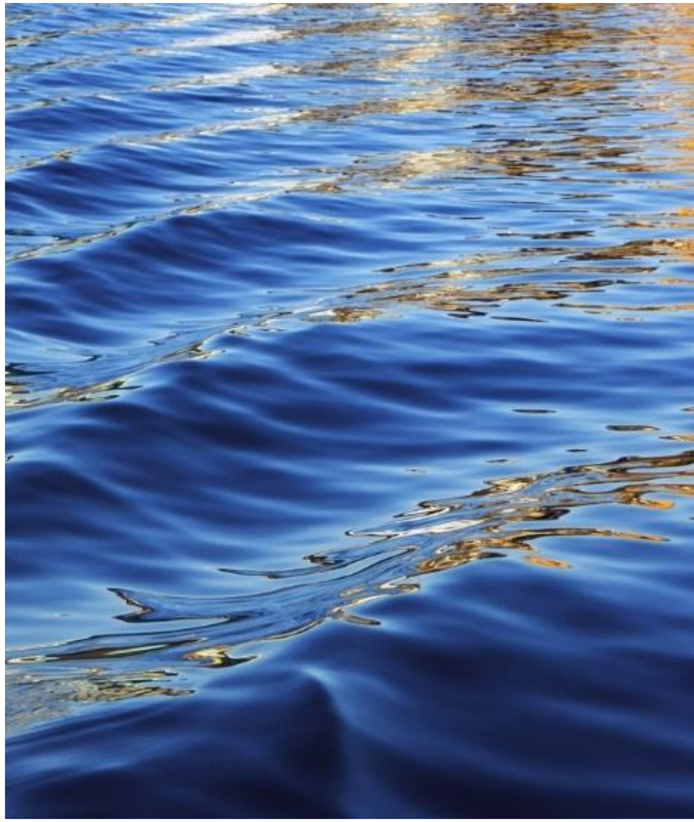
Big Data

Implementation Guide

Contents

- Big Data
- Implementation: engineering requirements
- Implementation: general requirements

1. Big Data



1. Big Data

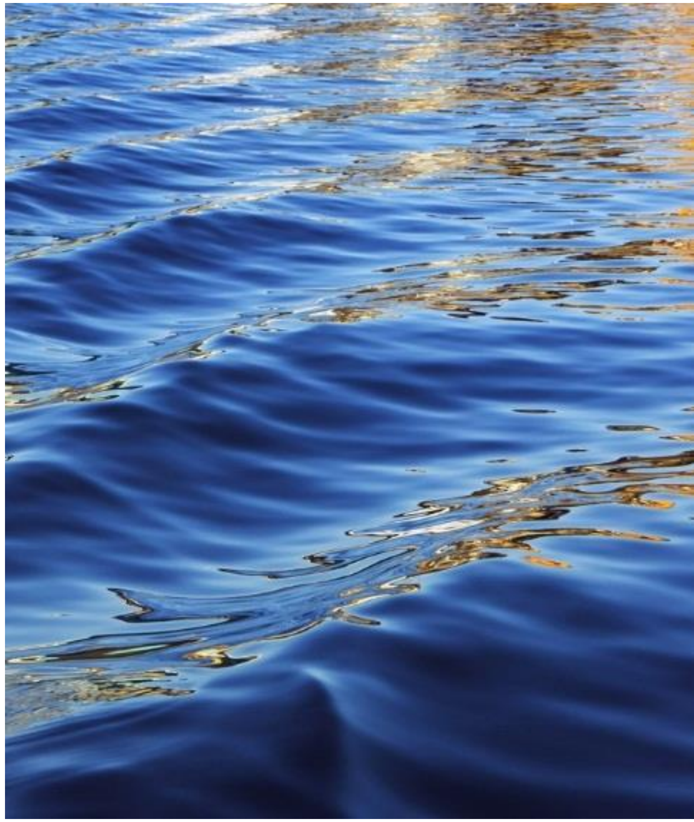
Big Data involves the automated collection and automated analysis of large volumes of data. A well-known early example is the automated collection and analysis of data through loyalty cards etc., in the retail sector

Big Data encompasses the addition of micro-electronic devices to components and assemblies that go into logistics vehicles including trucks and ships.

The addition of micro-electronic devices enables the Internet of Things, which can facilitate remote monitoring and reporting of performance in use.

