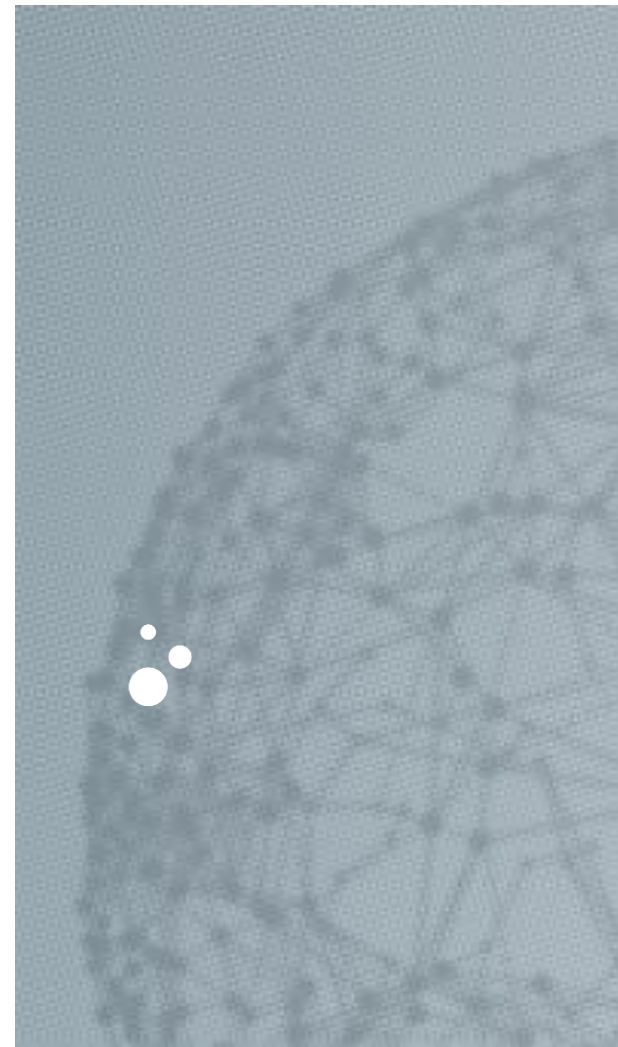
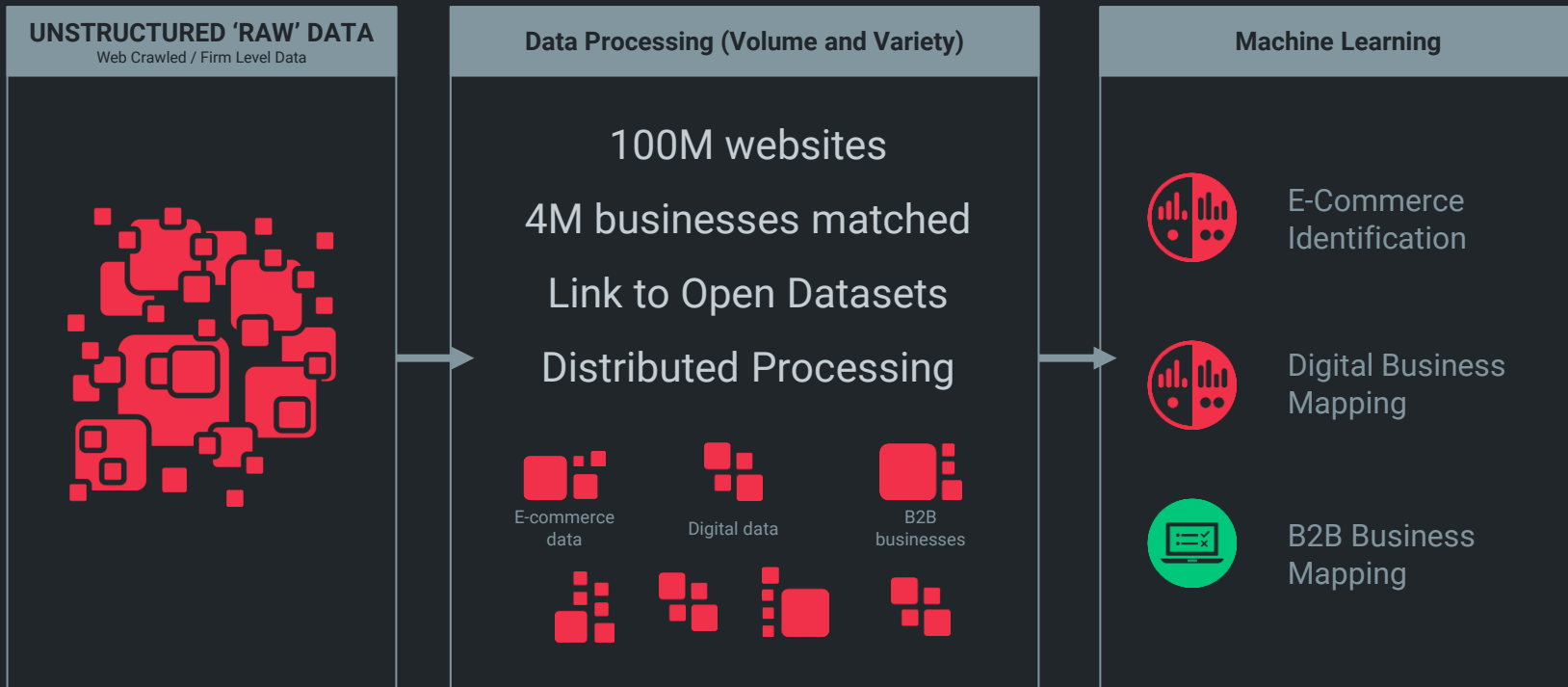


