

Simplified Universe

Presented by - LMDoctors

Debjyoti
Vivek
Sandeep
Aniruddha
Mrigank



Lexical & syntactic complexity



Patient Health Reports



Legal Documents

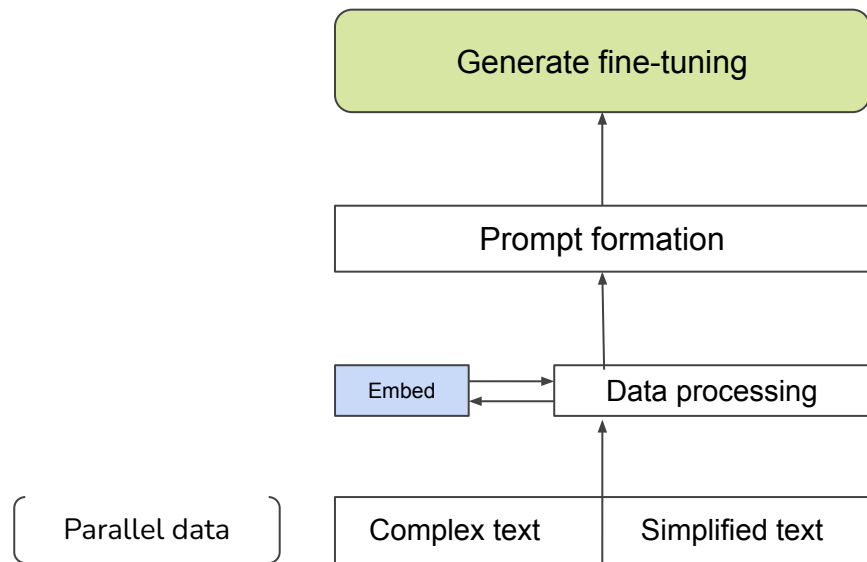
Text Simplification

- Documents from domains like medical, administration, science - are really hard to interpret (for non-experts of the domain or someone from different native language).
- We try to simplify such text using lexical & syntactic simplification.

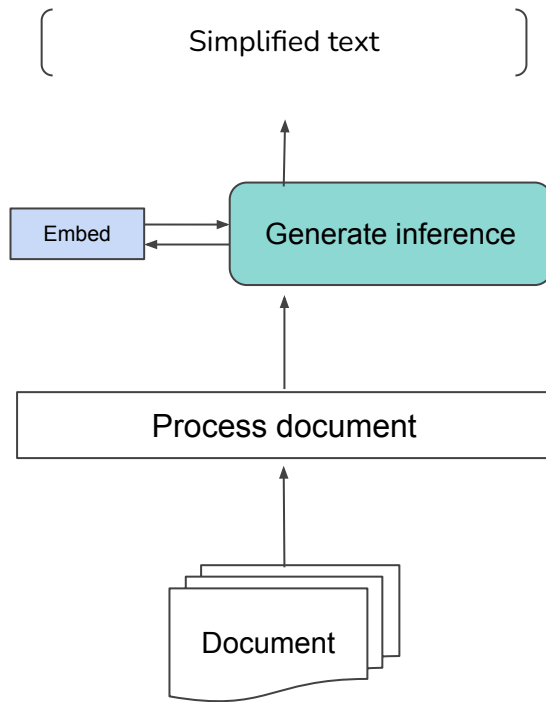
—

- To do so, we employ vast lexical knowledge learned by LLMs (“**Generate**” api).
- We used prompt-engineering to ensure model learns the task along with domains terminology.
- Also leveraged the “**Embed**” api to extract well simplified samples from parallel datasets.

Our Approach - pipeline



Fine-tuning



Inference

Datasets

- **Asset Corpus by Facebook Research**
The corpus is composed of 2000 validation and 359 test original sentences that were each simplified 10 times by different annotators
- **Medical Wiki Parallel Corpus for Medical Text Simplification**
The corpus consists of 3.3k pairs of sentences of medical domain, with one sentence from English Wikipedia and a corresponding sentence from Simple English Wikipedia.
- **SimPA Corpus**
Released as a part of SIMPATICO - a Horizon H2020 project
Domain: Public Administration Documents
3,300 lexical simplifications (3 for each original sentence)
1,100 syntactic simplifications (1 for each original sentence)

Simplified universe

Let us simplify the world for you

This reflects the fact that the Parliament's President is not part of the executive branch but still ranks higher in (symbolic) importance than the Chancellor.



This is because the Parliament's President is not part of the executive branch but still ranks higher than the Chancellor.

Simplified universe

Let us simplify the world for you

According to natural law theory, all people have inherent rights, conferred not by act of legislation but by "God, nature, or reason."

1

The law of nature is that all people have full rights. These rights are determined by "God, nature, or reason."

Simplified universe

Let us simplify the world for you

A lymph node is a small ball-shaped organ of the immune system, distributed widely throughout the body including the armpit and stomach/gut and linked by lymphatic vessels. Lymph nodes are garrisons of B, T, and other immune cells.



The lymph nodes are small ball-shaped organs in the body. They are found throughout the body, including the neck, armpit and stomach. They contain lymphocytes that work together to help fight infections.

Simplified universe

Let us simplify the world for you

Leptospirosis (also known as Weil's disease, Weil's syndrome, canicola fever, canefield fever, nanukayami fever, 7-day fever, Rat Catcher's Yellow, Fort Bragg fever, and Pretibial fever) is a bacterial zoonotic disease caused by spirochaetes of the genus *Leptospira* that affects humans and a wide range of animals, including mammals, birds, amphibians, and reptiles.

It is caused by the genus *Leptospira*.

It is an infectious disease of humans and a wide range of animals. It is caused by *Leptospira* species as well as other bacteria.

The bacteria which cause it are

Simplified universe

Let us simplify the world for you

Although the central ideas of natural law had been part of Christian thought since the Roman Empire, the foundation for natural law as a consistent system was laid by Aquinas, as he synthesised ideas from his predecessors and condensed them into his Lex Naturalis (lit. Natural law)

The basis for natural law as a system was laid by Aquinas. He combined ideas from his predecessors and condensed them into his Lex Naturalis (Lex or language), which he called Lex humanitatis.

Simplified universe

Let us simplify the world for you

Because many consider the information in medical records to be sensitive private information covered by expectations of privacy, many ethical and legal issues are implicated in their maintenance, such as third-party access and appropriate storage and disposal



Because medical records are not sold to anyone but the individual who owns them, they are protected private information. Third-party access may be available and appropriate storage and disposal can be done properly.

Sample Training point

--separator--

Simplify the below text:

Complex input: Career He studied with the English-born composer Henry Hugo Pierson in Stuttgart, and with William Sterndale Bennett and the pianist Edward Dannreuther in London.

Simplified output: He studied with Henry Hugo Pierson as well as William Sterndale Bennett and Edward Dannreuther.

--separator--