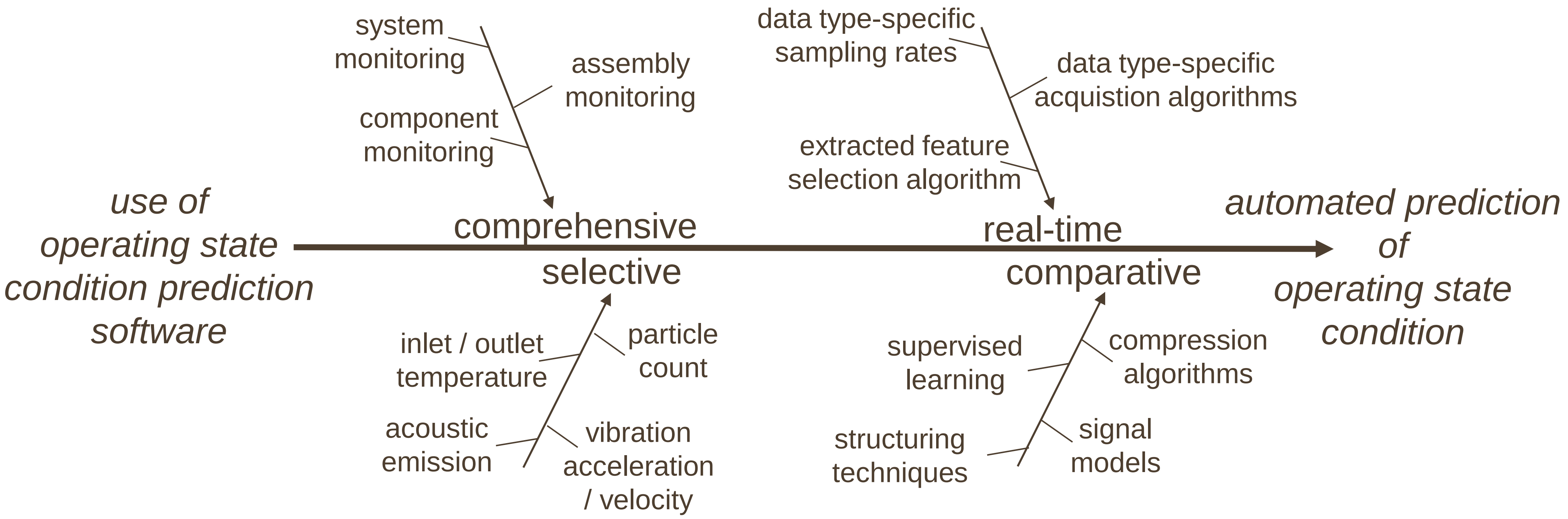
For example, Big Data for logistics vehicles can include automated prediction of vehicles' operating state condition, which can inform maintenance that prevents vehicle break-downs.

2. Implementation:
engineering requirements



2. Implementation: engineering requirements

The diagram below shows engineeringiii requirements for a Big Data implementation for the example of automated prediction of a logistics vehicle's operating state condition



2. Implementation: comprehensive data

Big Data is comprehensive data when it is collected from components, assemblies, and the entire logistic vehicle's system.

Comprehensive monitoring can generate huge amounts of data. For example, large industrial engines for ocean going vessels generates massive volumes of data.

However, the volume of data can be managed by determining the smallest number of component, assembly, and system positions that can be combined to provide comprehensive data about operating state condition.

2. Implementation: selective data

Data variety can be managed by carefully planning what types of data will be collected from what components, assembly, and system positions.

The selection of positions for data collection should be based on the significance of component, assembly, and system positions for the whole life cycle of a vehicle.

For example, vibration data are recommended for gear teeth fault detection. Acoustic emission sensors are used to detect the change in material structure. The particle count technique is also widely used in industrial applications, especially for gearbox condition monitoring. The cooling outlet temperature or exhausted gas temperature is a useful parameter to predict the working condition of cylinder liner for engines.

2. Implementation: real-time data

Collecting data in real-time involves data being collected as things happen – not retrospectively after they have happened.

Data sampling rates for prediction of operating state condition should be informed by the speed at which faults are likely to progress at different component, assembly, and system positions.

Specific data acquisition algorithms and feature extraction algorithms should be applied to acquire filtered data at pre-set sampling rates, which enable a fixed size amount of data to be collected at fixed intervals.

