

OR and Data Science: a match made in heaven?

Dr. Giles A. Hindle

Professor Richard Vidgen

YHORG: 23 February 2016

Hull University Business School

Connected Thinking!

Agenda

- What is OR?
 - How is the world of the OR practitioner changing?
 - What is data science?
 - Are OR and data science the same?
 - The contribution of soft OR – a business analytics methodology
 - Putting the pieces together
 - Discussion
-

What is OR?

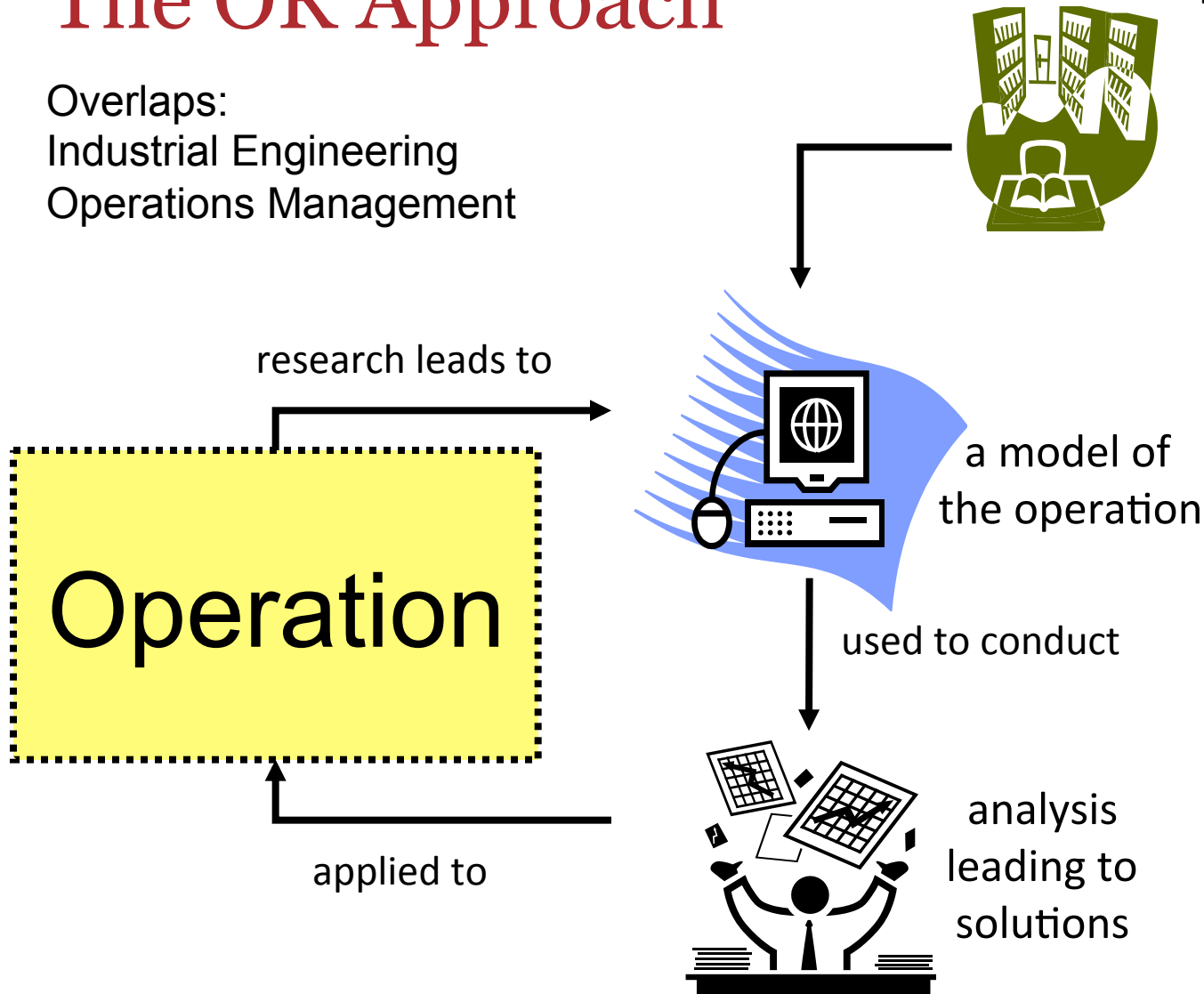
“The discipline of applying advanced analytical methods to help make better decisions”
(INFORMS and ORS)

NOTE: The definition is very broad. Could apply to OR, Soft OR
Data Science, Business Analytics, Business Analyst, Market Research.

