Data to Text

Task and motivation

Technology

Applications and evaluation

WHAT IS DATA TO TEXT

The situation

Massive numerical data produced by machines

Quasi-contiuous sensor data over time lines

Selection of interesting data constellations

System (Arria) gives a natural language summary

Motivation

Human inspection of data tedious

Experts are discharged of routine tasks, concentrate on assessment

Text summaries are an alternative, may be superior to visual display of data

Application areas

Medicine ("Baby talk")

Meteorology

Geology

https://www.arria.com

A CONSENSUS ARCHITECTURE

Signal analysis

Analysing numerical and other data

(e.g., measures of body functions, medication treatments) for patterns

Looking for trends

Data interpretation

Identifying more complex (domain-specific) messages from patterns and trends

Identifying causal and other relations between messages

Document planning

Deciding which messages should be mentioned

Creating a document and rhetorical structure around messages

Microplanning and realization

Building an actual text for the document plan

SIGNAL ANALYSIS

Functionality

Building symbolic expressions out numerical data

Standard signal analysis algorithms for specific pattens

Examples: Short-term changes (spikes, steps)

Long-term changes (values in- or decreasing over time)

Artifacts (corrupted data)

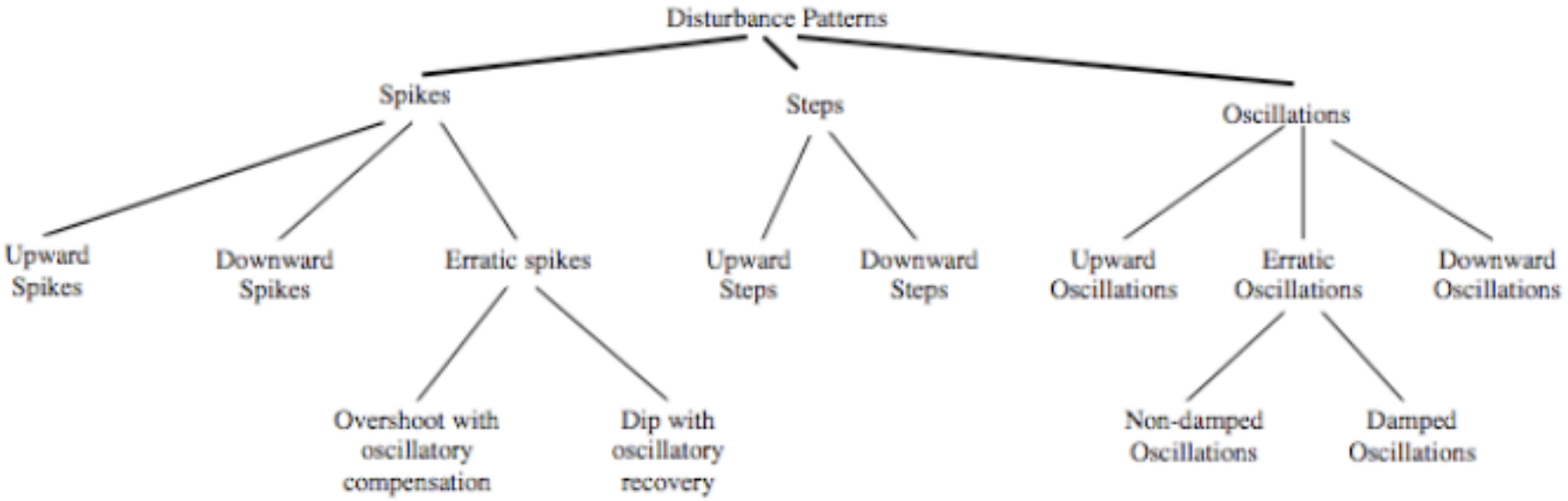
System developer capabilities required

Understand the (standard) signal analysis algorithms

Understand the domain well enough

Understand the way humans talk about the domain

PATTERN ONTOLOGY IN THE GAS TURBINE DOMAIN



DATA INTERPRETATION

Functionality

Map basic patterns and events into messages and relationships

typically used in communication about the domain

In some domains, basic pattens define level of communication

-> no data interpretation needed (e.g., marine forecast)

