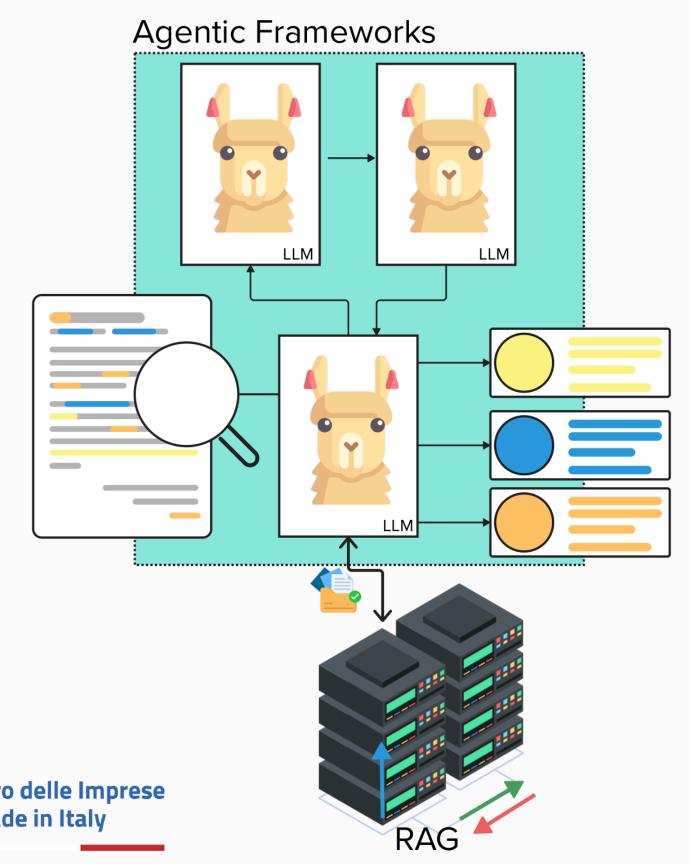
# Complex Applications with Language Models



