



The
University
Of
Sheffield.



Sheffield
Water Centre

Machine Learning data-driven analytics for smart water networks

Dr. Steve Mounce, Research Fellow in Hydroinformatics,
Civil and Structural Engineering, The University of Sheffield





The
University
Of
Sheffield.



Sheffield
Water Centre

Overview

- Data to information: the Big data issue - why is less than 0.5% of data ever analysed and used?
- Data mining using artificial intelligence for knowledge discovery and predictive analytics
- Data-driven online event detection for clean and wastewater networks and real-time control
- Future directions



©marketoonist.com



The
University
Of
Sheffield.



Sheffield
Water Centre

Roads help us navigate to
a location.
Data help us navigate to
a decision.





The
University
Of
Sheffield.



Sheffield
Water Centre

“I have a million Big Data ideas, but, they all point to certain death”

“Information is the oil of the 21st century, and analytics is the combustion engine”



“More data has been created in the past two years than in the entire previous history of the human race”

“Big Data is not about the data – it is about the analytics”

You can have data without information, but you cannot have information without data

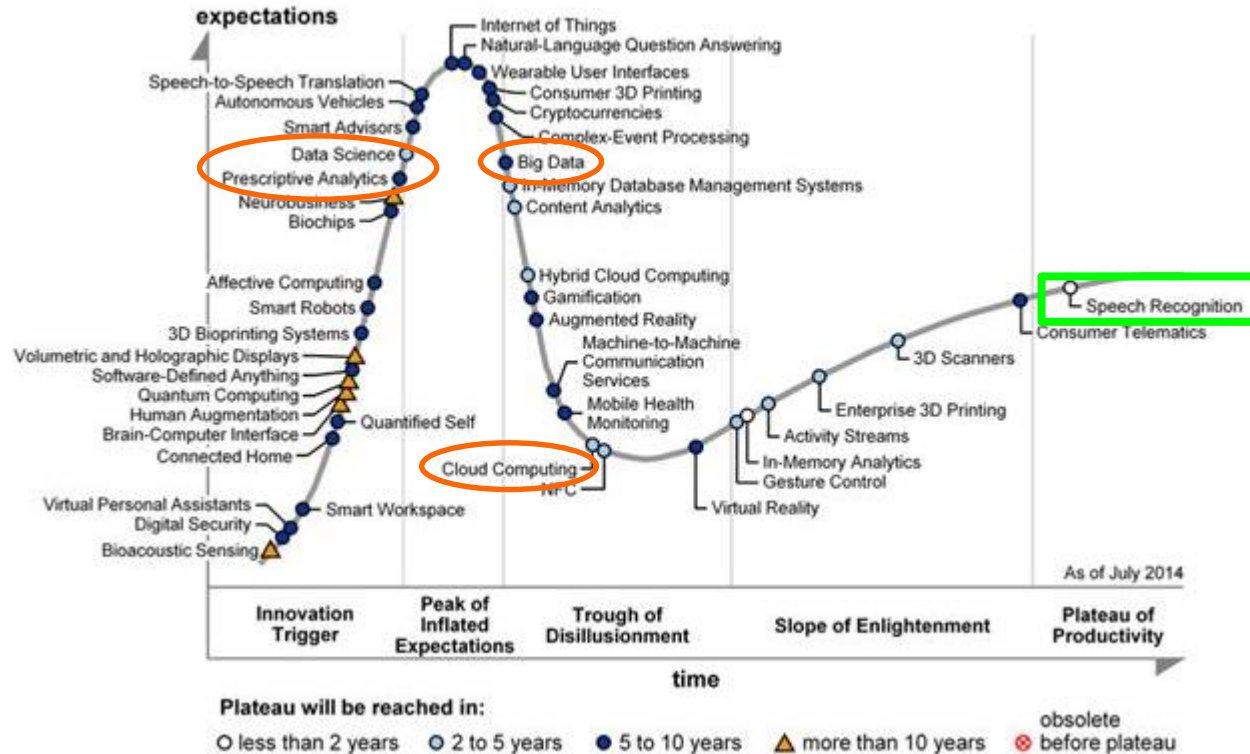


The
University
Of
Sheffield.