Tasks to be accomplished

Message building

e.g., "heart rate temporarily too low" for downward spikes below 100

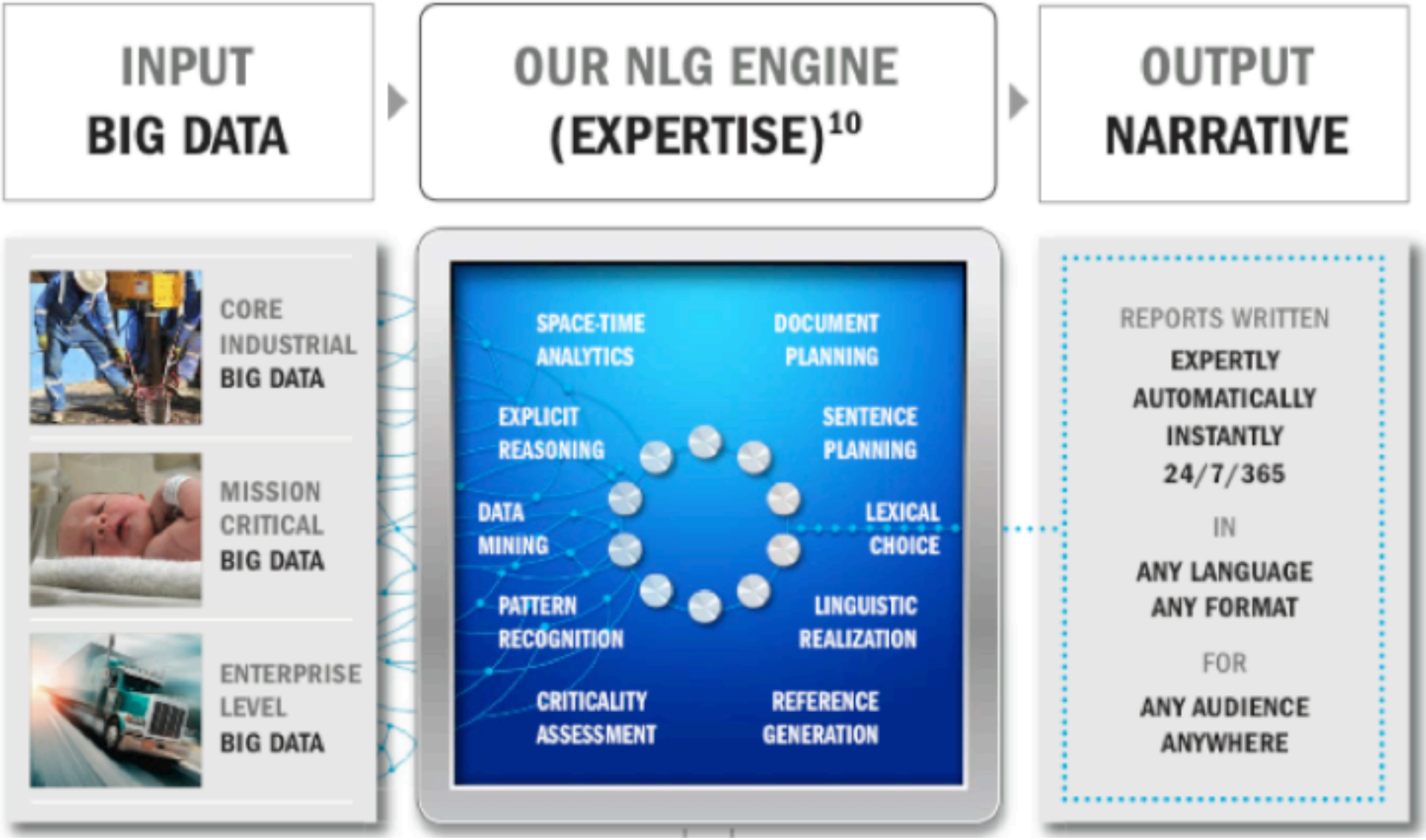
Assessing degrees of importance

e.g., depending on duration of a pattern

Inferring relations

e.g., causality as instances of (causal) domain rules

associations as temporal cooccurrence


INPUT
BIG DATA
OUR NLG ENGINE
(EXPERTISE)10
OUTPUT
NARRATIVE
CORE INDUSTRIAL BIG DATA
MISSION CRITICAL BIG DATA
ENTERPRISE LEVEL BIG DATA
SPACE-TIME ANALYTICS
DOCUMENT PLANNING
EXPLICIT REASONING
SENTENCE PLANNING
DATA MINING
LEXICAL CHOICE
PATTERN RECOGNITION
LINGUISTIC REALIZATION
CRITICALITY ASSESSMENT
REFERENCE GENERATION
REPORTS WRITTEN
EXPERTLY
AUTOMATICALLY
INSTANTLY
24/7/365
IN
ANY LANGUAGE
ANY FORMAT
FOR
ANY AUDIENCE
ANYWHERE

OUR NLG BIG DATA ANALYTICS

Space-Time Analytics
Document Planning Uses low level data to identify what objects exist and when they exist across space and time

Explicit Reasoning
Captures the logical skills of experts in the field, emulating their analytic capabilities

