Data Mining Applications in Manufacturing

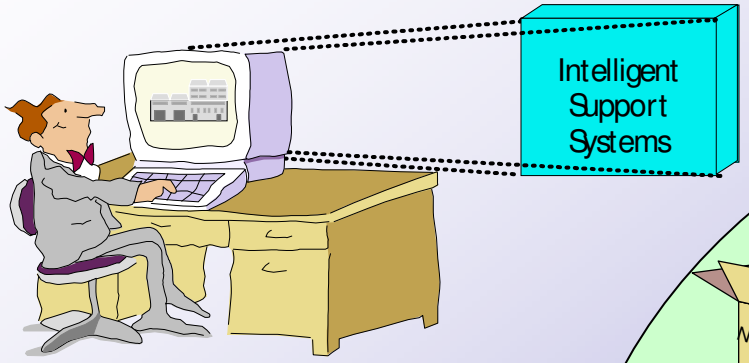
Dr Jenny Harding

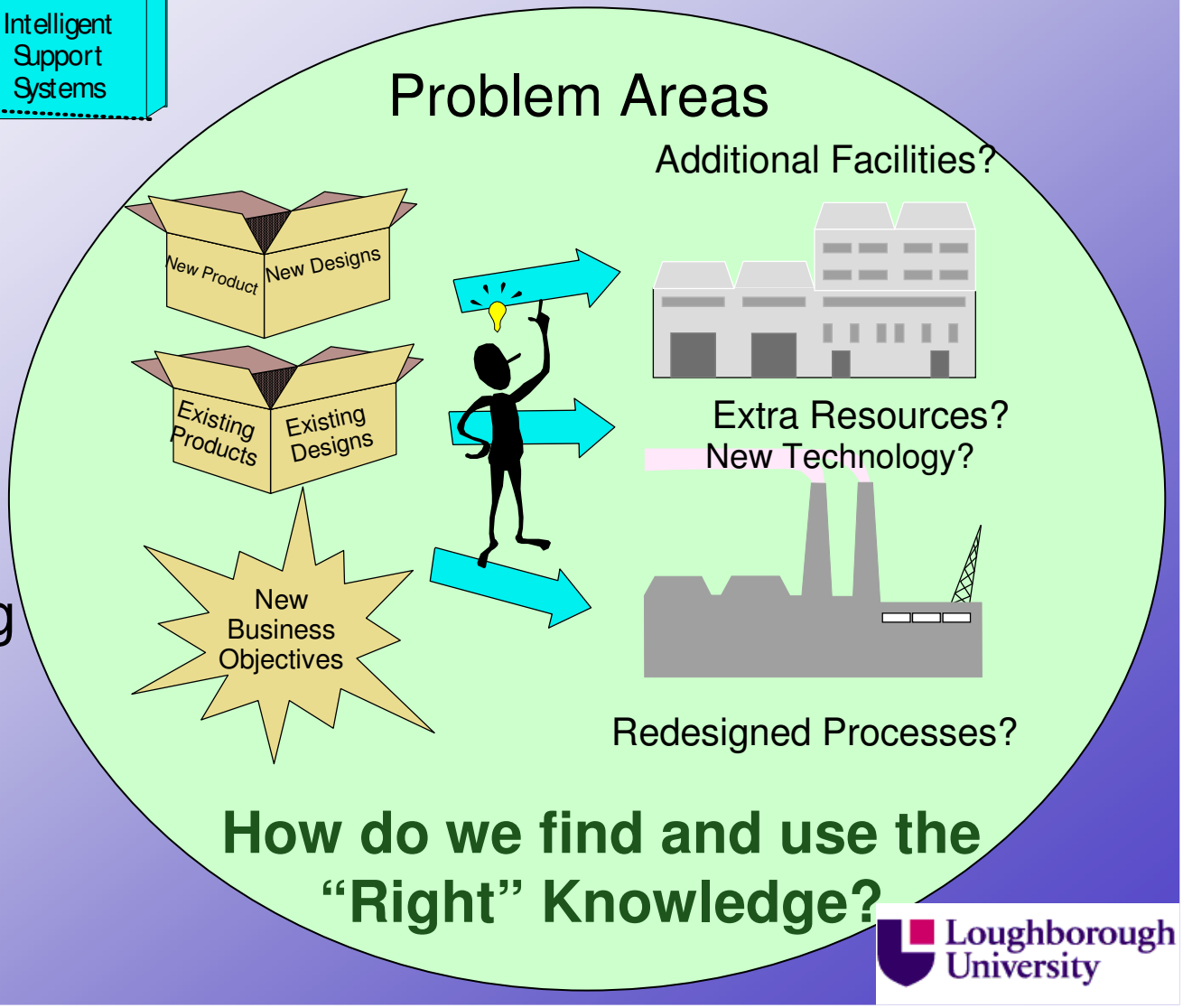
Senior Lecturer

Wolfson School of Mechanical & Manufacturing Engineering, Loughborough University

