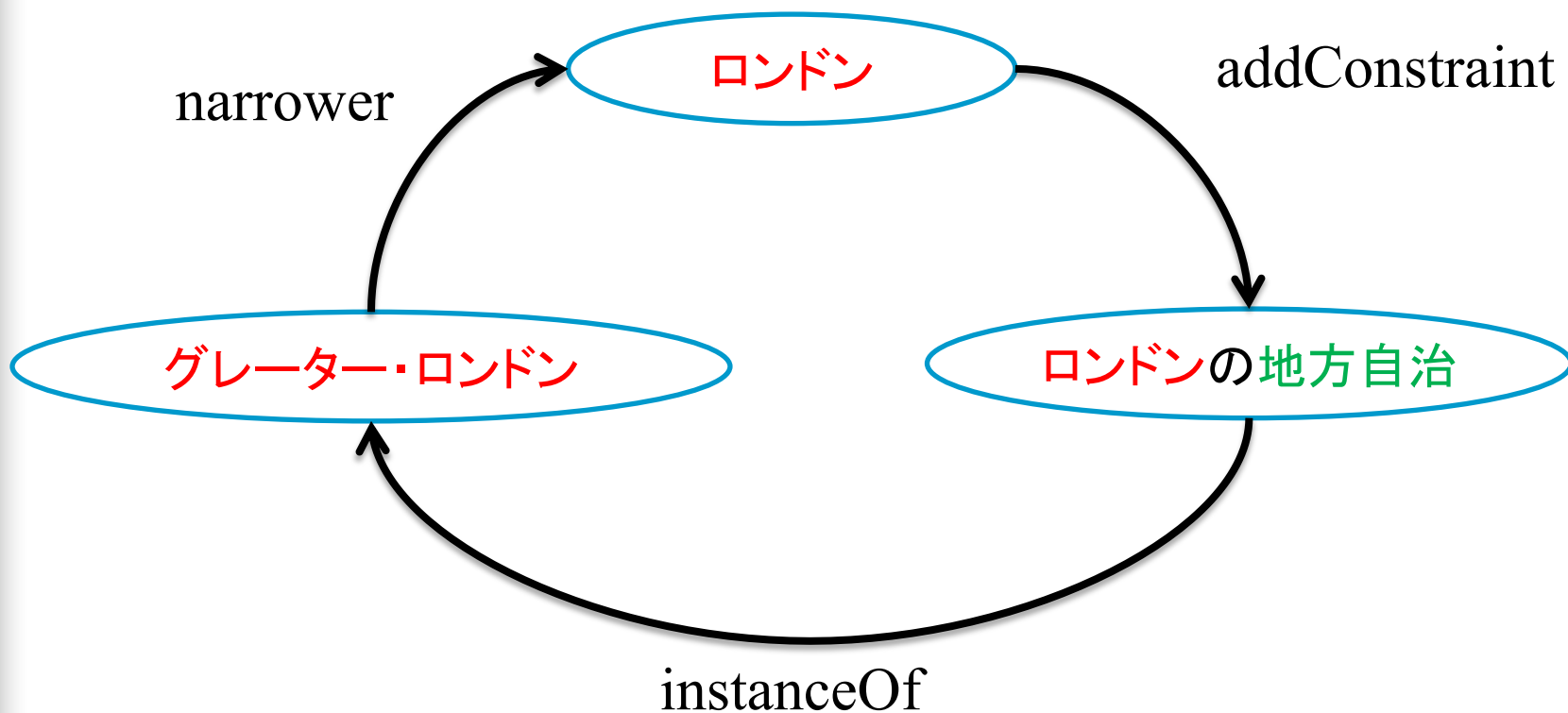






# 階層構造中のループの解消

- Wikipediaの階層構造には、ループが存在
- 包含関係が成り立たない階層関係が問題





# <http://wcontology.org>

- 全てのデータはWebサイトで公開
  - SPARQLのエンドポイントも公開

## Wikipedia Category Ontology (WCO)

updated on June 9, 2020

[Japanese explanation](#) follows with English explanation.

[日本語の説明](#)は、英語の後にあります。

### Summary

Wikipedia Category Ontology (WCO) is an ontology for representing Wikipedia category information from knowledge engineering aspect. At this moment, All categories of Japanese Wikipedia (Dump: 2017/10/20) are organized as instances of the ontology. This ontology can be used for manipulating Wikipedia category (e.g., extracting transitive category hierarchy, selection of topic/set categories). As a result you can also obtain class-subclass hierarchy used for the ontological engineering.

### Download (2020/06/09 version)

- [RDF file](#) for definition of core terms
- [RDF file](#) for representing Japanese Wikipedia categories



# 関連発表論文

- Masaharu Yoshioka and Takanori Nakagawa: [Wikipedia Category Ontology: A Framework for Utilization of the Wikipedia Category Structure by Knowledge Engineers](#). In *Proceedings of the ISWC 2020 Demos and Industry Tracks: From Novel Ideas to Industrial Practice co-located with 19th International Semantic Web Conference (ISWC 2020)*, pp. 300-305, 2020.
- 吉岡真治: [Wikipediaカテゴリオントロジー - Wikipediaカテゴリの知識工学的観点からの活用を目指して -](#). 人工知能学会第55回セマンティックウェブとオントロジー研究会, SIG-SWO-055-05, 2021.