As data size can soon grow to terabytes, optimization algorithms are needed for real-time processing. These algorithms should have as low computational complexity as possible.

2. Implementation: comparative data

In order to enable automated prediction, incoming data needs to be structured to enable analysis and presented in comparison with nominal models of operating state condition.

Data mining algorithms can be used to extract relational features from all gathered data.

Prediction analyses can encompass statistics, signal models or supervised learning.

In order to transfer final operating state condition information to remote places, the information needs to be compacted by compression algorithms.

2. Implementation: engineering requirements Q&A

Q) What are engineering requirements to enable:

Comprehensive data?

Selective data?

Real-time data?

Comparative data?

2. Implementation: engineering requirements Q&A

Q) What are engineering requirements to enable:

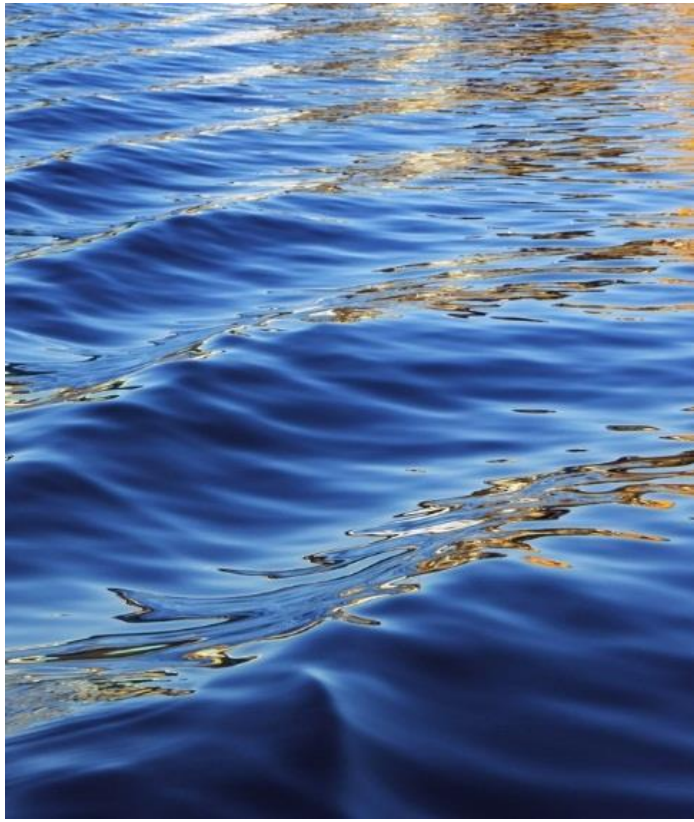
Q) Comprehensive data?
A) system monitoring, assembly monitoring, component monitoring

Q) Selective data?
A) acoustics, temperatures, vibration, particles

Q) Real-time data?
A) sampling rates, acquisition algorithms, extracted feature selection algorithms

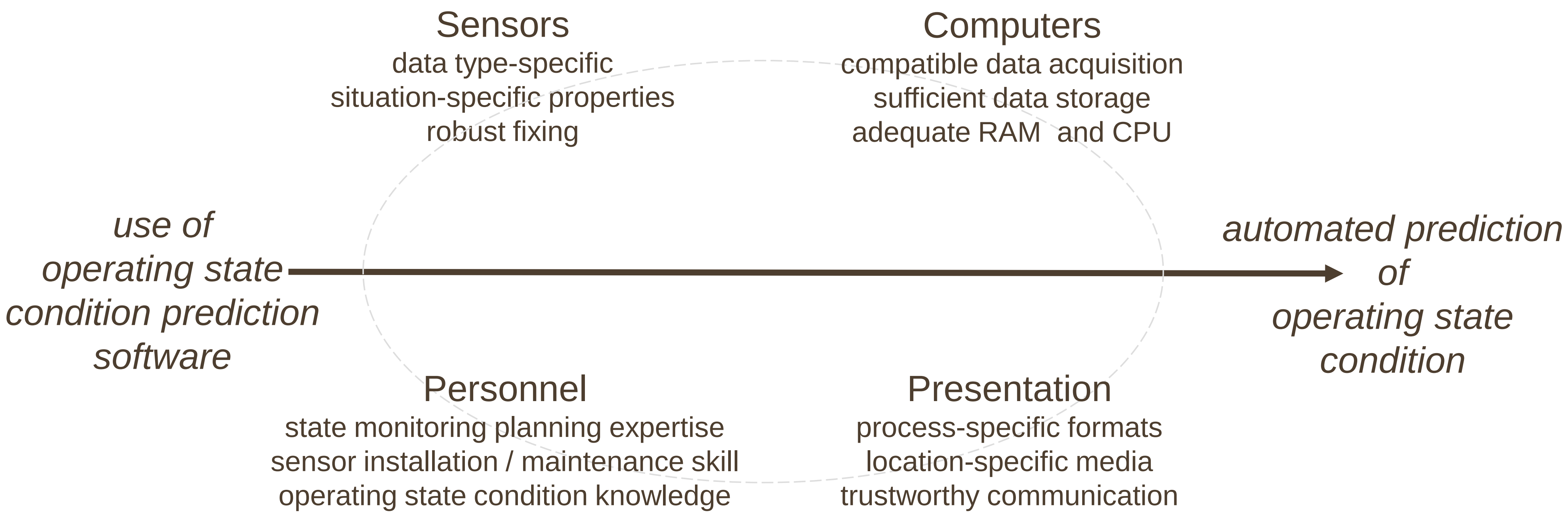
Q) Comparative data?
A) structuring techniques, signal models, supervised learning, compression algorithms

