

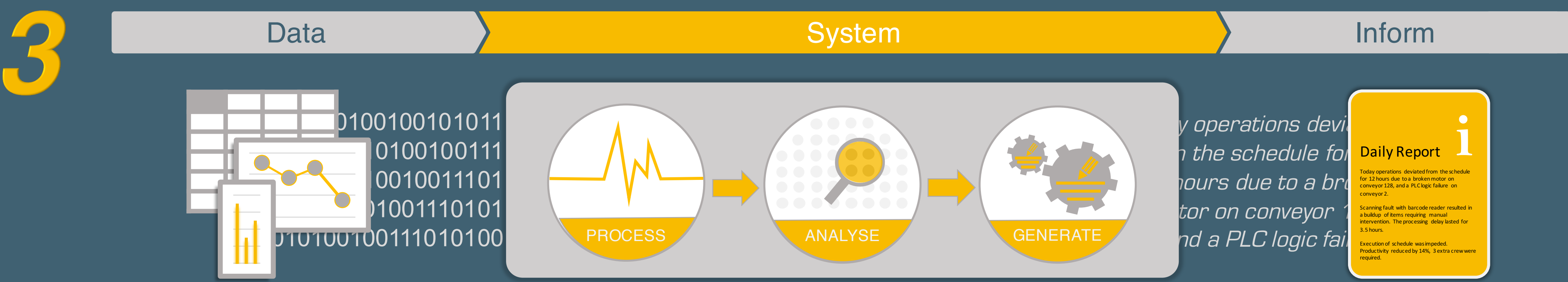
Automated Data Contextualisation for Decision Support

Author: Thomas McKeon
Supervisors: Tim French, Melinda Hodkiewicz, Wei Liu

1 Dispatch operators have to synthesise multiple streams of data to form a mental model of site operations. This process requires a significant cognitive load on the operator. Operators also write routine reports summarising daily events which can contain unconscious bias. Ideally, data would be the source used for reporting the causes and effects of unplanned events.