Big Data: Perspectives and Progress





ESCoE Conference on Economic Measurement 2018

Prash Majmudar
@prashmaj



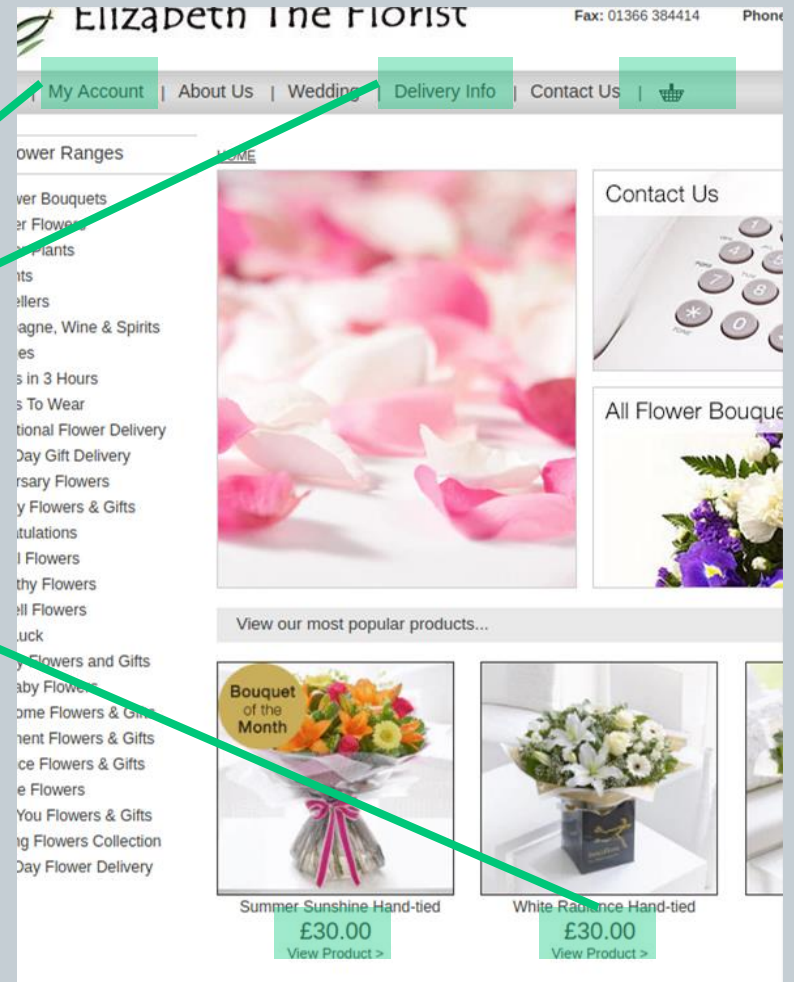


Feature extraction and Supervised ML Modelling

 Shopping Cart Links	 Product Pricing
 Delivery Page Info	 E-commerce technologies

