

# The economic impact and organisational aspects of online data analytics

Juan Mateos-Garcia, Nesta

**Information and Knowledge for All:  
*Towards an Inclusive Innovation***

**The World Conference on Intellectual Capital for Communities**

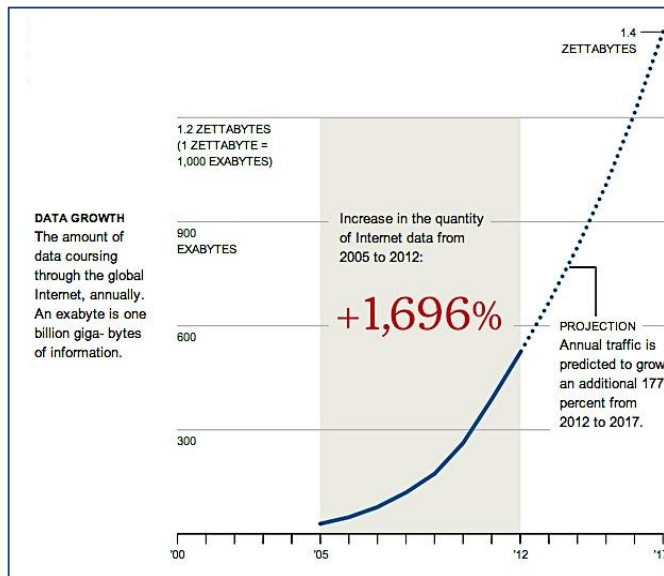
**UNESCO, July 3&4 2017**

# Structure

- Context: A big data revolution
- Our paper: questions, data and method
- Findings
- Conclusions

