

# Intelligent title answering system to combat click-bait culture

Proposal for customized master thesis  
Markus Sverdvik Heiervang

---

Clickbait is a popular method used by modern media to generate revenue. Typically, news articles are titled in a way that hides an important piece of information from the reader to bait them into reading the article, when their only goal is to find the hidden piece of information.

Examples of clickbait titles:

“Here’s the reason why ...”, “She got fired for this”

“One thing you didn’t know about ...”

A trend commonly referred to as “saved you a click”, or #SavedYouAClick has arose in response to this inflation of clickbait, where one reader reads the article and finds the piece of information that the title hides, then immediately reveals it in e.g. the comment section so people can find it without having to read the entire article.

## The goal of this thesis is

to automate this “saved you a click” trend using NLP. To create system that reads a news article and a headline, and outputs the missing information that the title is hiding, and possibly even embed it in a browser extension (either through web api, or running on client side).

Example:

Title: What the 'Someone Is Typing' Bubbles in Messaging Apps Actually Mean

Answer: It’s an indicator someone is typing

Title: The Reason Cops Touch Your Car’s Taillight When Pulling You Over

Answer: To leave fingerprints, as proof that they pulled you over in case you decide to flee

Title: BREAKING: Has Cristiano Ronaldo raped someone?

Answer: No.

These are actual news headlines, taken from the top posts of the reddit page:  
<https://www.reddit.com/r/savedyouaclick/top/?t=all>

The plan is to implement such a solution for primarily English text data, but possibly also experimenting with a solution for Norwegian

## **Methodology**

Application and experimentation with state-of-the-art techniques in data-driven question answering systems and text summarization models.

## **Curriculum and Milestones**

Autumn 2020 (30 credits):

Complete courses

IN4080: Natural Language Processing

IN-STK5000: Adaptive Methods for data based decision making

TEK5040: Deep learning for autonomous systems

Spring 2021 (40 credits):

Complete courses

IN4030: Introduction to bioinformatics

IN4200: High performance computing and numerical projects

IN5550: Neural methods in natural language processing

Write and complete essay

Autumn 2021 (20 credits):

Thesis introduction, data collection and analysis, experimentation

Spring 2022 (30 credits):

Complete implementation and finish written thesis.