Sheffield
Water Centre

Gartner Hype Cycle for Emerging Technologies, 2014



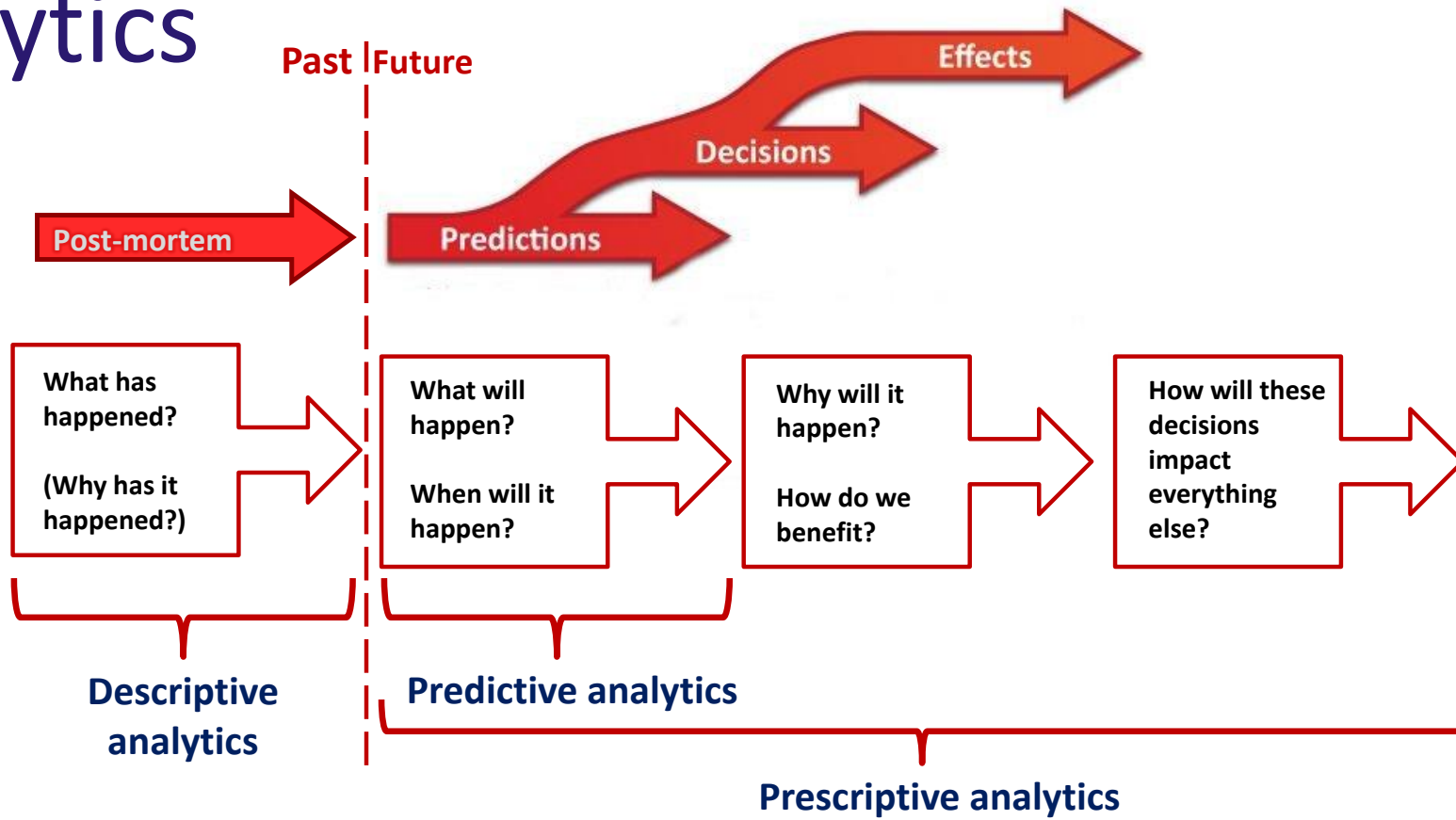


The
University
Of
Sheffield.



Sheffield
Water Centre

Analytics





The
University
Of
Sheffield.



Sheffield
Water Centre

Why do we care?

- By 2020, at least a third of all data will pass through the cloud
- Retailers who leverage the full power of big data could increase their operating margins by as much as 60%.
- 73% of organizations have already invested or plan to invest in big data by 2016.
- At the rate at which data and our ability to analyse it are growing, it is reasonable to expect that most companies will be using the impact of big data analytics in the next five years.



The
University
Of
Sheffield.