Loughborough University

Identification of Knowledge - Context



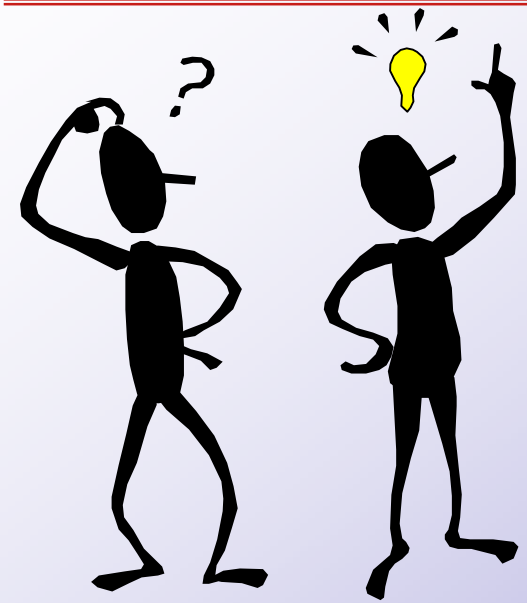


Tools:

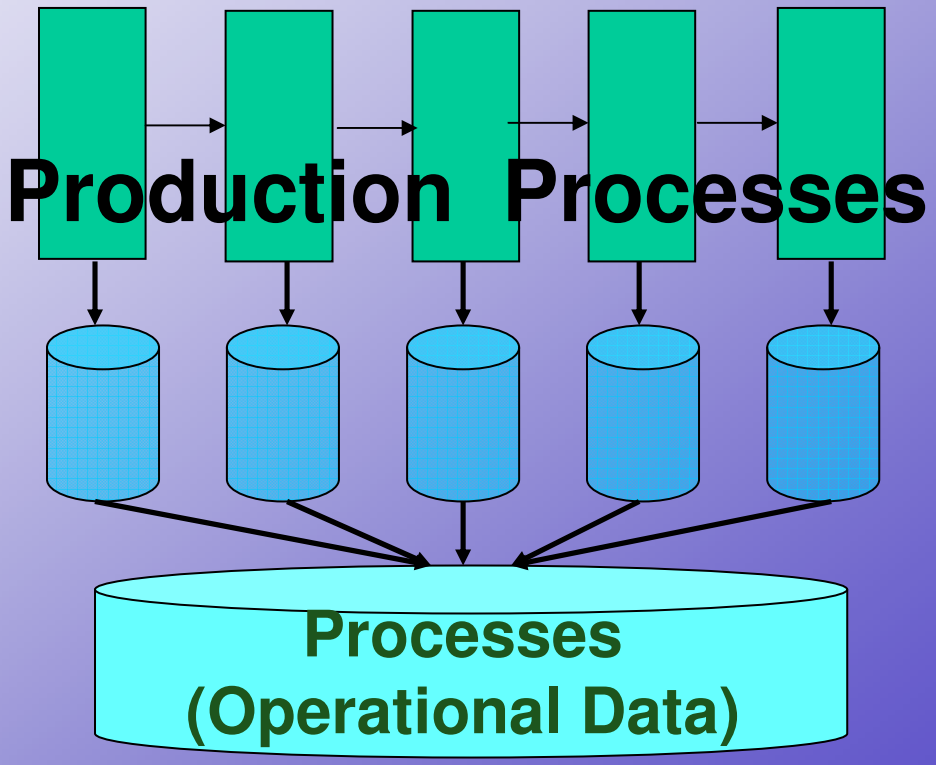
Information Modelling
Process Modelling
Artificial Intelligence
Simulation
Data Mining???