Data Mining
Finds the unexpected regularities and co-occurrences in the analyzed Big Data sets

Pattern Recognition
Detects the describable higher-level phenomena existing in the big data sets

Criticality Assessment
Works out the relative importance of the available information being analyzed

NATURAL LANGUAGE GENERATION

Document Planning
Works out the overall structure of the text being generated: what gets said, and in what order

Sentence Planning
Works out the correct amount of information to pack into each sentence being generated

Lexical Choice
Works out what words to use so the meaning is clear and the style appropriate to the reader

Linguistic Realization
Works out how to express the content of each sentence using the correct grammar

Reference Generation
Works out how entities are referred to so the reader knows what is being talked about

Knowledge Capture

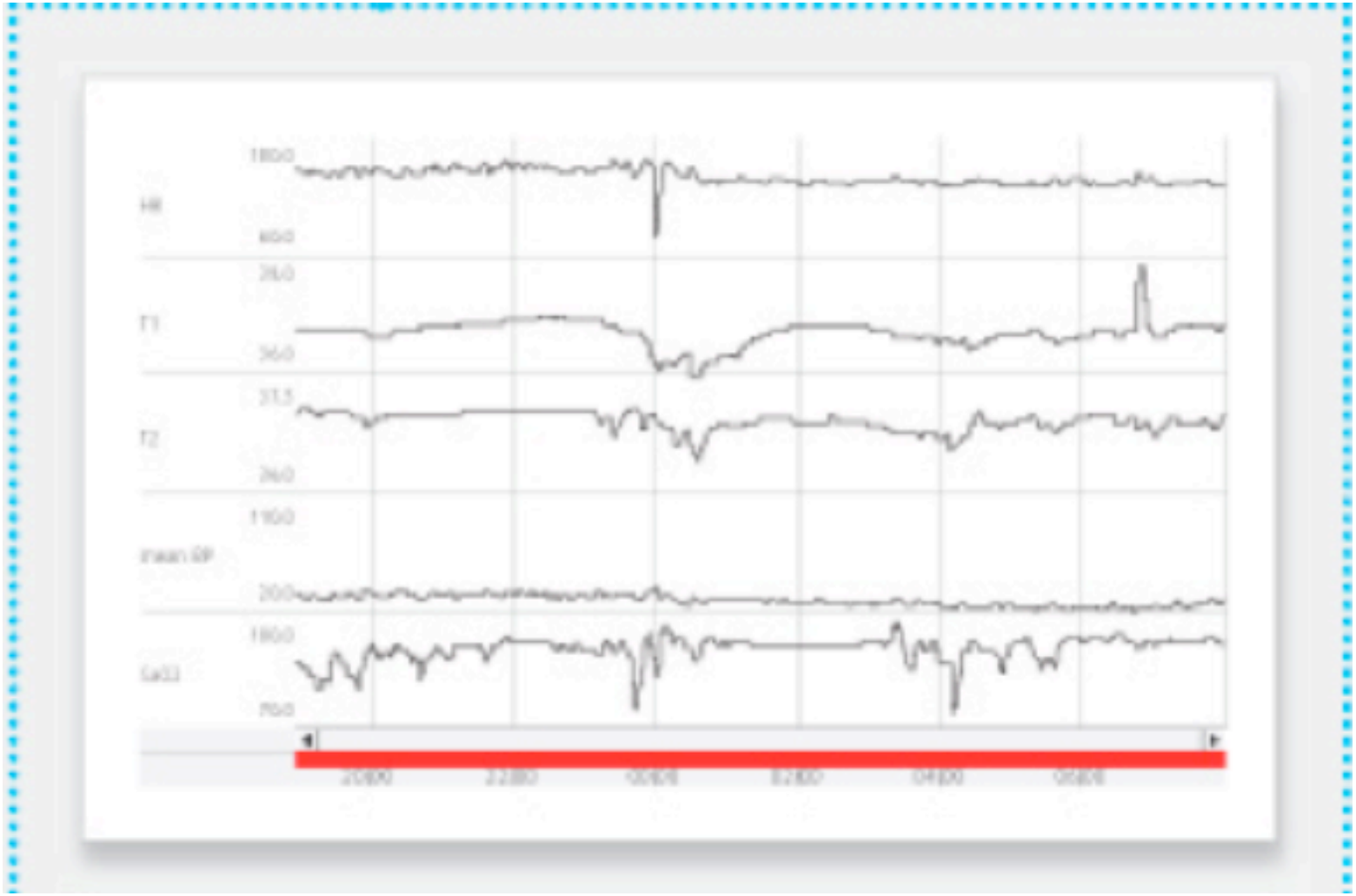
The Arria NLG Engine can be programmed to:

- incorporate an expert's knowledge as to what is important about the data;
- incorporate the reasoning processes which an expert would use in analysing the data;
- enable the "best practice" Knowledge of an organisation's most experienced experts to be permanently captured in the software, allowing the most efficient use of its scarce resources of expertise and avoiding knowledge loss when staff leave;
- improve quality control by standardising analytic and reporting practice;
- enable the knowledge of expert resources at the centre of an organisation to be distributed to the operational edges of the organisation; and
- reduce the effects of expert down-time.

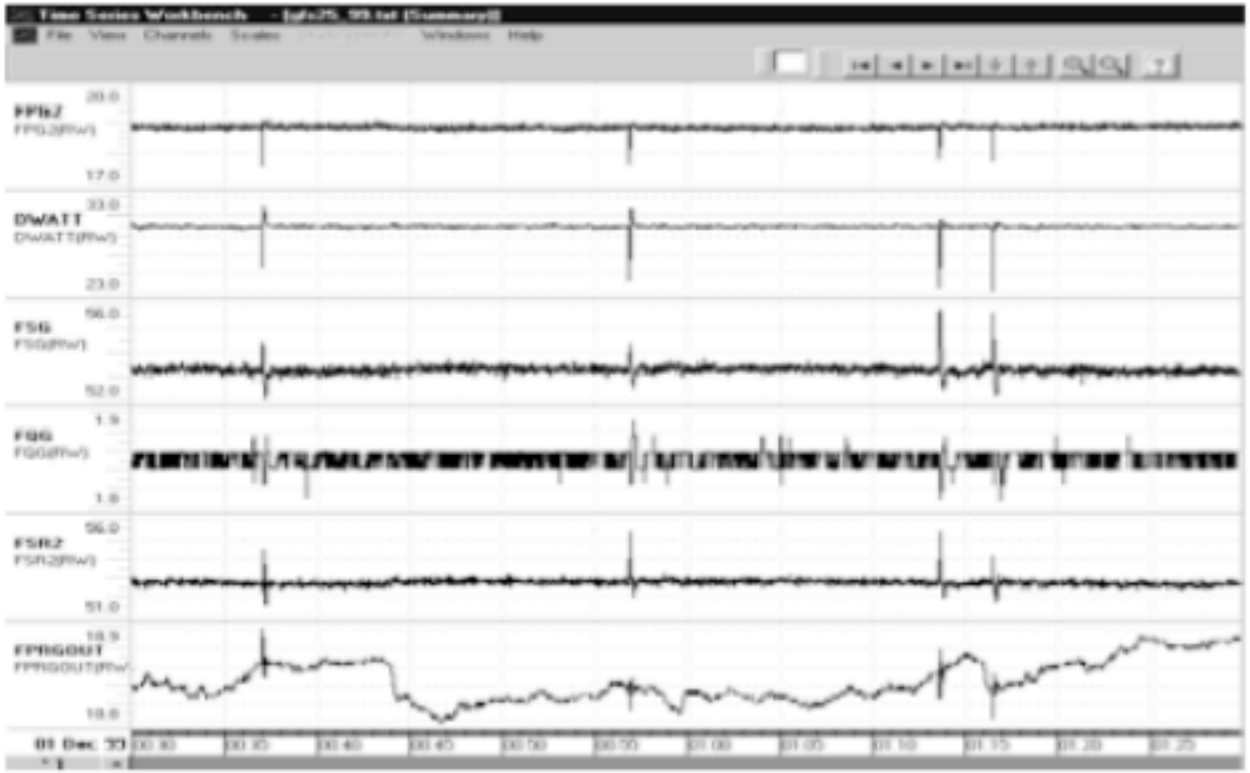
ARRIA - OVERALL ARCHITECTIRE

Problems with traditional generation systems

EXAMPLE BABY TALK



EXAMPLE GAS TURBINE



%FPG2 Intervalve gas fuel pressure input
%FSR2 Gas fuel stroke reference from fuel splitter
%FPRGOUT Gas ratio valve servo command
%DWATT Generator load
%FQG Gas fuel flow
%FSG Fuel stroke reference

Fig. 2. Visualisation of channels from one of the subsystems

SAMPLE BABYTALK NURSE REPORT

Respiratory Support: Current Status
Currently, the baby is on CMV in 27 % O2. Vent RR is 55 breaths per minute. Pressures are 20/4 cms H20. Tidal volume is 1.5. Sa02 is variable within the acceptable range and there have been some desaturations. The most recent blood gas was taken at around 07:45. Parameters are acceptable.

SAMPLE BABYTALK FAMILY REPORT

YOUR BABY, JOHN, is receiving intensive care at the Royal Infirmary of Edinburgh. He is being looked after in Blackford nursery in cot space five.

John is now 2 days old with a corrected gestation of 24 weeks and 2 days.

A SAMPLE GAS TURBINE REPORT (SYSTEM SUM-TIME-TURBINE)

[Background information]

Gas turbine: aylesford
Subsystem: exhaust temperature