# Context: A (big) data revolution

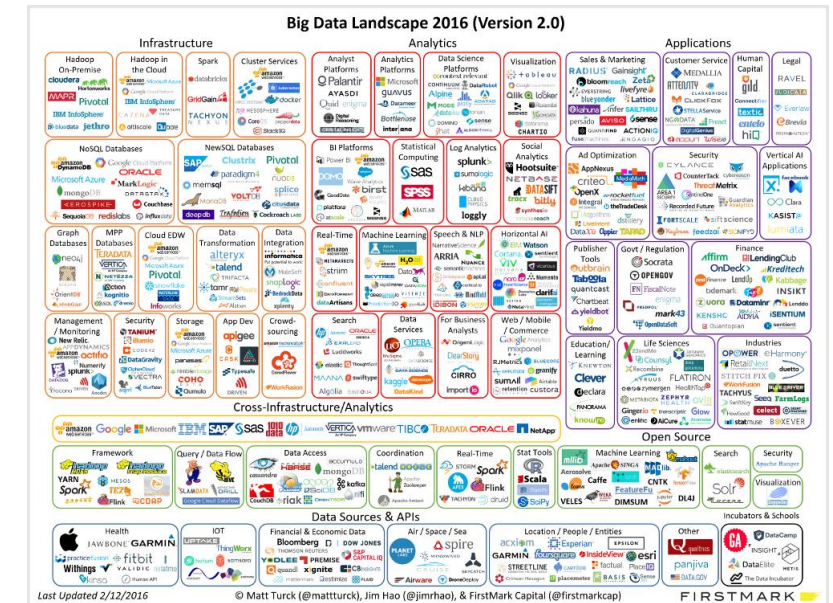
## More data



## More analysis

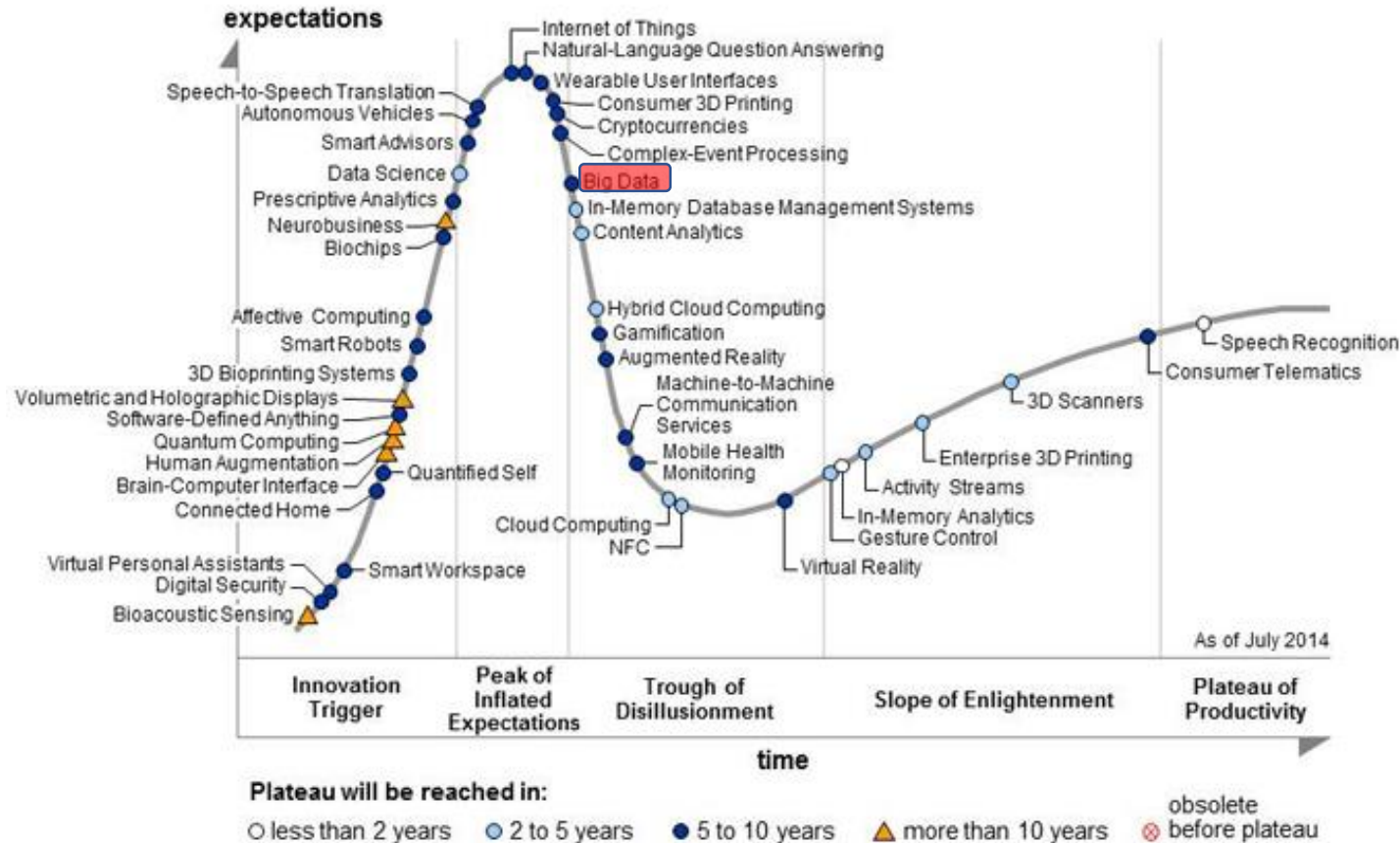


## More applications



**Source:** Cisco via New York Time  
<http://nyti.ms/12fNxdrs>

# But what is the impact?



And how are the benefits distributed?

Big questions with implications for...

