# **Honours Project Ideas**

# Data Cleaning (Horizontal research/study)

* Tools (any open source?):
  + Trifecta
  + OpenRefine
  + Paxata
  + Alteryx
  + Distributed system data cleansing – Optimus (Apache Spark)
  + Cloud Dataprep (google)
* Quality Screens
  + Column screens – testing individual columns
  + Structure screens – testing integrity of relationships between columns: primary/foreign keys
  + Business rule screens – testing if data follows specific business rules
* Limitations
  + Costs – hundreds of thousands of dollars
  + Time – mastering large-scale data-cleansing software is time consuming
  + Security – cross-validation requires sharing information, giving the application access across the system, including sensitive legacy systems.
* Challenges and problems
  + Error correction & loss of information
  + Maintenance of cleansed data
  + Data cleansing in virtually integrated environments
  + Data-cleansing framework

# Internet of things

* Applications:
  + Consumer
    - Smart homes
      * Control house functions from office – turn down aircon etc.
      * Amazon Alexa – listen to news cutting vegetables
      * Sensors monitoring heat, water, electricity etc. and adjusting to be more efficient
    - Assistance for disabled/elderly
      * Voice control for deaf
      * Monitors for medical emergencies – falls/seizures
    - Intelligent shopping systems
      * Monitor user shopping habits to give more user-defined offers on products they bought
      * Location of items they need – via fridge creating list on phone saying what has run out (milk/eggs etc)
  + Enterprise Internet of Things (EIoT)
    - * Used by businesses
        + Media

Studying user habits to better present advertising

* + - * + Manufacturing (Industrial Internet of Things (IIoT))

Digital control systems – automate process controls, operator tools, service information systems – optimize safety and security

Asset management

Integrated with the smart grid

Intelligent maintenance systems

* + - * + Environmental monitoring

Air/water quality

Changes by climate

Movements of wildlife and habitats

Improve earthquake or tsunami early-warning systems

Span large geographic areas, can be mobile

# Software Evolution

* Leads to open source development
  + Large number of new ideas discovered and generated
* Types of software maintenance:
  + Corrective
  + Adaptive
  + Perfective
  + Preventative
* Stage model
* Laws of software evolution
* Converting legacy system to cloud based micro-services