Monitoring channels: TTXD-1, TTXD-2, TTXD-3, TTXD-4, TTXD-5 and TTXD-6
Turbine running state: part load
Time interval of these channels: from 12 to 15 on 27 Nov 99

[Overview information]

There were large erratic spikes in all channels at 12:59, 13:01, 13:41 and 14:40.

[Most significant patterns]

At 12:59, there were large erratic spikes in TTXD-1, TTXD-2, TTXD-3, TTXD-4, TTXD-5 and TTXD-6. These patterns violated the pairs and follows check. In more detail, there were dips with oscillatory recoveries in TTXD-3 and TTXD-4, followed 1s later by dips with oscillatory recoveries in TTXD-1, TTXD-2, TTXD-5 and TTXD-6. This occurred between 12:59:17 and 12:59:54.

A COMPARATIVE EVALUATION

	J. Doctor	J. Nurse	S. Doctor	S. Nurse	Overall score	Mean time
G	.37 (.15)	.40 (.19)	.40 (.16)	.44 (.09)	.40 (.15)	73.16
H	.42 (.11)	.48 (.10)	.44 (.10)	.47 (.12)	.45 (.10)	77.23
C	.44 (.16)	.36 (.10)	.38 (.12)	.47 (.10)	.41 (.13)	78.81

Table 1

Mean decision-making performance score and standard deviations per group and overall, with reaction times in seconds

Subject groups **(35 subjects in total, each subclass represented equally)**

senior (s)/junior (j) doctor or nurse

Tested versions

visual graphics (G), human generated (H), computer generated (C)

OVERALL RESULT

Performance in decision making based on report

Computer-generated text as effective as visual display

Human texts are superior to the other versions

-> NL text is a better modality than visual display

-> Computers are competitive, but still require improvement

Some sources for improvement

No context in on assessment scores which content selection is based on

Additional adaptation to target action (communicative purpose) possible

Temporal order to always conveyed unambiguously

(better sentence planning required)

Incorporation of meta-knowledge (gas turbime domain)