Loughborough University

Sources of Manufacturing Knowledge



People
Expertise and Experience



Facts or Folklore?

Under Exploited?

Loughborough University

Background History

Why Data Mining?

What is the “Right” Knowledge?

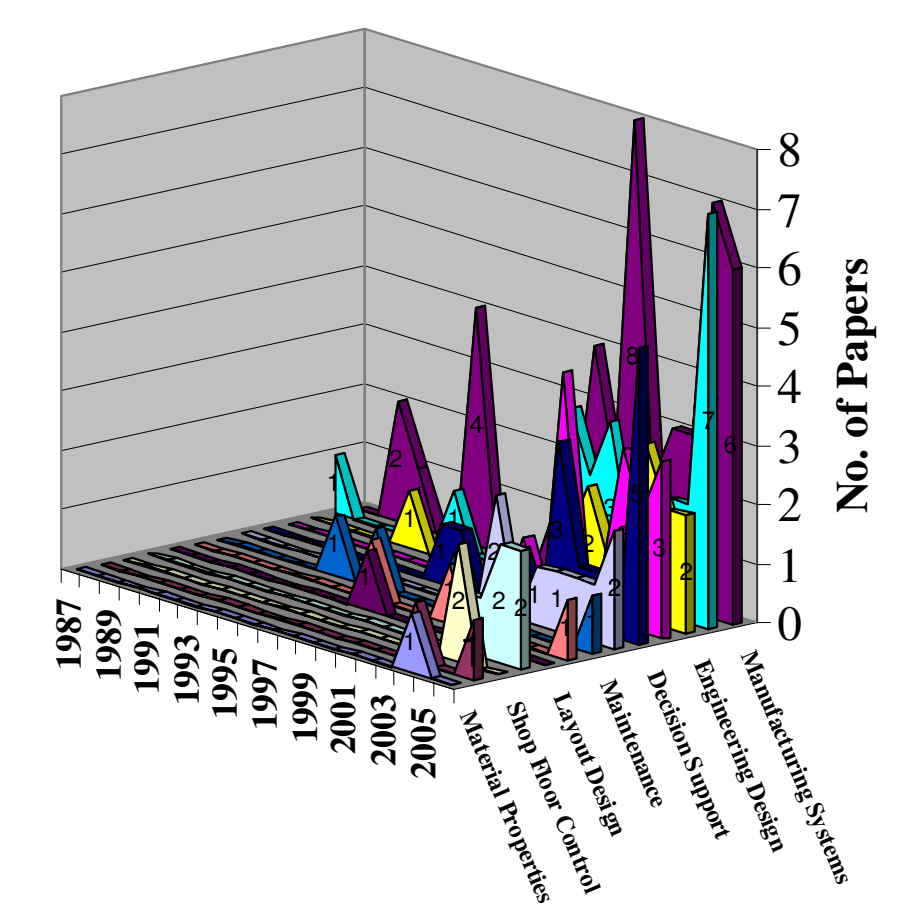
Are “experts” the only sources?

Is there any substance in folklore?

Or is the knowledge hidden somewhere else as well?

Loughborough University

Reported Applications of DM in Manufacturing



Harding, J A, Shahbaz, M, Srinivas and Kusiak, A, 2006, "Data Mining in Manufacturing: A Review", American Society of Mechanical Engineers (ASME): Journal of Manufacturing Science and Engineering.

Earliest - Fault diagnosis (1987)

Substantial increase in (manufacturing) data mining publications in last 2 years

Loughborough University

Reported Applications of DM in Manufacturing (2)

Summary:

Use of data mining applications in Manufacturing is relatively new

To date – most of the research done has focused on semi-conductor industry (fault detection and productivity improvement)

To date – most of the reported research uses decision trees, clustering and neural networks

Loughborough University

Challenges for DM in Manufacturing

Why slow adoption?

