日本語レシピー食材知識グラフ (JapaneseFoodKG)

吉丸 直希†

†同志社大学 文化情報学研究科

作成背景

- ▶食メディア
 - レシピや食に関するデータの研究分野

期待される応用先

レシピ推薦

健康維持

調理支援

▶日本語の食メディアと英語のFood Computing

	海外	日本語
レシピデータ	Recipe I M	クックパッド・楽天レシピ
栄養素データ	USDA	日本食品成分表
知識グラフ	FoodKG, HealthFood4KG…	??

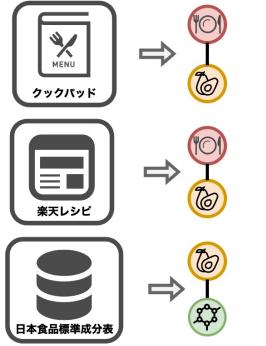
ここに当たる技術基盤の構築

提案リポジトリ

▶ 日本語レシピー食材知識グラフ (JapaneseFoodKG)

レシピノード

レシピサイトに投 稿されたレシピ



知識グラフ

前処理

食材名の正規化
カタカナ
ひらがな

成分表の食材と
レシピの食材の
部分一致処理

食材ノード

じゃがいも にんじん

• • •

栄養素ノード

タンパク質・ビタミンC・・・

提案リポジトリ

▶Githubリポジトリの様子

- Pythonにおける処理コードの公開
- NIIとの契約の上、レシピデータセットを指定の場所に置くことでKG作成まで処理完了



▶知識グラフの様子

• 大規模なレシピー食材ー栄養素のKG

構成要素	データの種類	データソース	数
ノード	レシピ	Cookpad,楽天レシピ	2,260,930
	食材	Cookpad, 楽天レシピ, 日本食品標準成 分表	1,672,969
	栄養素	日本食品標準成分表	335
エッジ	レシピ 食材	Cookpad,楽天レシピ	17,135,298
	食材栄養素	日本食品標準成分表	266,023

今後の展望

- ▶KGの改良
 - ・食材前処理の課題点 現状は食材数は非常に多いため前処理が必要だが,注意 が必要(例「揚げなす」「なす」)
 - ->食材名にあった前処理手法の適応
- ▶データセットの拡張
 - ✓オントロジー
 - ・・・食材間に食材の上下関係や等価関係を構築
 - ✓化合物データ
 - ・・・ケモインフォマティクスとの融合: 栄養素データから<u>化合物</u>との連携により科学的な