3. Implementation:
general requirements



3. Implementation: general requirements

The diagram below shows general requirements for a Big Data implementation



3. Implementation: sensors

Sensors need to be data type-specific. For example, accelerometers are needed for vibration acceleration data; thermometers are needed for temperature data.

Sensors need to have situation-specific properties. For example, low sensitivity accelerometers can obtain frequency range from 10Hz to 1Mhz. They are designed for capturing the vibration data from the system that has a wide frequency range. However, they are not sensitive for measuring the data that have low frequency range and low amplitudes. The high sensitivity accelerometers can capture very low frequency like 0.5 Hz with low amplitudes; however, they do not work well for high frequency range data.

Sensors must be fixed robustly in their selected positions. For example, if accelerometers become loose their data outputs can be misleading..

3. Implementation: computers

Data acquisition (DAQ) hardware devices provide interfaces between the signals from sensors and the computer hardware that will process and store the signals.

DAQ devices can be connected directly to the computer hardware or to an external breakout box. DAQ software is needed in order for the DAQ hardware to work with the computer hardware.

Sensors, DAQ hardware, DAQ software, and computer hardware all need to be compatible. When there is high volume of data inputs, high volume data storage is needed: especially as only some data are processed online and most of data are stored offline for later references and uses.

Adequate RAM capacity and CPU speed are essential for timely software operation, data processing, and displaying of analyses results.

3. Implementation: personnel

With regard to personnel, it is essential that planning expertise is available to formulate process models, signal models, etc. Then, based on these, determine what types of data are needed; where sensors should be located; and how often different types of data should be collected from which locations.

Subsequently, personnel are needed who have skill in the installation of sensors and DAQ devices.

Moreover, personnel are needed who have knowledge of the Big Data application: in the example, operation state condition during the use of ships engines. This knowledge is needed to facilitate appropriate decisions about, for example, what maintenance, repair, and/or replacement work needs to be done in what order in response to predictive data.

3. Implementation: presentation

It is important that the results of analyses are shown in relation to operating state condition process models. This can be done through the use of, for example, statistical process control charts.

Also, presentation should be made via media that are location-specific. Paper, for example, does not depend upon broadband internet access, or new immersive facilities such as headsets, at points-of-use, such as underneath ship engines.

Also, the communication of analyses results should be trustworthy. Three aspects of communication trustworthiness are security, reliability, and safety. Safety can be enabled by automated backup and multiple data storage locations.

3. Implementation: general requirements Q&A

Q) What are general requirements for Big Data implementation

Sensors?

Computers?

Personnel?

Presentation?

3. Implementation: general requirements Q&A

Q) What are general requirements for Big Data implementation

Q) Sensors?

A) data type specific, situation-specific properties, robust fixing

Q) Computers?

A) compatible data acquisition, sufficient data storage, adequate speed

Q) Personnel?

A) state monitoring expertise, sensor installation/maintenance skill, application knowledge

Q) Presentation?

A) process-specific formats, location-specific formats, trustworthy communication