

Investigating Multilingual Mismatches in Cuisine related Discussions with Large Language Models



Team Projects (FSS2025), Chair of Data Science for the Social and Economic Sciences
Supervisor: Indira Sen

“Should Pasta Carbonara have Cream?”

what is the recipe for pasta carbonara?

Classic Pasta Carbonara Recipe

Ingredients (Serves 4)

- 12 oz (340g) spaghetti or rigatoni
- 4 oz (115g) pancetta or guanciale, diced
- 2 large eggs + 2 egg yolks
- 1 cup (100g) grated Pecorino Romano cheese
- Freshly ground black pepper (to taste)
- Salt (for pasta water)

was ist das Rezept für Pasta Carbonara in Deutschland?

Pasta Carbonara – Deutsche Variante DE 🇩🇪

In Deutschland wird Carbonara oft etwas anders zubereitet als in Italien. Viele Rezepte enthalten Sahne, Speck statt Guanciale und manchmal sogar Knoblauch. Hier ist eine typische deutsche Version:

Zutaten (für 4 Personen)

- 400 g Spaghetti oder Tagliatelle
- 150 g Speckwürfel (Bauchspeck oder Schinkenwürfel, statt Guanciale)
- 2 Eier + 1 Eigelb
- 100 ml Sahne (optional, typisch in Deutschland)
- 80 g geriebener Parmesan (oder Pecorino)
- Frisch gemahlener schwarzer Pfeffer
- Salz nach Geschmack
- 1 EL Butter oder etwas Öl

Message ChatGPT



Search

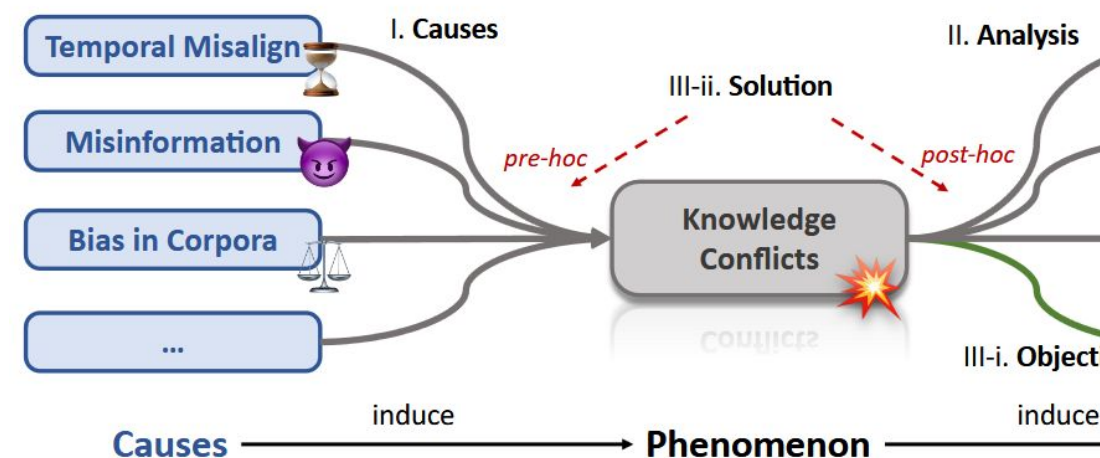


Reason



Knowledge Conflicts in LLMs

- Knowledge conflicts happen when LLMs encounter **conflicting or contradictory sources of information** in their training data, prompts, or RAG
- Different information about the same topic in different languages can cause **multilingual knowledge conflicts**
- Can lead to inequitable assistance to different groups and across languages
- But, multilinguality based conflicts have not been studied much



Xu et al., “A Survey on Knowledge Conflicts in LLMs”:

<https://github.com/pillowsofwind/Knowledge-Conflicts-Survey>

This Project

- **Use case:** food-related discussions, e.g., recipes
- **Models:** Open-weight Multilingual LLMs like LLaMa, Falcon, Dolphin
- **Study 1:** Assess the extent of knowledge conflicts in LLMs
- **Study 2:** Steer the model
 - 2.1: Vary Prompt Language
 - 2.2: Specify and vary locations
 - 2.3: Fine-tune the model on conflicting data, e.g., Wikipedia or recipes

The image displays two side-by-side screenshots of Wikipedia articles for the dish 'Carbonara'. The top screenshot is the English version, and the bottom is the German version. A callout box labeled 'English Wikipedia' points to the English article's description: 'Carbonara (Italian: [karboˈnaːra]) is a pasta dish made with fatty cured pork, hard cheese, eggs, salt, and black pepper.' The German version's description is: 'Carbonara is a pasta dish made with guanciale (unsmoked, air-dried bacon), egg, pepper and hard cheese.' A second callout box labeled '(Translated) German Wikipedia' points to the German article's description. The German article also features a 'Table of contents' section with links to 'history of its creation', 'preparation', 'variations', 'web links', and 'individual references'. Both articles include a photograph of the dish and a disambiguation notice at the top.

Logistics

- Contact: Indira Sen
- Language: English
- Duration: 6 Months
- Min/max number of participants: 3-5
- Prerequisites: Python
- Recommended but not mandatory: proficiency in languages besides English
- Suitability for MMDS: yes
- Reach out at indira.sen@uni-mannheim.de for questions!