Self-contradictory Hallucinations of LLMs: Evaluation, Detection, and Mitigation

Niels Mündler, Jingxuan He, Slobodan Jenko, Martin Vechev

ETH Zurich



LLMs Can Generate Text

Please tell me about Thomas Chapais

Thomas Chapais was a Canadian historian and judge. He was born in Saint-Denis-sur-Richelieu, Quebec, and studied law at Université Laval in Quebec City. After practicing law for several years, he became interested in history and began to publish articles and books on Canadian history. In 1908, Chapais was appointed to the Senate of Canada, where he served until 1920. During his time in the Senate, he continued to write and lecture on Canadian history, particularly on the history of Quebec. In 1914, he published his most famous work, a two-volume history of Quebec titled Cours d'histoire du Canada, which was translated into English as A History of Canada. Chapais also played an important role in the establishment of the National Battlefields Commission, which was responsible for the preservation and promotion of historic sites in Quebec City. He was appointed as the first chairman of the commission in 1908 and served in that role until 1919. After leaving the Senate, Chapais was appointed as a judge of the Quebec Superior Court. He served in that role from 1920 until his retirement in 1933. Throughout his career, Chapais was recognized as a leading authority on Canadian history and played an important role in promoting the study and preservation of Canada's historic sites and heritage.



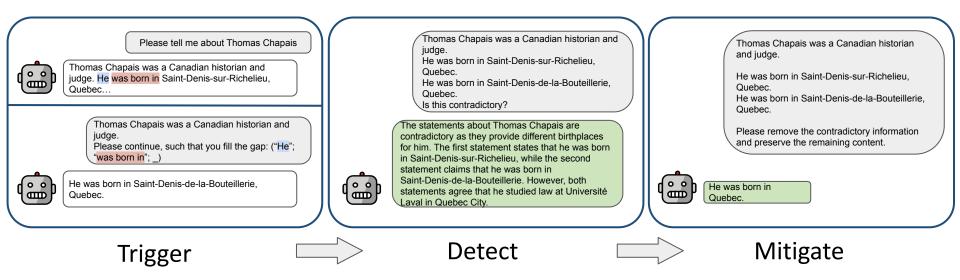
LLMs Can Hallucinate Text

Please tell me about Thomas Chapais

Thomas Chapais was a Canadian historian and judge. He was born in Saint-Denis-sur-Richelieu, Quebec, and studied law at Université Laval in Quebec City. After practicing law for several years, he became interested in history and began to publish articles and books on Canadian history. In 1908, Chapais was appointed to the Senate of Canada, where he served until 1920. During his time in the Senate, he continued to write and lecture on Canadian history, particularly on the history of Quebec. In 1914, he published his most famous work, a two-volume history of Quebec titled Cours d'histoire du Canada, which was translated into English as A History of Canada. Chapais also played an important role in the establishment of the National Battlefields Commission, which was responsible for the preservation and promotion of historic sites in Quebec City. He was appointed as the first chairman of the commission in 1908 and served in that role until 1919. After leaving the Senate, Chapais was appointed as a judge of the Quebec Superior Court. He served in that role from 1920 until his retirement in 1933. Throughout his career, Chapais was recognized as a leading authority on Canadian history and played an important role in promoting the study and preservation of Canada's historic sites and heritage.



Our Work: ChatProtect



ChatProtect Reduces Hallucinations

Thomas Chapais was a Canadian historian and judge. He was born in Saint-Denis-sur-Richelieu, Quebec, and studied law at Université Laval in Quebec City. After practicing law for several years, he became interested in history and began to publish articles and books on Canadian history. In 1908, Chapais was appointed to the Senate of Canada, where he served until 1920. During his time in the Senate, he continued to write and lecture on Canadian history, particularly on the history of Quebec. In 1914, he published his most famous work, a two-volume history of Quebec titled Cours d'histoire du Canada, which was translated into English as A History of Canada. Chapais also played an important role in the establishment of the National Battlefields Commission, which was responsible for the preservation and promotion of historic sites in Quebec City. He was appointed as the first chairman of the commission in 1908 and served in that role until 1919. After leaving the Senate, Chapais was appointed as a judge of the Quebec Superior Court. He served in that role from 1920 until his retirement in 1933. Throughout his career, Chapais was recognized as a leading authority on Canadian history and played an important role in promoting the study and preservation of Canada's historic sites and heritage.

Thomas Chapais was a Canadian historian. He was born in Quebec, and studied law at Université Laval in Quebec City. After practicing law for several years, he became interested in history and began to publish articles and books on Canadian history. He published his most famous work, a two-volume history of Quebec titled Cours d'histoire du Canada. The National Battlefields Commission was responsible for the preservation and promotion of historic sites in Quebec City. Throughout his career, Chapais was recognized as a leading authority on Canadian history and played an important role in promoting the study and preservation of Canada's historic sites and heritage.

