

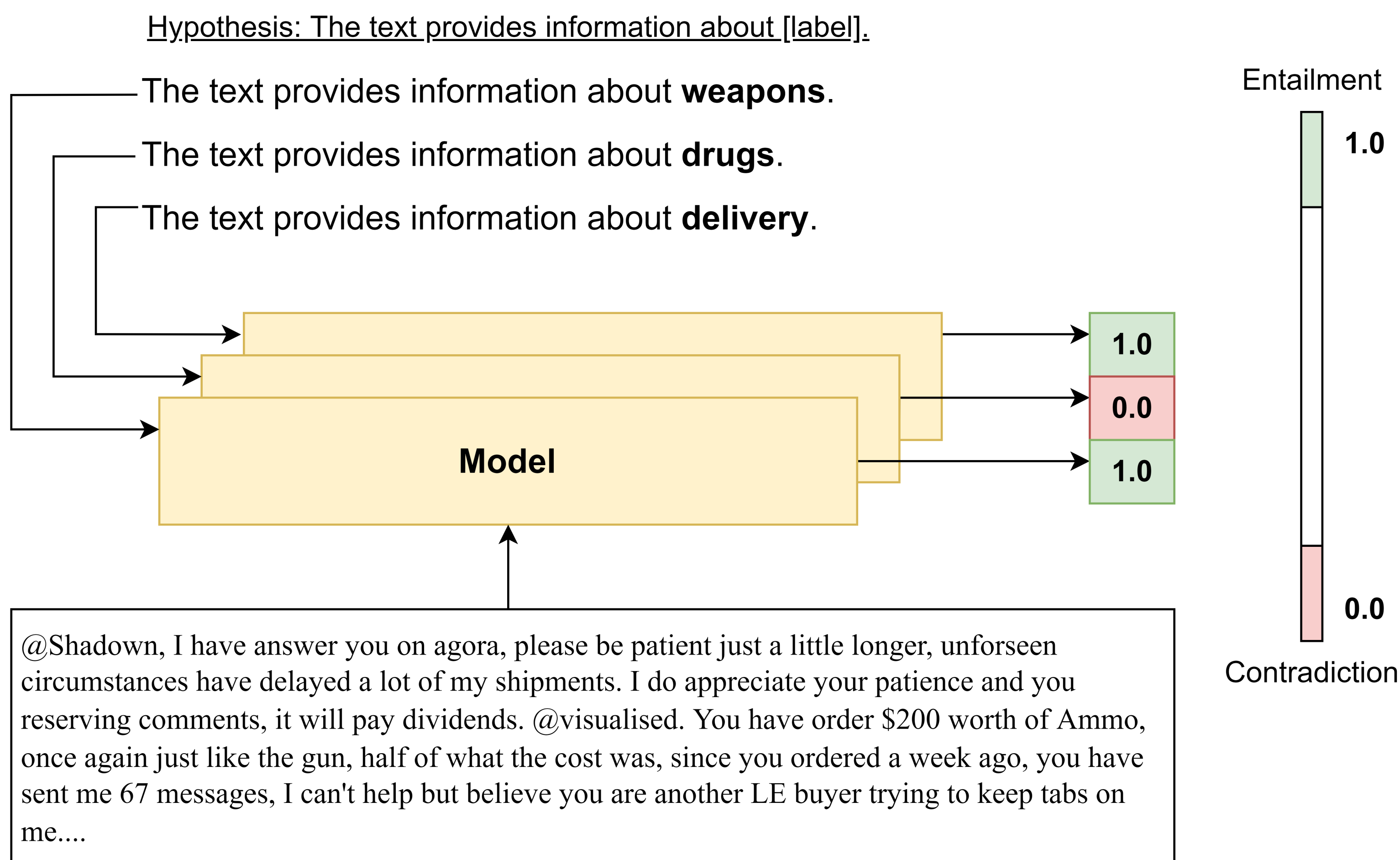
Automatic Webpage Content Categorisation and Extraction

Author: Michal Rein

Supervisor: Ing. Daniel Dolejška



Textual Entailment



The Stanford Natural Language Inference (SNLI) Corpus

Premise:

Two doctors perform surgery on patient.