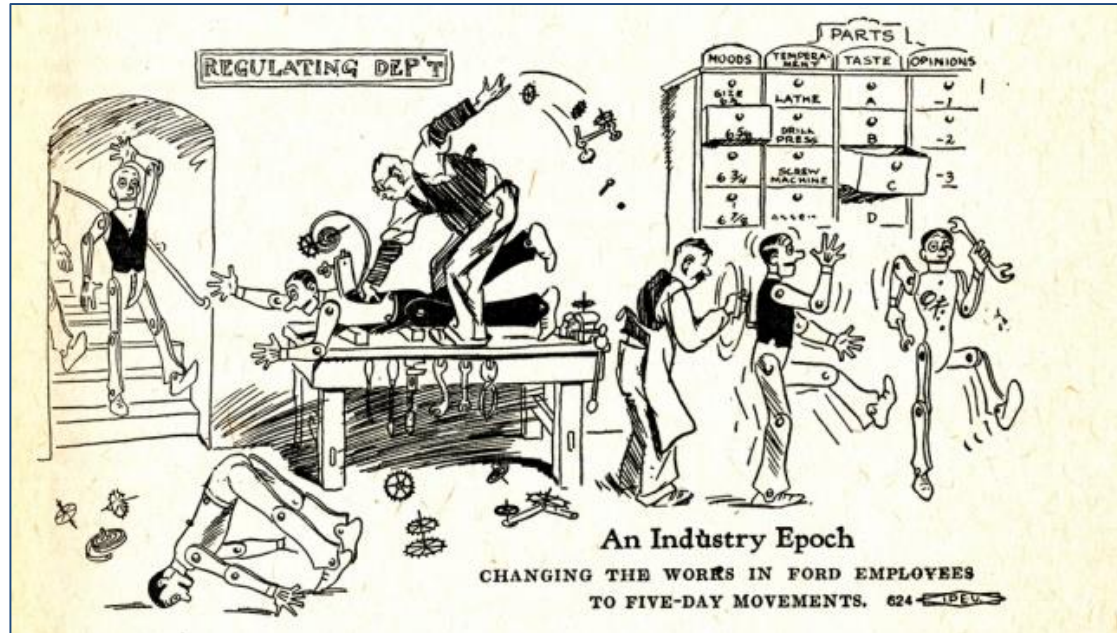
- Innovation policy
- Regulation (data protection, competition)
- Management

## Some evidence of impact

- **Brynjolfsson et al. (2011)** surveys 179 US publicly listed companies, and finds that a one-standard deviation increase in data-driven decision-making adoption has a positive effect on productivity levels (5%-7%) and on business profitability.
- **Tambe (2013)** uses LinkedIn data to study the complementarities between 'big data' adoption and skills in the US and finds a strong relationship between productivity and firms' investment in 'big data', proxied by their employment of individuals mentioning Hadoop skills in their personal LinkedIn profile.
- But still very limited



# And how to organise to benefit from data?



- + data = Less need to consult with managers
- + data = More scope for experimentation
- + data = More scope for automation
- + data = Easier monitoring

Net effect on organisation?

# Our paper: questions

1. What is the link between online data use and firm productivity?
2. Are there complementarities between online data use and employee autonomy?
3. Are there complementarities between online data use and process innovation?

# Method

Use a standard value-added production function equation to estimate the impact of data on TFP:

$$\ln(Y) = \beta_0 + \beta_1 \ln(K)_{it} + \beta_2 \ln(L)_{it} + \beta_{3x} Data_x + controls + \varepsilon$$

Allow for industry heterogeneity in factor shares

Control for:

- IT intensity (proxied by IT employment share)

- Human capital (proxied by average wages)

- Web “relevance”

- Innovation

- Industry and year fixed effects



# Data

## Survey data:

Telephone survey of 500 UK companies commercially active online, with more than 50 employees (2012). Focus on online customer (not necessarily 'big') data, survey targeted to marketing activities/functions (and complementary survey of IT managers)

- Online data use including:
  - Collection of data
  - Analysis & communication
  - Deployment in the business
- Index of all the above
- See annex for summary

## Financial data (2006-2012):

- FAME data, sourced from mandatory annual accounts fillings with Companies House (majority private companies).