Sheffield
Water Centre

Data use

- Approximately 80% of data is closed (organisational) or personal – large volumes untapped
- ‘If data is the new oil, then personal and closed data is the new rocket fuel’
- Water sector issues with proprietary data – trust etc.
- Secondary or tertiary uses not realised or known
- Open closed data up in the future to data innovators?



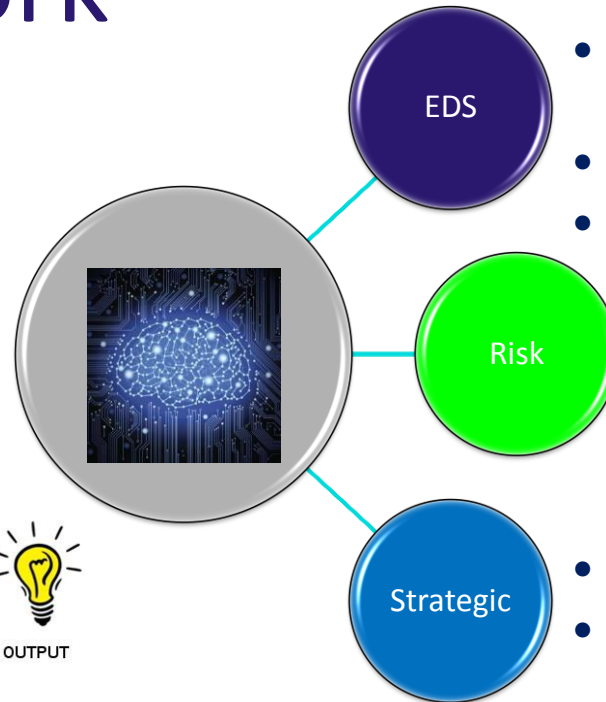
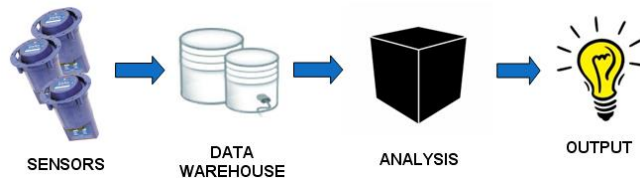
The
University
Of
Sheffield.



Sheffield
Water Centre

Previous and current machine learning work

Common theme: data driven models, especially ANNs



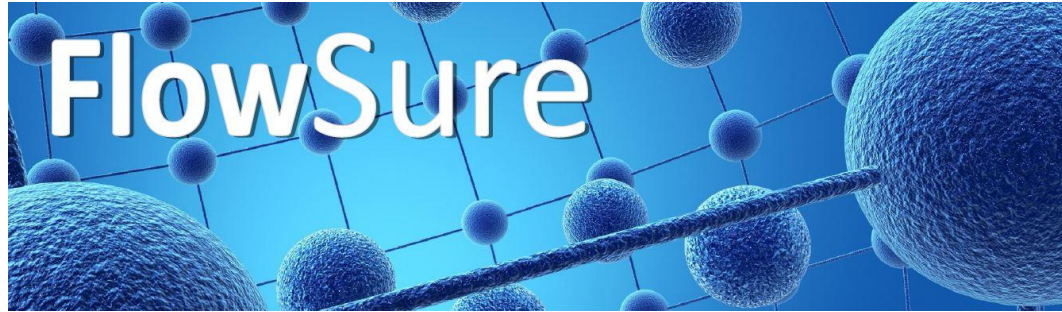
- ANN/ FIS for leak detection
 - SVR and BED for hydraulic anomalies
 - CANARY WQ EDS (USEPA)
 - CSO Analytics (ANN/ FIS)
- SOM and FIS stagnation pipe risk model
 - Ensemble DT predictive model for DMA iron risk
- Semblance analysis
 - EPR and SOMs for regen rates



The
University
Of
Sheffield.



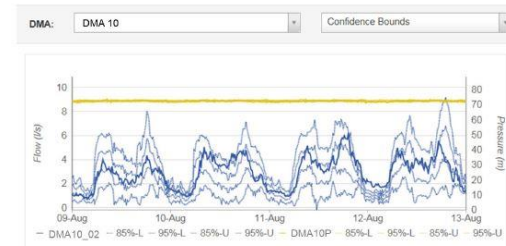
Sheffield
Water Centre



- FlowSure uses readily available meter data and smart algorithms to automatically identify when a significant burst or other large, unusual flow has occurred
- The software uses an artificial neural network to learn how the water supply network behaves normally and then applies fuzzy logic to detect abnormality
- Commercialisation of PhD research and R&D pilots (ADA)



Winner of IWEX
University
Challenge 2010





The
University
Of
Sheffield.