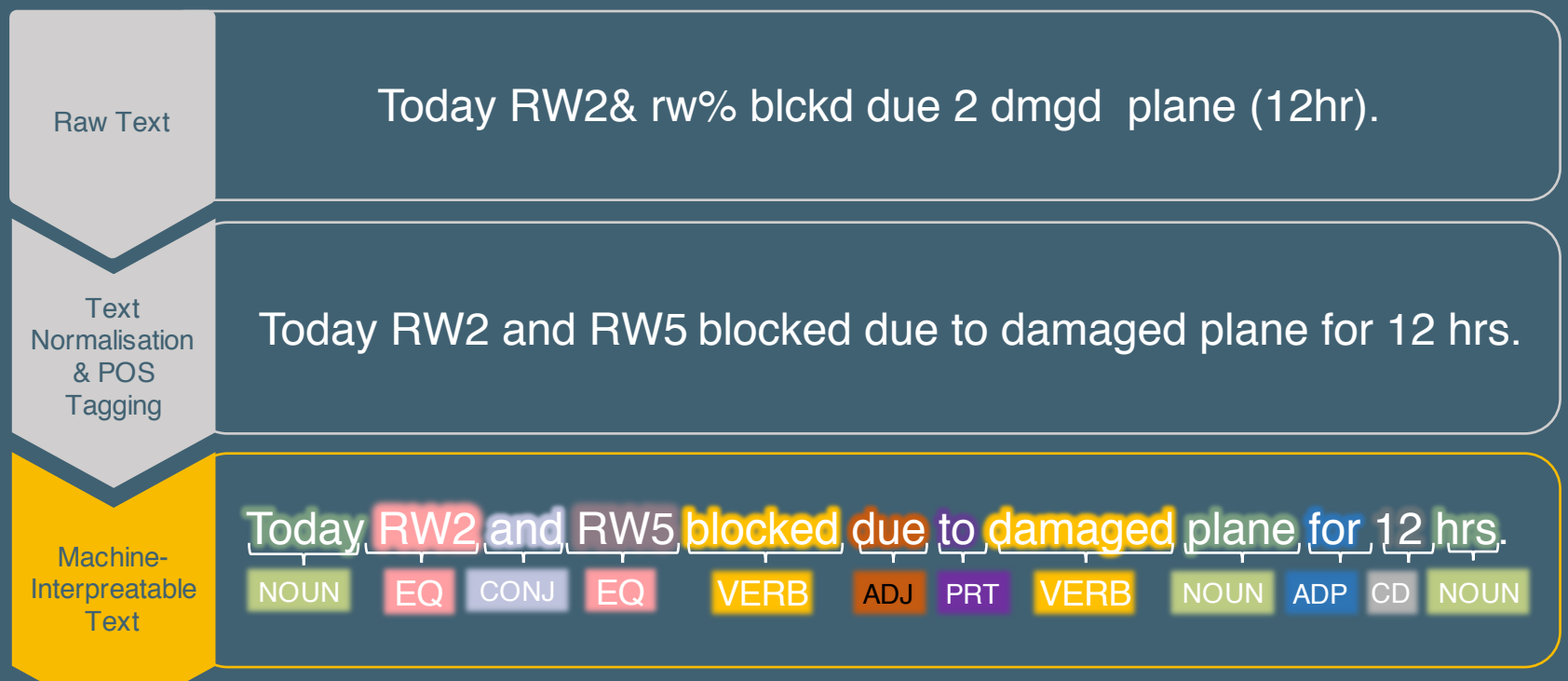


2 Aim to generate fast, accurate and unbiased textual summaries in a natural language from statistical data to support operators and management.



The system is comprised of three modules: a Text Processing Module used to clean the raw company text into a machine-interpretable form and form a domain-specific vocabulary; a Template Generation Module used to analyse word sequences and produce blank templates; and a Language Generation Module that uses the results from the prior modules to produce data-driven reports.

4 **Text Processing**
Text is cleaned and standardised into a machine-interpretable form for analysis through a series of tasks including: text normalisation and part-of-speech tagging. This process creates and makes use of the Domain Vocabulary.
82% Reduction in noisy text



Domain Vocabulary

The domain vocabulary is created using rule based extraction over the domain corpus.

For example...