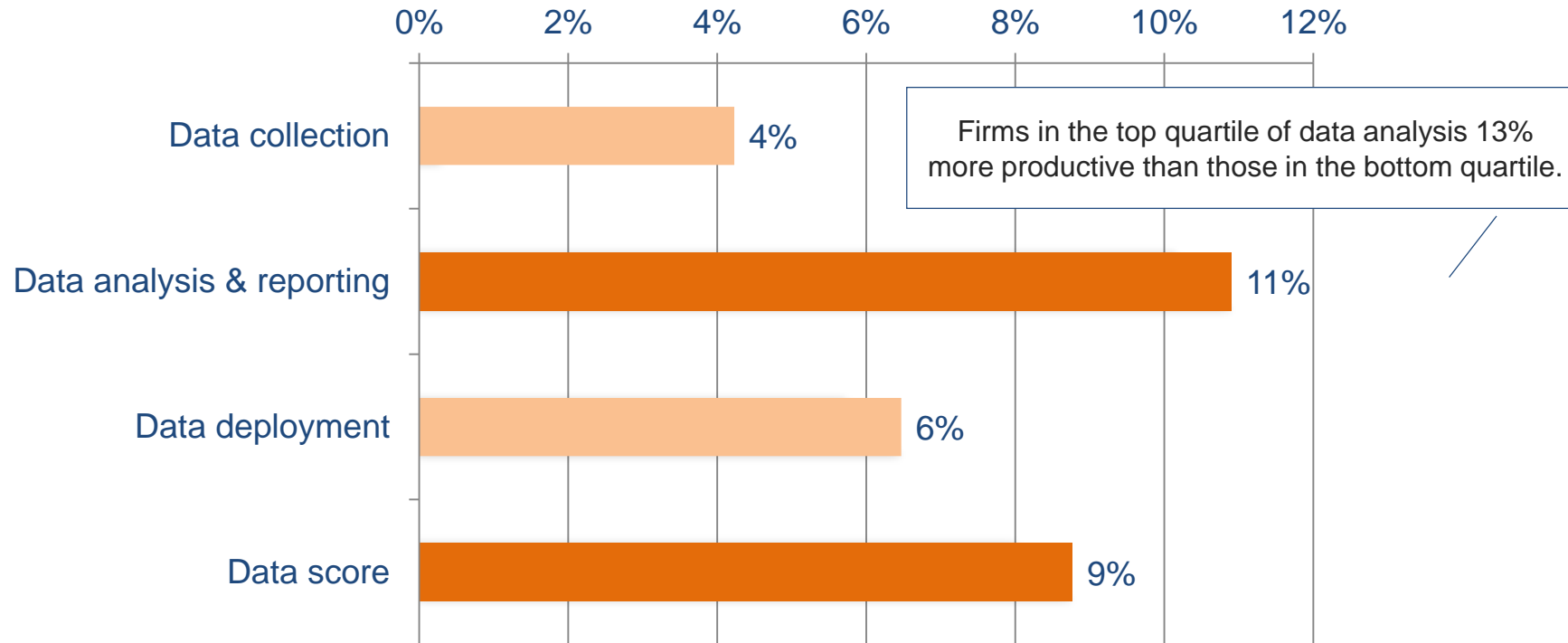
# Findings: Data measures correlation

	Data collection	Data analysis & reporting	Data deployment	Data score
Data collection	1			
	500			
Data analysis & reporting	0.5613*	1		
	500	500		
Data deployment	0.6342*	0.5230*	1	
	500	500	500	
Data score	0.8653*	0.8215*	0.8503*	1
	500	500	500	500

# Online data & value added

**Value added** = Turnover minus costs of goods sold

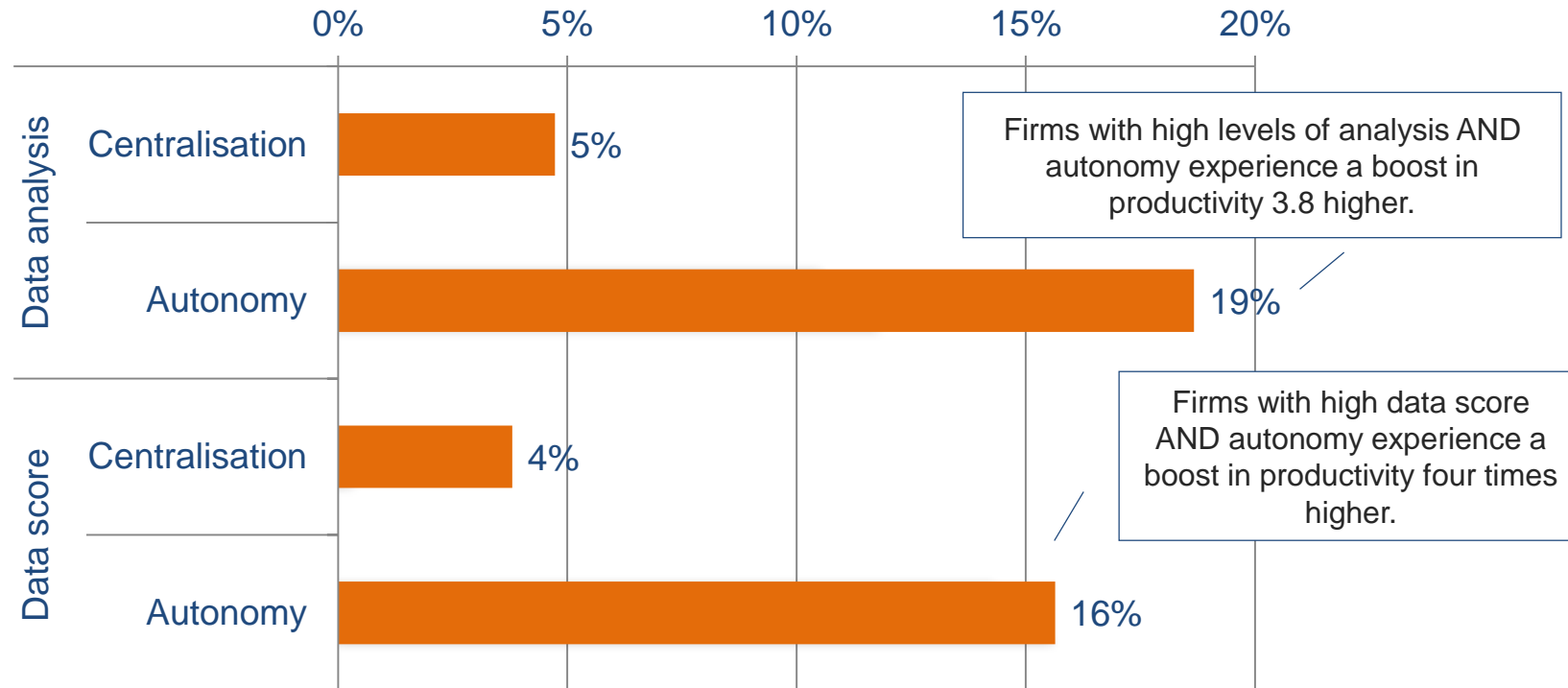
% improvement in productivity for firms with higher levels of adoption (with all controls)