- ? Investment – very wide range of tools and techniques and not obvious in most applications which will produce best results and rewards (DM preparation costs are high).
- ? Availability and quality of data – substantial data cleaning and pre-processing required.
- ? Combinations of expertise needed – Data Mining Expertise + Expertise in Processes / Technologies under consideration
- ? Inadequate Exploitation of Resulting Knowledge - “one-off” applications addressing specific problems

Loughborough University

Challenges for DM in Manufacturing (2)

Why slow adoption?

- ? Uncertainty and risk - there are no guarantees that useful knowledge will be discovered. In highly competitive situations high risks may be unacceptable
- ? Complexity and diversity - extremely difficult to devise generic data mining processes even for particular types of manufacturing problems

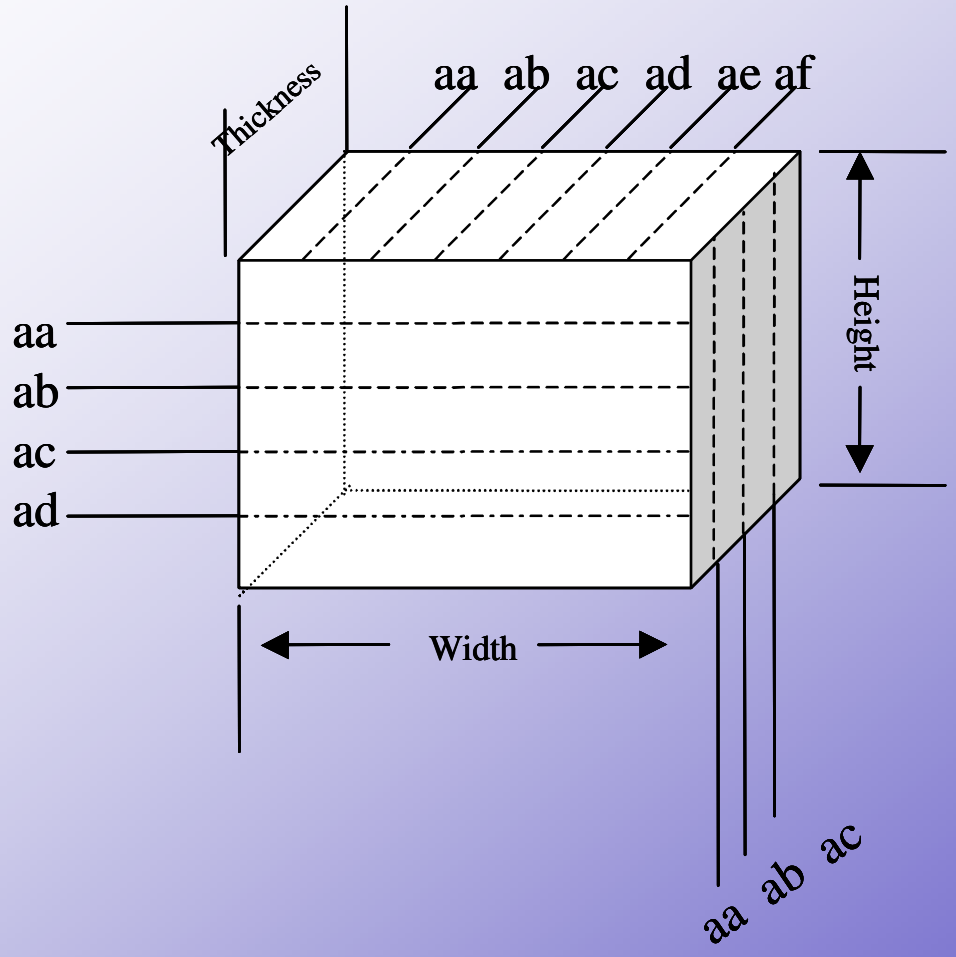
Loughborough University

Case Studies

Experience from literature and industrial case studies shows.....