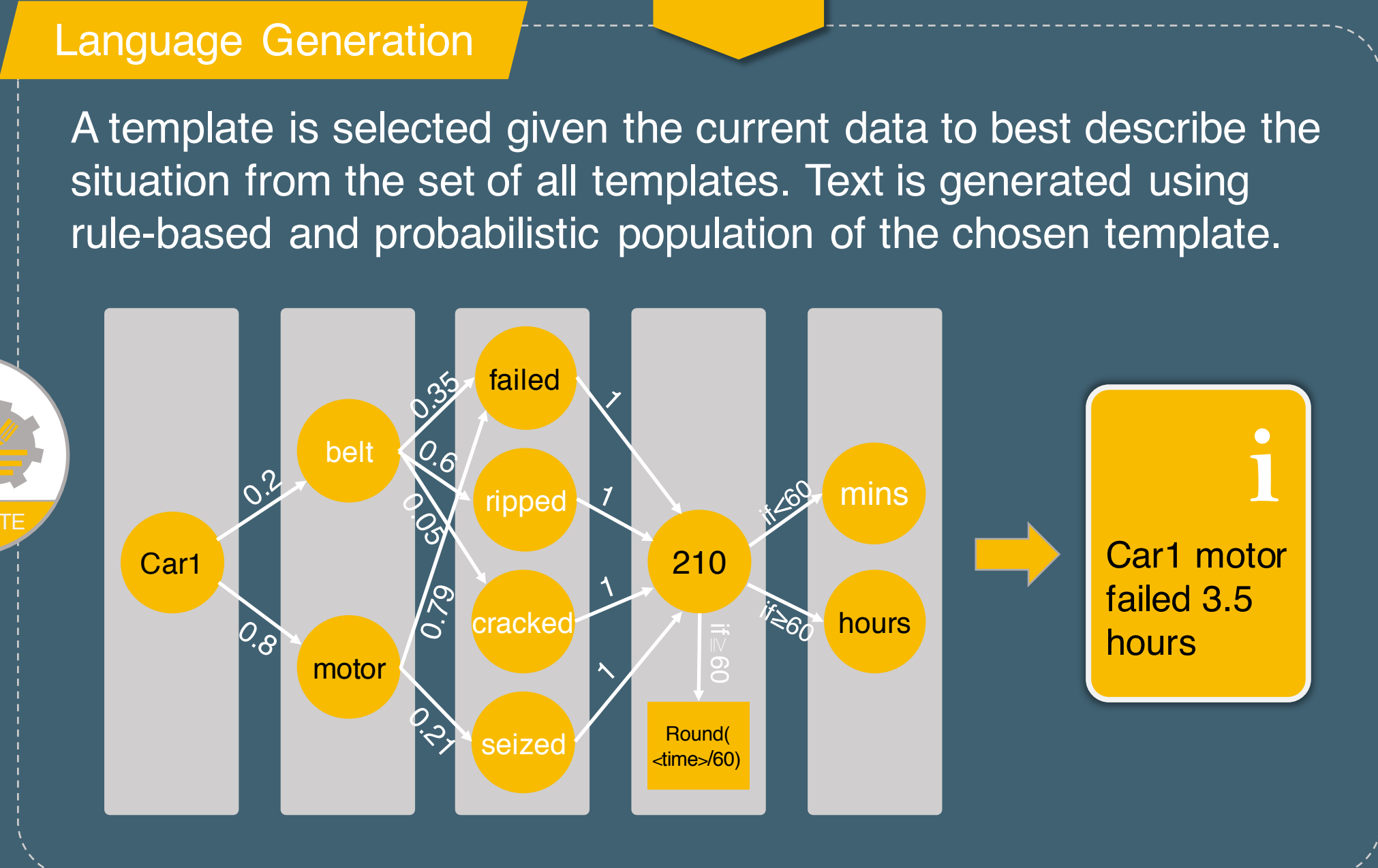
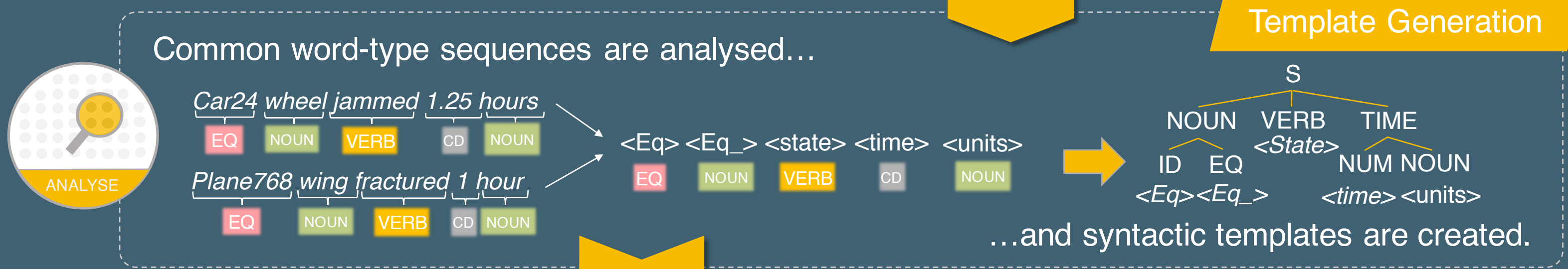
Today RW2& rw% blocked due 2 damaged plane.

Regular Expression: $\backslash(?!\backslash D)\backslash N$
[a-zA-Z]{2,4}
[0-9%#\$*]\

RW2, rw%

The extracted instances are stored in one of the five created dictionaries.

- Asset ID List
- Noun List
- Acronym List
- Abbreviation List
- Standard Dictionary



5 The automated reporting system reduces the cognitive load on operators as events are presented in a natural language. The system improves reporting by using data as the source of knowledge and removes bias. It can also scale, writing reports for management who are not directly involved in the detailed operations. Productivity is increased as no time is required to handwrite reports.