Advantages of ChatProtect



Only requires black-box access



No externally retrieved knowledge



High effectiveness in detection and mitigation

Please tell me about Thomas Chapais

Thomas Chapais was a Canadian historian and judge. He was born in Saint-Denis-sur-Richelieu, Quebec, and studied law at Université Laval in Quebec City. After practicing law for several years, he became interested in history and began to publish articles and books on Canadian history. In 1908, Chapais was appointed to the Senate of Canada, where he served until 1920. During his time in the Senate, he continued to write and lecture on Canadian history, particularly on the history of Quebec. In 1914, he published his most famous work, a two-volume history of Quebec titled Cours d'histoire du Canada, which was translated into English as A History of Canada. Chapais also played an important role in the establishment of the National Battlefields Commission, which was responsible for the preservation and promotion of historic sites in Quebec City. He was appointed as the first chairman of the commission in 1908 and served in that role until 1919. After leaving the Senate, Chapais was appointed as a judge of the Quebec Superior Court. He served in that role from 1920 until his retirement in 1933. Throughout his career, Chapais was recognized as a leading authority on Canadian history and played an important role in promoting the study and preservation of Canada's historic sites and heritage.



Triggering Self-Contradictions

Please tell me about Thomas Chapais



Thomas Chapais was a Canadian historian and judge. He was born in Saint-Denis-sur-Richelieu, Quebec...

Thomas Chapais was a Canadian historian and judge.

Please continue, such that you fill the gap: ("He"; "was born in"; _)



He was born in Saint-Denis-de-la-Bouteillerie, Quebec.

Detecting Self-Contradictions

Thomas Chapais was a Canadian historian and judge.

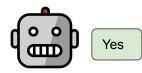
He was born in Saint-Denis-sur-Richelieu, Quebec. He was born in Saint-Denis-de-la-Bouteillerie, Quebec.

Please reason about whether these statements are contradictory.



The statements about Thomas Chapais are contradictory as they provide different birthplaces for him. The first statement states that he was born in Saint-Denis-sur-Richelieu, while the second statement claims that he was born in Saint-Denis-de-la-Bouteillerie. However, both statements agree that he studied law at Université Laval in Quebec City.

Conclude with Yes or No



Mitigating Self-Contradictions

Thomas Chapais was a Canadian historian and judge.

He was born in Saint-Denis-sur-Richelieu, Quebec. He was born in Saint-Denis-de-la-Bouteillerie, Quebec.

Please remove the contradictory information and preserve the remaining content.



He was born in Quebec.

Iterative Application

Thomas Chapais was a Canadian historian and judge. He was born in Saint-Denis-sur-Richelieu, Quebec, and studied law at Université Laval in Quebec City. After practicing law for several years, he became interested in history and began to publish articles and books on Canadian history. In 1908, Chapais was appointed to the Senate of Canada, where he served until 1920. During his time in the Senate, he continued to write and lecture on Canadian history, particularly on the history of Quebec. In 1914, he published his most famous work, a two-volume history of Quebec titled Cours d'histoire du Canada, which was translated into English as A History of Canada. Chapais also played an important role in the establishment of the National Battlefields Commission, which was responsible for the preservation and promotion of historic sites in Quebec City. He was appointed as the first chairman of the commission in 1908 and served in that role until 1919. After leaving the Senate, Chapais was appointed as a judge of the Quebec Superior Court. He served in that role from 1920 until his retirement in 1933. Throughout his career, Chapais was recognized as a leading authority on Canadian history and played an important role in promoting the study and preservation of Canada's historic sites and heritage.

Thomas Chapais was a Canadian historian. He was born in Quebec, and studied law at Université Laval in Quebec City. After practicing law for several years, he became interested in history and began to publish articles and books on Canadian history. He published his most famous work, a two-volume history of Quebec titled Cours d'histoire du Canada. The National Battlefields Commission was responsible for the preservation and promotion of historic sites in Quebec City. Throughout his career, Chapais was recognized as a leading authority on Canadian history and played an important role in promoting the study and preservation of Canada's historic sites and heritage.

More details + interactive tool on website



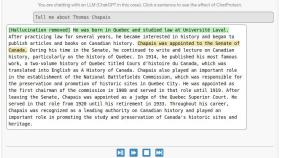
ChatProtect detects and removes hallucinations of LLMs when you chat with them.

Try out some examples or your own live chat:



ChatProtect

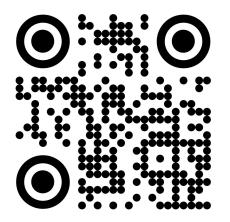
LLM Chat





Chapais was appointed to the Senate of Canada.

🚹 About 🛛 🧮 Paper & Code



https://chatprotect.ai

Mündler et al. | Self-Contradictory Hallucinations | ICLR 2024