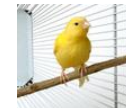
Sheffield
Water Centre

EDS in WQMS

- Collaboration with Intellitect for water quality EDS for improving real-time system knowledge allowing proactive network management
- EPA/Sandia National Labs has developed open source EDS software platform called CANARY(MATLAB/Java based)
- Software tuned for Intellisondes and optimised for response speed, number of alerts, false positive/negative ratio etc. Online system running, evaluated & lessons learned
- Now developing **bespoke platform** for detection, and to enable progress towards classification and diagnosis: potentially to inform both strategic and tactical operations



intellitect
water monitoring systems



EPA
United States
Environmental Protection
Agency



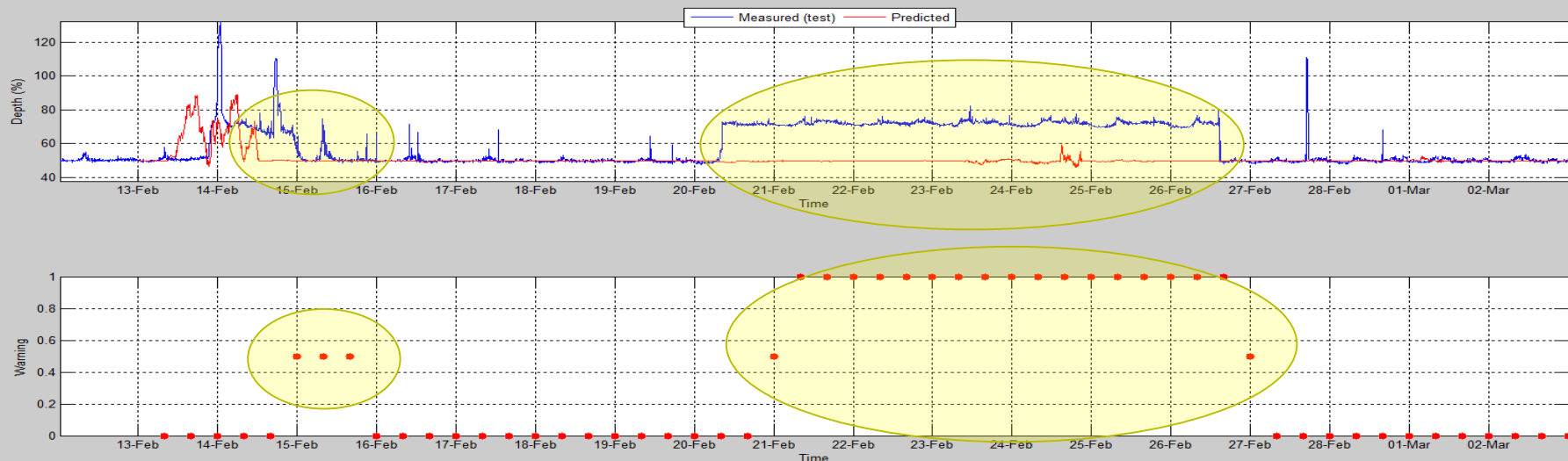
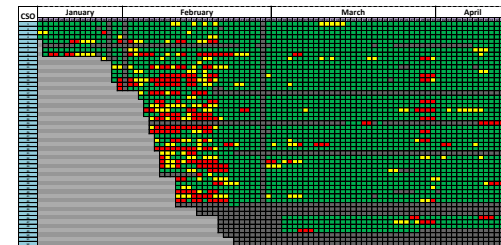
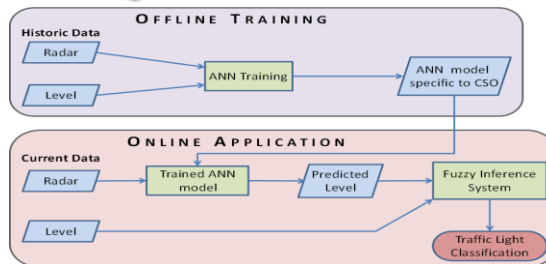
The University
Of Sheffield.