Loughborough University

An Example

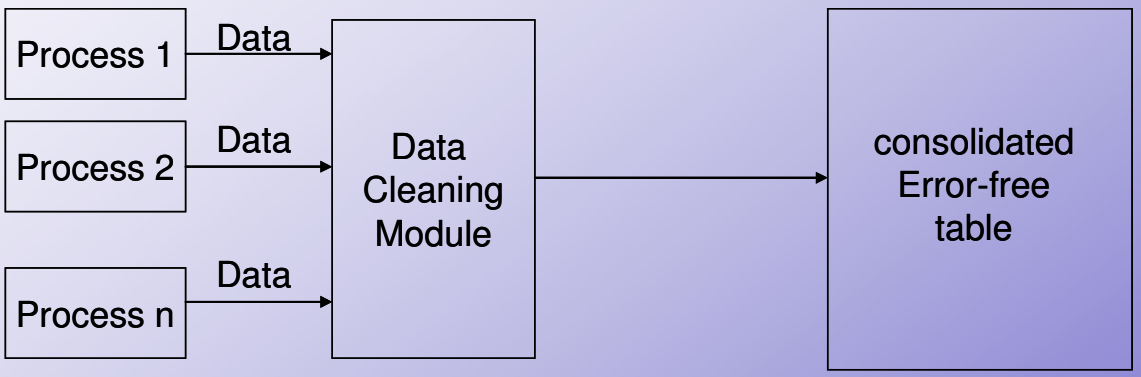


A simple sample product

Loughborough University

Data Cleaning

- Very important! (But very time-consuming)
- Iterative Process



General Classifications

Identical Records
Some Duplication in Records
Confusing Records
Missing Values in Records
etc

Possible Causes

Breakdowns / Restarts
Rework cycles
Operator Error
Operator Training
Etc., etc.