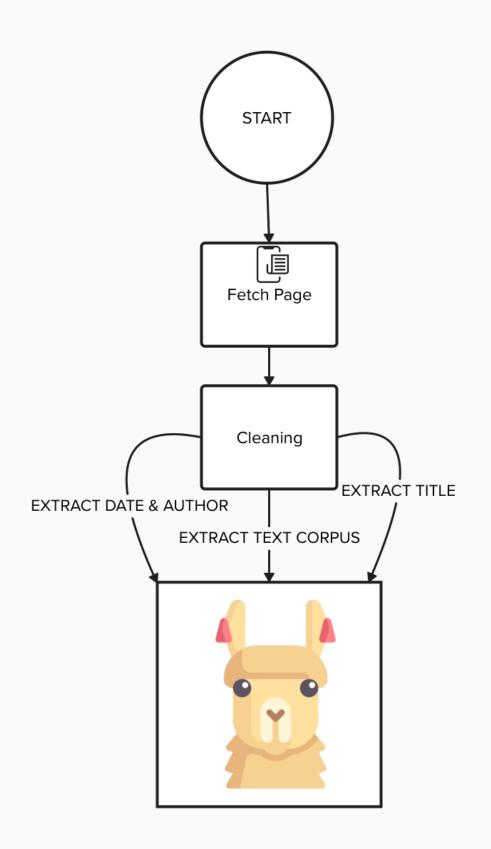




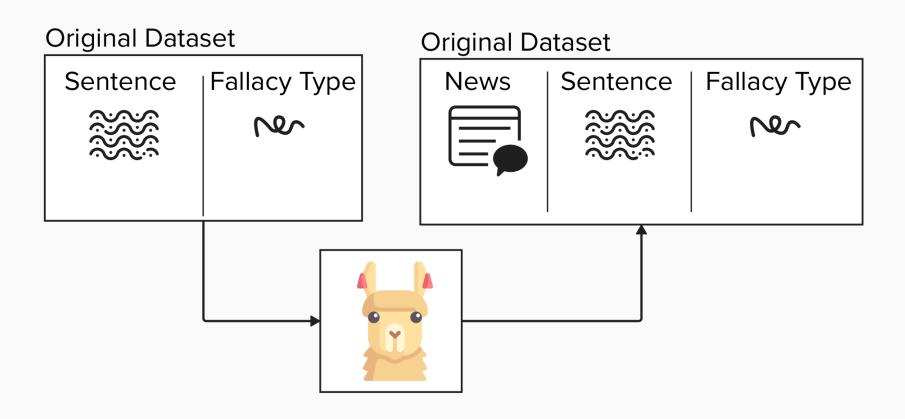




# 1. WebScraping Agent



# 2. Data Augmentation



	sentence	fine_class	generated_news
0	All judges are fair-minded individuals; there	fallacy of logic	A recent study has found that the number of ap
1	Good people don't lie. You told a lie. Therefo	fallacy of logic	A local politician has come under fire after b
2	We should stop using hairspray because it is s	fallacy of logic	A severe winter storm has hit the northeastern
3	"Live with me or live on the streets"	false dilemma	A new homeless shelter in downtown Los Angeles
4	Using a double standard or arguing for an unju	intentional	A recent proposal to increase funding for loca

Luigi Serreli Text Classification and NLP Seminar 2/4

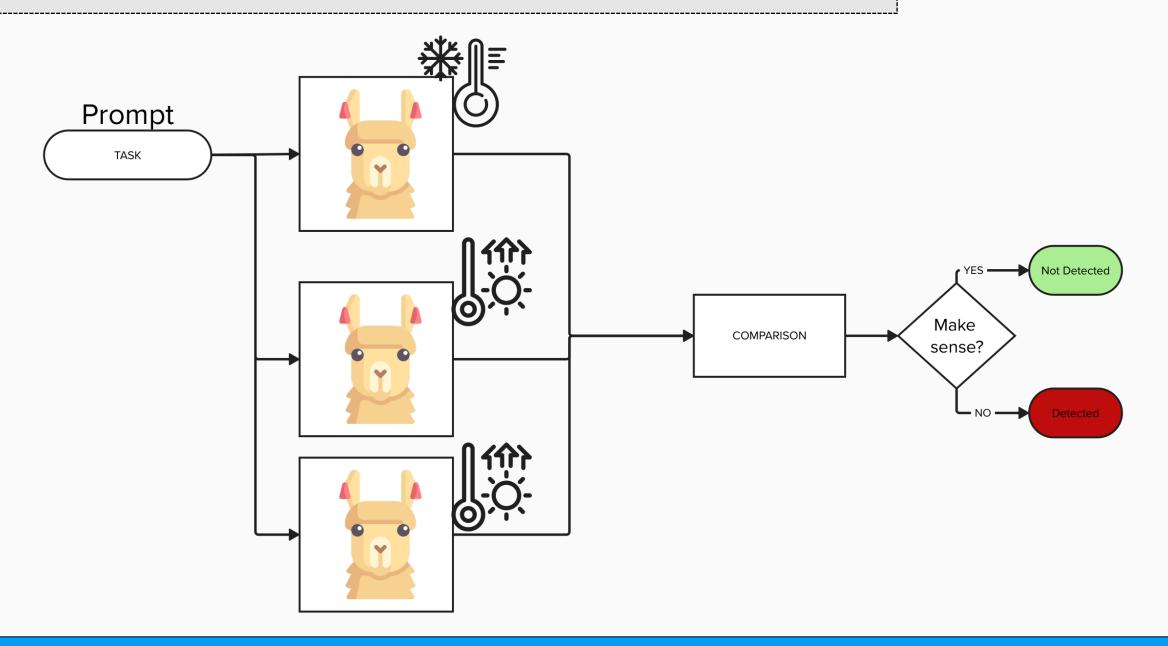
### 3. Simple Hallucination Detection

#### **Definition according to IBM - Hallucinationc**

AI hallucination is a phenomenon wherein a large language model (LLM)—often a generative AI chatbot or computer vision tool perceives patterns or objects that are nonexistent or imperceptible to human observers, creating outputs that are nonsensical or altogether inaccurate.

#### **Famous Hallucinations**

- Google's Bard chatbot incorrectly claiming that the James Webb Space Telescope had captured the world's first images of a planet outside our solar system.
- Microsoft's chat AI, Sydney, admitting to falling in love with users and spying on Bing employees.
- Meta pulling its Galactica LLM demo in 2022, after it provided users inaccurate information, sometimes rooted in prejudice.



# 4. Writer / Reporter

## 5 Different Agents – same model, different task

- Planner
- Researcher
- Generator
- Judge/ Critic
- Thinker

Standards