Sheffield
Water Centre

CSO Analytics

'Red alert'	'0'	['']	['']
'Amber alert'	'2'	['']	['']
'Green'	'40'	['']	['']
'..'	['']	['']	['']
'Time'	'CSO No.'	'Warning'	'Qual'
'12/03/14 01:20'	'1'	'Amber'	'Good'
'12/03/14 01:00'	'2'	'Green'	'Poor'
'11/03/14 10:00'	'3'	'Green'	'Good'
'12/03/14 02:10'	'4'	'Green'	'Good'
'11/03/14 11:00'	'5'	'Green'	'Poor'
'12/03/14 03:45'	'6'	'No Data'	['']
	'7'	'Green'	'Good'



- Automated level and rainfall analysis

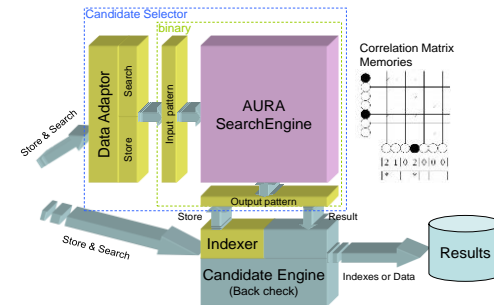
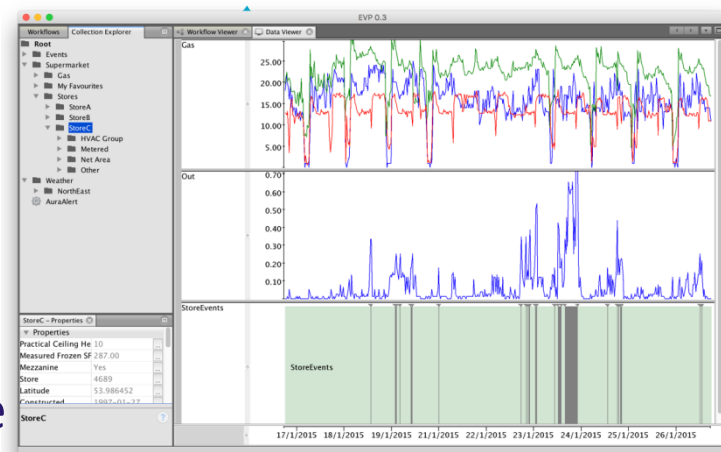


The
University
Of
Sheffield.



Analytics: AURA-Alert

- Advanced Uncertain Reasoning Architecture (AURA) Alert utilises the power of associative memories for novelty/ leakage detection:
 - Provides means to train system on normality without complex underlying models
 - Validate unknown data against stored states to see if unusual
 - Highly scalable approach for Big Data
 - Benefits from one-shot and high speed training



CYBULA
high performance pattern recognition systems

SW4EU

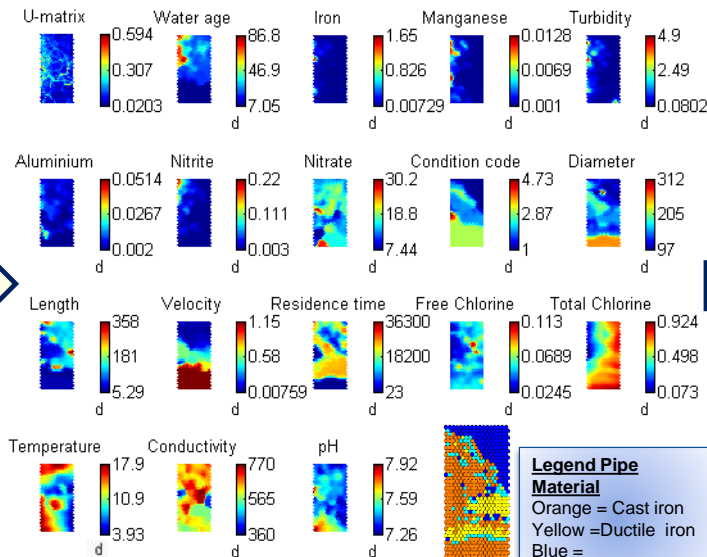
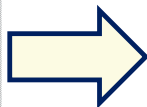
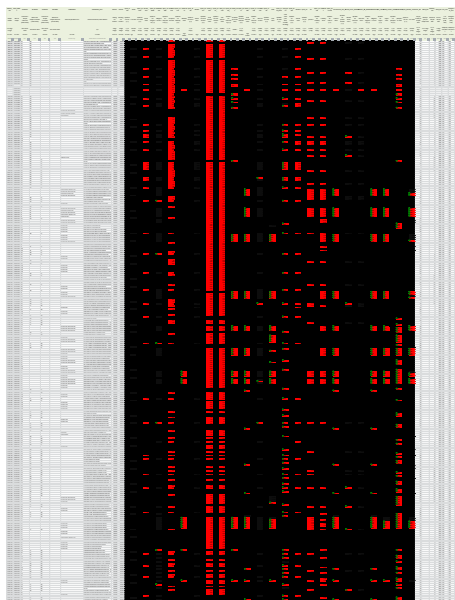


The University
Of
Sheffield.