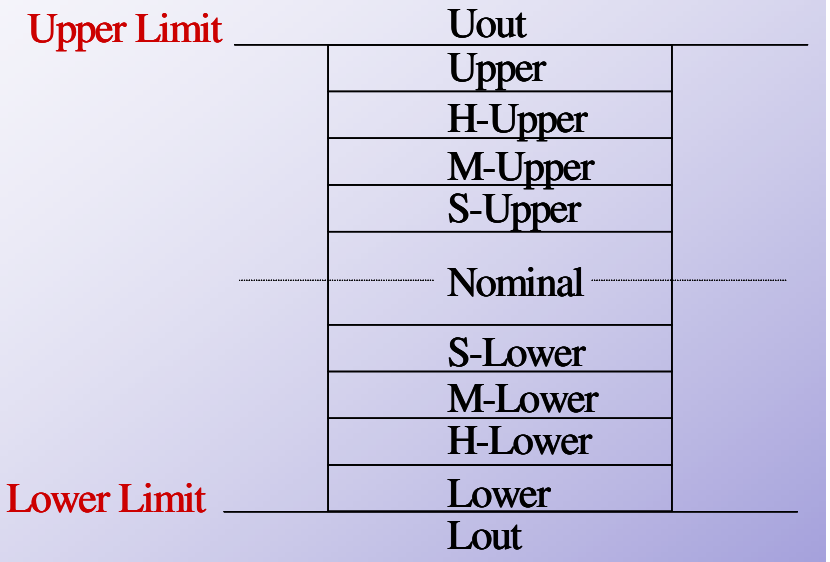
Loughborough University

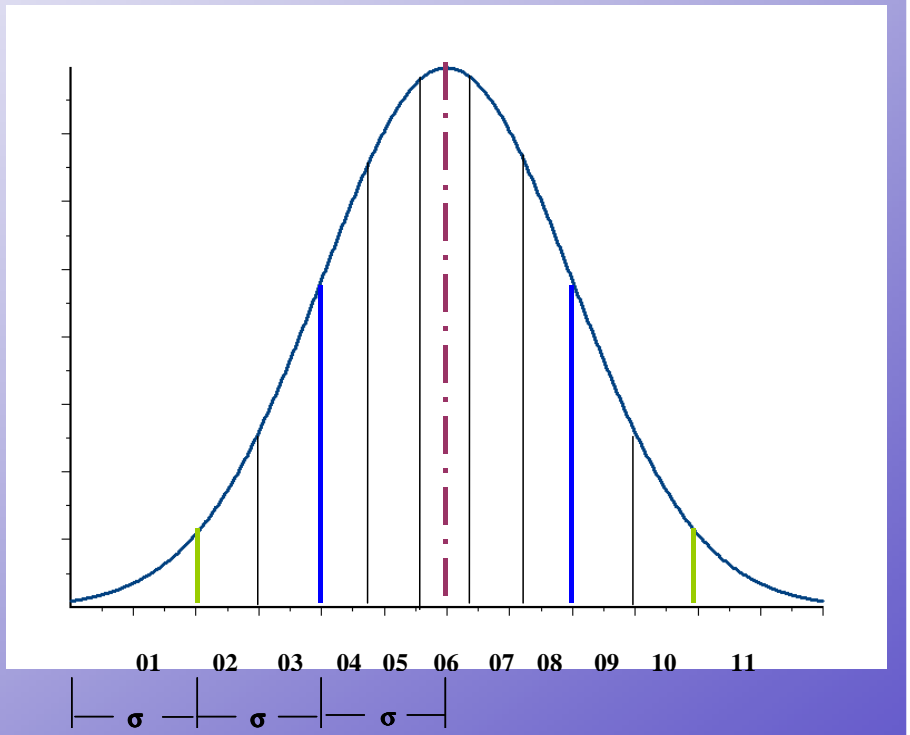
Example Thoughts during Cleaning

- Should Reworked records be included or ignored?
- Can Missing values be replaced or should the whole record be rejected?
- Can domain experts of more detailed files be consulted about replacing the abnormal data?
- All the replaced values need to be documented for future consultation.

Loughborough University

Data Transformation





Products - split tolerance range for dimension.

Equal divisions???

Process variables

Normal distributions?

Equal divisions???

Loughborough University

Data Mining

A complete set of dimensional / manufacturing data for each product needs to be collected and transformed to make each record

Regression analysis used first – are there any simple relationships that can be exploited?

Several data mining approaches tried

Association Rules the Apriori Algorithm has been most successful

Loughborough University

Data Mining

In Manufacturing - Data mining decisions should be made in the context of the physical realities of the data that is being examined.

In very highly controlled manufacturing processes, unusual combinations of results (representing problems) may be rare.

Support level kept very low to reduce risk of missing unusual but important combinations of manufacturing outputs

Loughborough University

Association Rule

Used in mostly *Market Basket Analysis*

Apriori Property:

"*Any subset of a large itemset must be large*"

Rules are in the form of

A ⇾ B or A ⇾ (B & C) etc.

e.g. Chips ⇾ Coke

•Each rule is based on certain Support and Confidence level.

Loughborough University

Rule Quality

Need to check Quality of the generated rules – use Support, Confidence, Statistical methods AND check with domain experts that the rules make sense in the Manufacturing Context.

Support: Support is the frequency of the occurrence of items in a transaction or record.

Confidence: Confidence is the certainty of the discovered rule.

Loughborough University

Example Rules

Types of Manufacturing Rules that may be identified include:

Product Design Errors or Constraints

Process Errors or Constraints

From Product Data - Rules of the form

Good Dimension output ← **Good Dimension output**

Bad Dimension output ← **Bad Dimension output**

Good Dimension output ← **Bad Dimension output**

From Product and Process Data

Process variable values that produce Good Product Results

Process variable values that produce Poor Product Results

Loughborough University

Example Rules

- Rules that indicate one dimension being “nominal” results in another dimension being “nominal” are reassuring – but generally not very interesting

- Rules that indicate one dimension being “nominal” results in another dimension being out of tolerance or away from “nominal” are more interesting.

Possible Causes include:-

Design Error
Process Error
Process Constraints

Loughborough University

Results

- Data Mining can be used to extract Manufacturing process limitations.
- Data Mining can be used to discover design constraints and then help to improve the product design, (if manufacturing process is acceptable).
- Data Mining can be used to improve the output quality by controlling the manufacturing process variables.

Loughborough University

Future Considerations.....