# Organising to create value from data

**Employee Autonomy** = workers try new ideas, set pace & direction of work

% improvement in productivity for firms with higher than average levels in a variable (with all controls). All statistically significant.



# Robustness

- Results robust to:
  - Excluding average wage (which measures human capital but also captures profits and employees' ability to extract rents)
  - Different data proxies (continuous vs. dichotomous, etc)
  - Reduced time period (2010 and 2011)
- Similar (but weaker) results when looking at firm profitability.
- But analysis doesn't account for endogeneity in our data variable
- Findings consistent with follow-up work focused on skills

# Conclusions

- Strong link between data and productivity
  - It is not collecting data that matters but analyzing it systematically.
- Need to adapt organisational structures and processes to make the most of data
  - Complementary between employee autonomy and data consistent with the idea that more & better data allows employees to make more decisions without consulting with their managers
  - Complementary between process innovation and data highlights the importance of updating processes to generate a return on data investment (as we know from ICT studies more widely)
- **Technology adoption is not enough: need to pay attention to organisational and human factors**



# Annex: Survey questions

Variable	Definition
Data collection	Online data collection indicator (standardised)  <i>Based on 1-5 scores for 5 items: Comprehensiveness in collection of online transaction data, service and support data, user online activity data, marketing data &amp; lifestage data.</i>
Data analysis & reporting	Online data analysis and communication indicator (standardised)  <i>Based on binary scores (1 or 0) for 7 items: Adoption of A/B tests, trend analysis and reporting, forecasting, dashboard and visualisations, segmentation, regression and propensity score modelling, and data and text mining.</i>
Data deployment	Online data use indicator (standardised)  <i>Based on 1-5 scores for 9 items: Importance of online data for making decisions regarding customer segmentation, tailoring of marketing and sales, developing products and services suited to customers, improving the website, predicting customer behaviour, reporting on performance, informing business strategy, optimising pricing, and designing and evaluating social media strategy.</i>
Data score	Combined online data use indicator (standardised)  <i>Based on the scores for Data collection, Data analysis &amp; reporting, and Data deployment above.</i>
Product innovation	Respondent launches goods and services ahead of competitors (standardised)
Process innovation	Respondent is willing to disrupt its business processes (standardised)
Online business share	Proportion of revenues generated through the website
IT employment share	IT employees as a proportion of the workforce in 2010
Autonomy	Decentralisation indicator (standardised)  <i>Based on 1-5 scores for three items: workers set the pace of work, workers decide how tasks should be performed, people are free to try new things.</i>