Combine with Firm-level data from Companies House

77K Firms in the UK classified as E-commerce

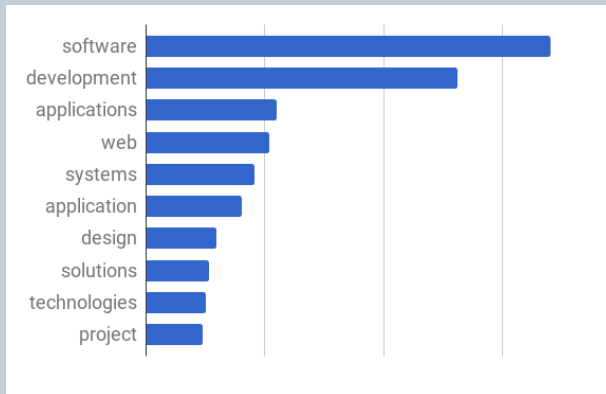


The screenshot shows the website for Elizabeth The Florist. At the top, there is a navigation bar with links for 'My Account', 'About Us', 'Wedding', 'Delivery Info', and 'Contact Us'. Below the navigation bar, there is a large image of pink and white rose petals. To the right of the petals is a 'Contact Us' section with a phone keypad. Below the petals is a section titled 'View our most popular products...' which displays two flower bouquets. The first bouquet is 'Summer Sunshine Hand-tied' priced at £30.00. The second bouquet is 'White Romance Hand-tied' also priced at £30.00. Green boxes highlight the 'My Account', 'Delivery Info', and the price tags on the bouquets. Green lines connect the 'Shopping Cart Links' and 'Product Pricing' boxes from the table to the 'My Account' and 'Delivery Info' links on the website. Another green line connects the 'Delivery Page Info' box to the price tag of the 'Summer Sunshine Hand-tied' bouquet.

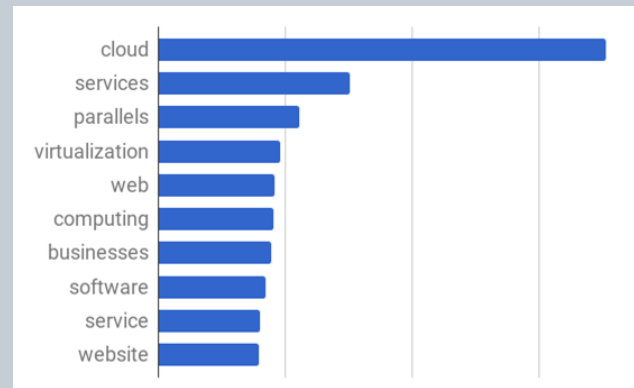
Unsupervised identification of Business Digital Sectors (LDA Topic Modelling)

100's of key business sectors / descriptors discovered

Software Development 'Topic'



Cloud Computing 'Topic'





Techniques maturing (NLP, Machine Learning)

Sample data at individual firm-level - capture nuance

Minimise 'lag' between collection and decision making

Firm data challenges: E.g. Non-Limited businesses difficult to observe

Coverage - businesses without online presence - data collection bias



Data matching and linking of industry and Gov / Open datasets

Workforce (Data Scientists) to complement Economists /
Statisticians



Projects

“Tech Nation 2016 Report” - NESTA and Tech City UK with GrowthIntel Core Project Partner

“Mapping Information Economy Business with Big Data: Findings from the UK” - 26 Nov 2014:
Max Nathan (NIESR), Anna Rosso with Francois Bouet (GrowthIntel)

“Measuring the UK’s digital economy with Big Data” - 22 July 2013: *Dr. Max Nathan (NIESR) and Anna Rosso with Prash Majmudar, Alex Mitchell, Tom Gatten*

THANK YOU