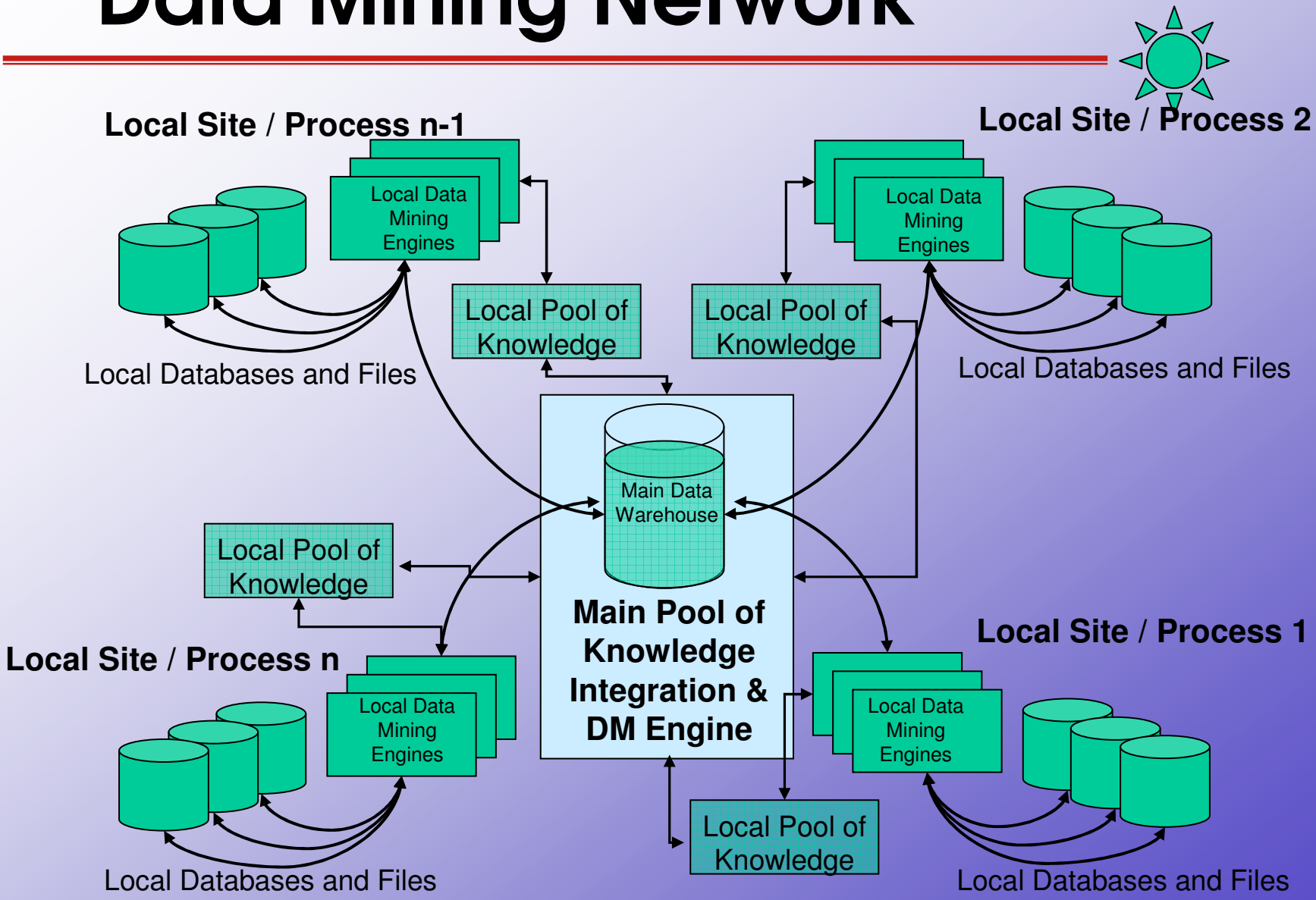
Data Mining projects tend to be "one-off" applications to solve a particular problem

But how do we make sure the acquired knowledge is made use of when it has been discovered?

Greater benefits would be gained if they were part of an ongoing Knowledge Reuse programme

Loughborough University

Data Mining Network



Any Questions

Thank You

J.A.Harding@lboro.ac.uk