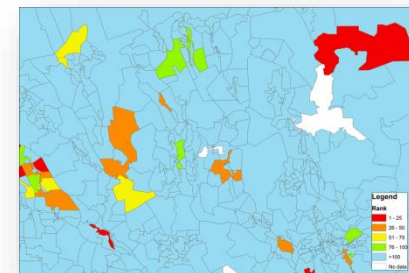


Sheffield
Water Centre

Data mining to risk ranking



Pipe level



DMA level

SOM results for water quality, asset, modelled data and with material type label

Sparse data, knowledge discovery, fuzzy/ ensemble ranking



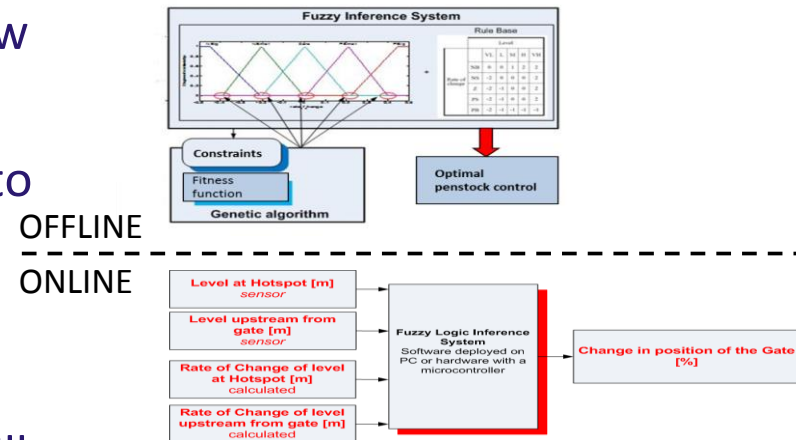
The
University
Of
Sheffield.



Sheffield
Water Centre

Cost Effective Neural Technique to Alleviate Urban flood Risk

- Cost effective, local autonomous sewer flow control system to reduce urban flood risk
- CENTAUR will use data driven approaches to develop real time control strategies to activate existing in sewer storage at the local scale
- Sophisticated computational techniques will be combined with specially designed flow control devices to reduce flood risk.



Horizon 2020
Programme





The
University
Of
Sheffield.



Sheffield
Water Centre

Deep Learning: Game Changer

- Deep Learning is when Big data hits Machine learning
- Why a hot topic now?
 - Big data availability
 - DL algorithmic breakthroughs (Hinton and others)
- Creation of tools allowing us to go beyond engineering for problems exceeding human understanding
- Examples: GoogleBrain (10 million videos), smartphone: transcription of speech/ SIRI, medical apps, fraud...





The
University
Of
Sheffield.

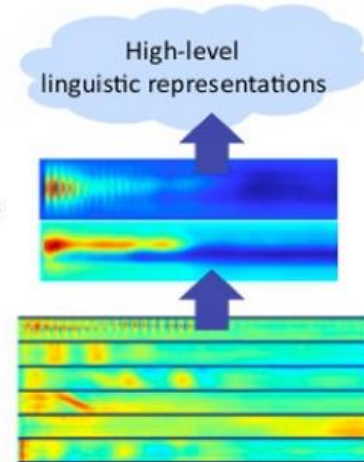
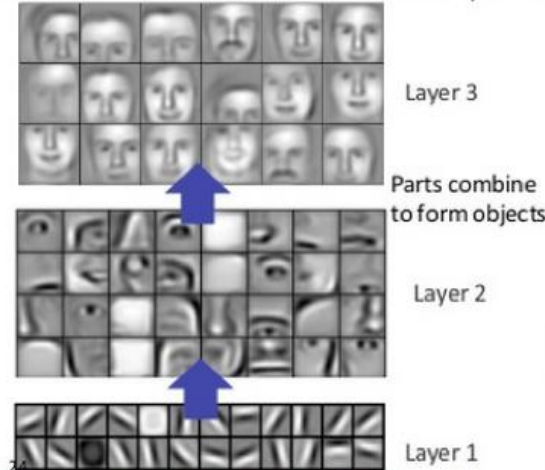