Hypothesis 1:

Doctors are performing surgery.
Label: **Entailment**

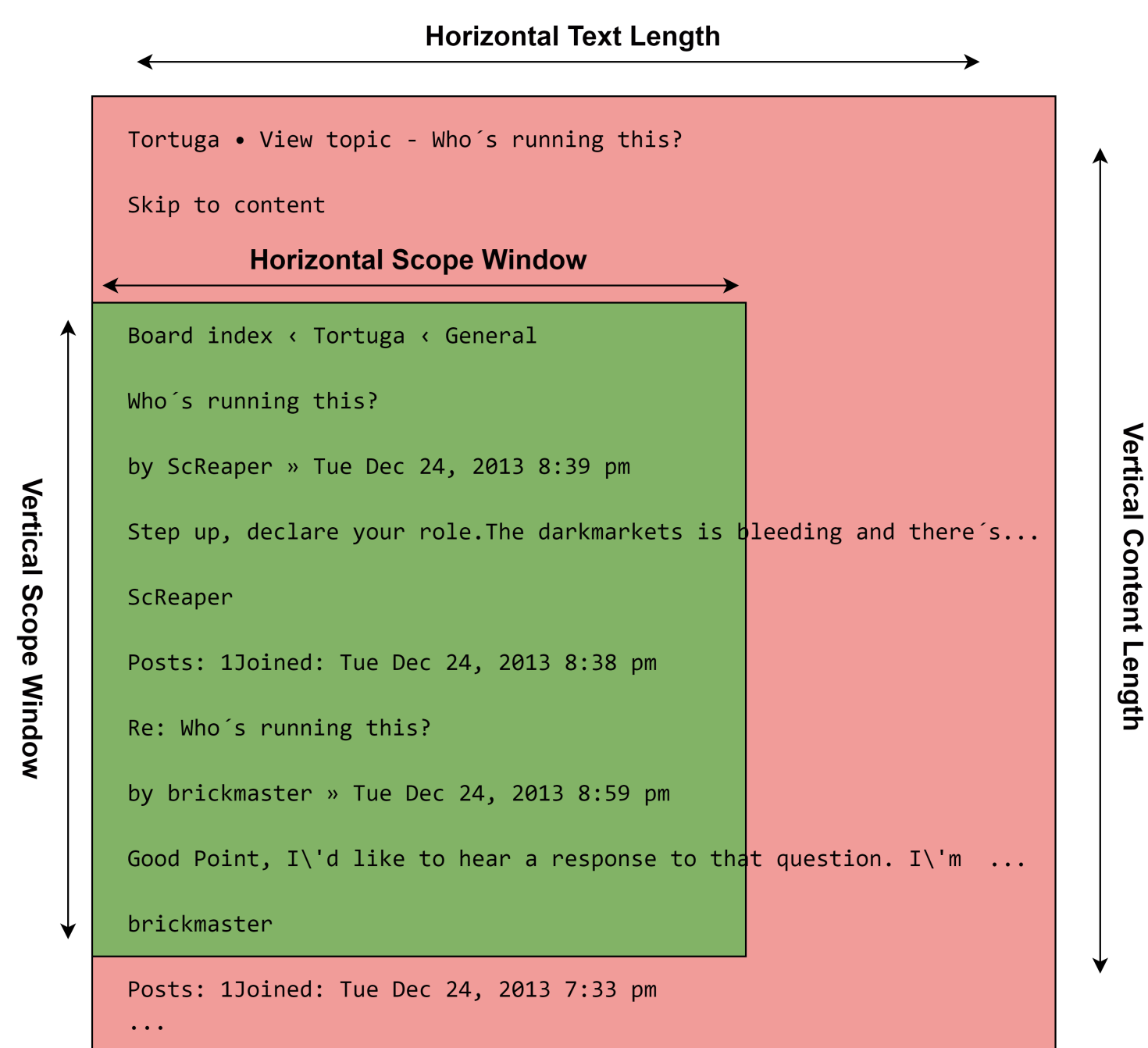
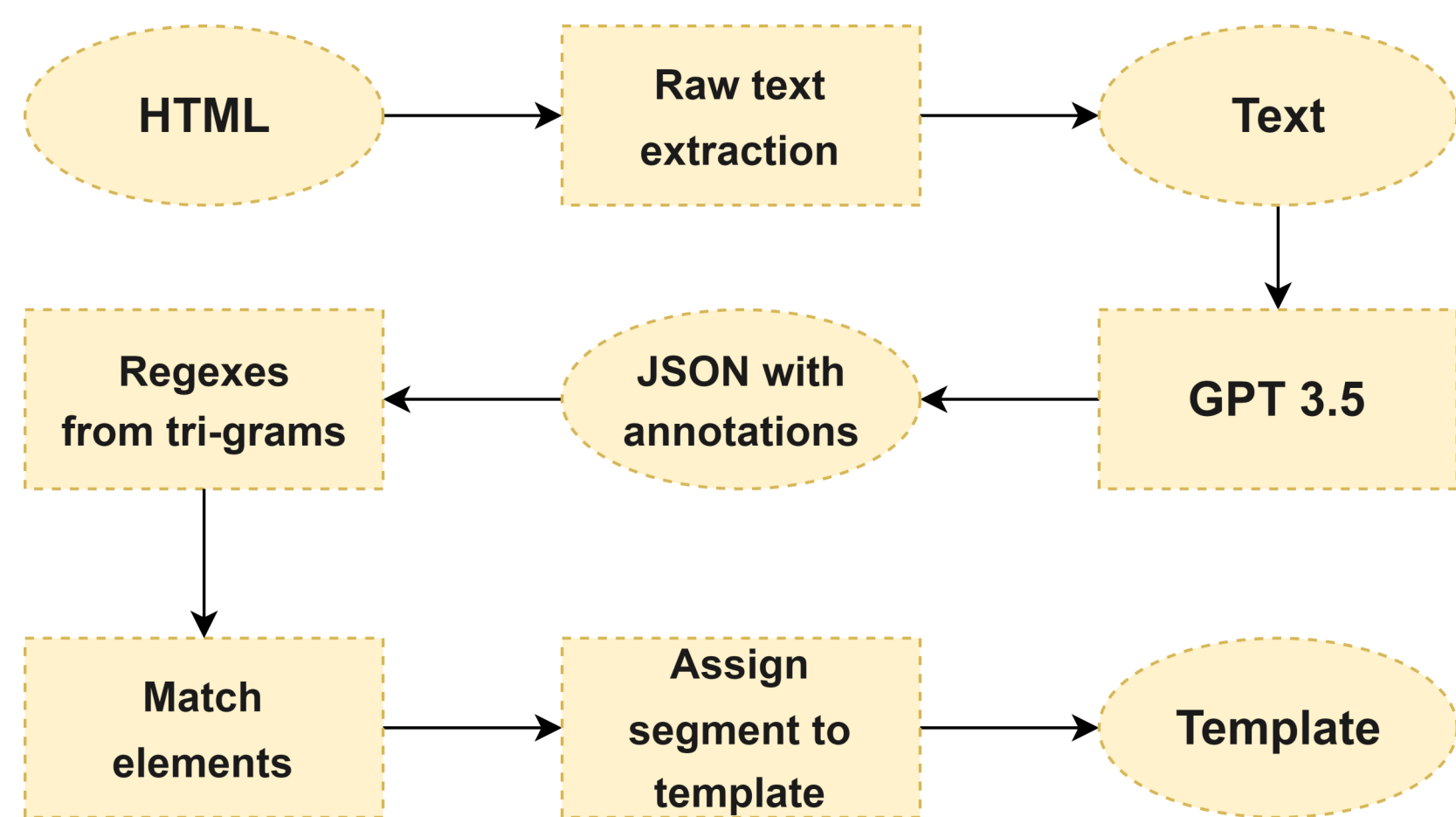
Hypothesis 2:

Two doctors are performing surgery on a man.
Label: **Neutral**

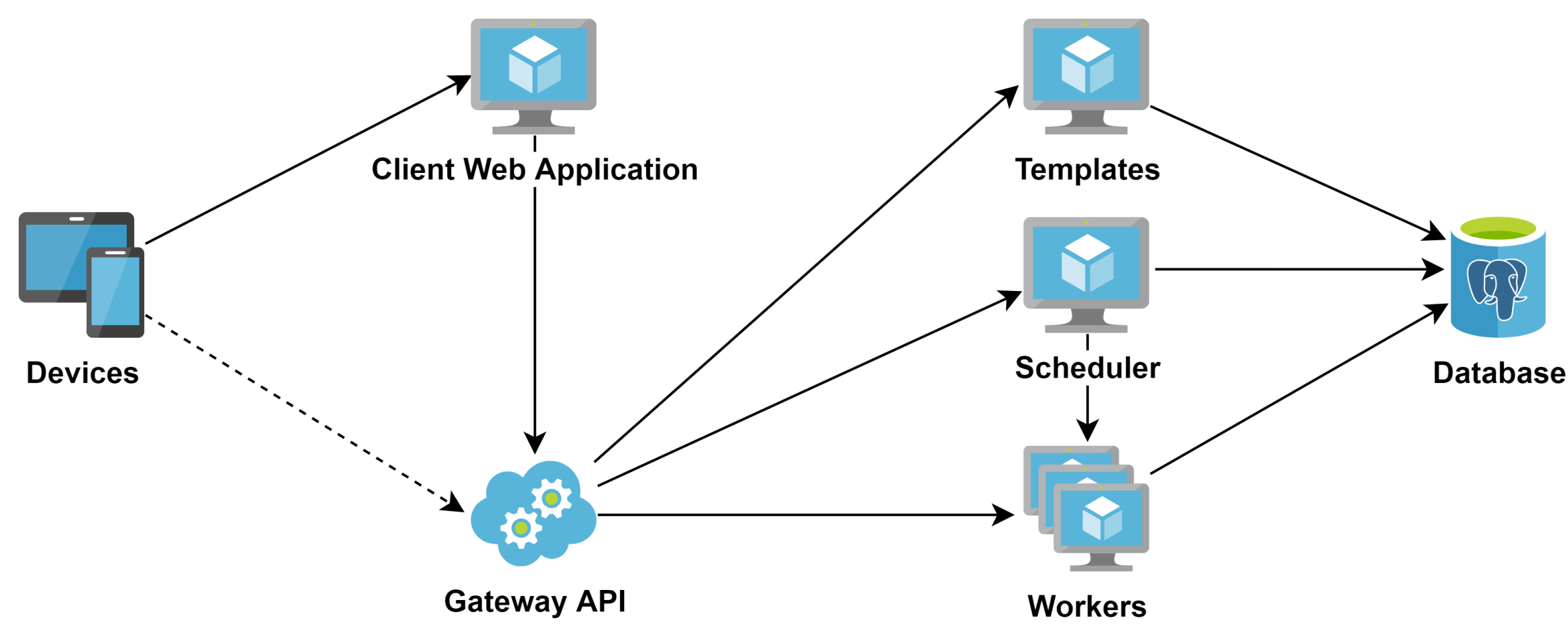
Hypothesis 3:

Two surgeons are having lunch.
Label: **Contradiction**

Templating



System



Processing Pipeline