Sheffield
Water Centre

Deep Learning

- New era of Machine Learning
- Use Artificial Neural Networks
- Large amounts of data...
- ...allow very deep structures
- ...and the automatic learning of features
- (Big) Data to Action

Successive model layers learn deeper intermediate representations



Prior: underlying factors & concepts compactly expressed w/ multiple levels of abstraction

Training data:
10-100M images

Network
architecture:
~10 layers
1B parameters

Learning
algorithm:
~30 Exaflops
~30 GPU days



The University
Of
Sheffield.



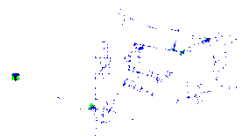
Sheffield
Water Centre

Data proliferation: Leakage R&D pilots

1998



Dial-up modem
15min
Daily download



5 DMAs
15 measurements

2008



GPRS
15min
30 min download



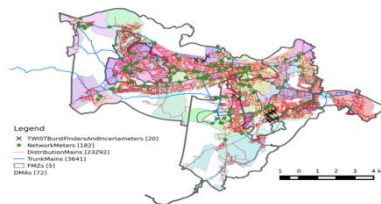
144 DMAs
431 measurements

2016

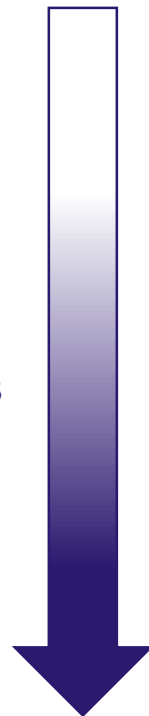
SW4EU



AMI
15min
Hourly download



1 DMA
4,500 AMR
measurements





The
University
Of
Sheffield.



Sheffield
Water Centre

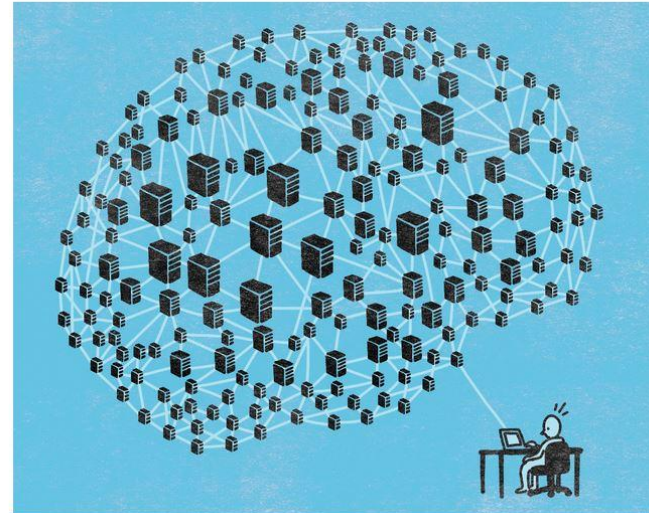
Platform and toolbox/API sharing for ML and Deep Learning



TensorFlow



Microsoft releases CNTK, its open source deep learning toolkit, on GitHub



- ML used to need expertise (PhDs etc): barrier to entry now lowered

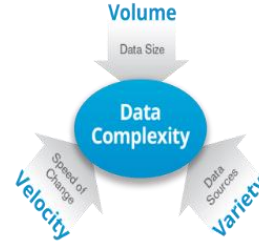


The
University
Of
Sheffield.



Sheffield
Water Centre

Summary



- Data (as) infrastructure?
- How to unlock water sector proprietary data in faster, better & more trusted ways
- Avoid being a slow adopter of data science
- *“Big data is a bit like an iceberg—most of the value we want to unlock is still lurking under the surface.”*
- Future possibilities of Deep Learning technologies



The
University
Of
Sheffield.



Sheffield
Water Centre

Any questions?

Further details:

<http://www.sheffield.ac.uk/penninewatergroup>

<http://www.smartwater4europe.com>

<http://www.sheffield.ac.uk/centaur>

s.r.mounce@sheffield.ac.uk



www.mouncehydrosmart.co.uk

With thanks to:



Prof. Joby Boxall