The OR Approach

Overlaps:
Industrial Engineering
Operations Management

library of techniques:
Inventory Control,
Linear Programming,
Optimisation,
Queuing Theory,
Simulation,
Forecasting, etc.



Examples of OR in Practice

- **Scheduling:** of aircrews and the fleet for airlines, of vehicles in supply chains, of orders in a factory and of operating theatres in a hospital.
 - **Facility planning:** computer simulations of airports for the rapid and safe processing of travellers, improving appointments systems for medical practice.
 - **Forecasting:** identifying possible future developments in telecommunications, deciding how much capacity is needed in a holiday business.
 - **Yield management:** setting the prices of airline seats and hotel rooms to reflect changing demand and the risk of no shows.
 - **Credit scoring:** deciding which customers offer the best prospects for credit.
 - **Marketing:** evaluating the value of sale promotions, developing customer profiles and computing the life-time value of a customer.
 - **Defence and peace keeping:** finding ways to deploy troops rapidly.
-

Typical of OR models Used

- **Descriptive models**
 - **Predictive models (Forecasting)**
 - **Optimising Models (Linear Programming, Heuristics)**
 - **Experimental Models (Simulation, System Dynamics)**
-

What's happening in the world?

- Data is getting bigger



What's happening in the world?

- The cloud



What's happening in the world?

- The Internet of Things

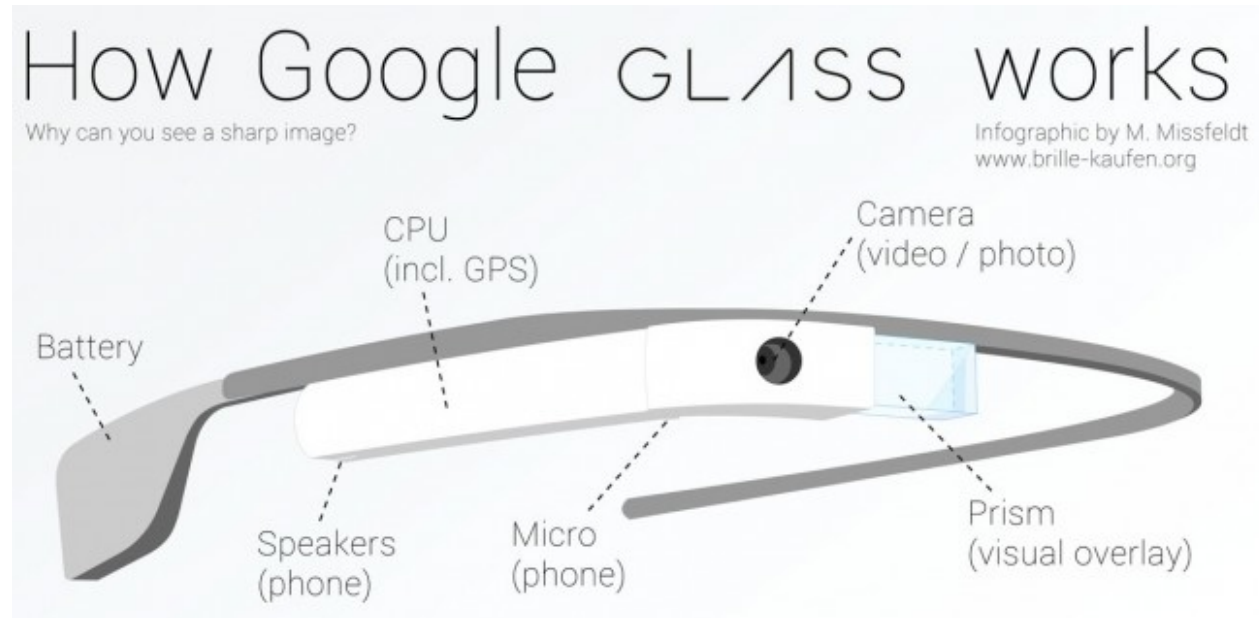


Everyday things get connected  for smarter tomorrow



What's happening in the world?

- Ubiquity



Data science: types of model

- Descriptive
 - e.g., what are the characteristics of customers who churn?
 - Predictive
 - e.g., which customers are most likely to churn?
 - Prescriptive
 - e.g., what action should we take for customers at high risk of churn?
-

What is a data scientist?

Class *DataScientist* {

Is skeptical, curious. Has inquisitive mind

Knows Machine Learning, Statistics, Probability

Applies Scientific Method. Runs Experiments

Is good at Coding & Hacking

Able to deal with IT Data Engineering

Knows how to build data products

Able to find answers to *known unknowns*

Tells relevant business stories from data

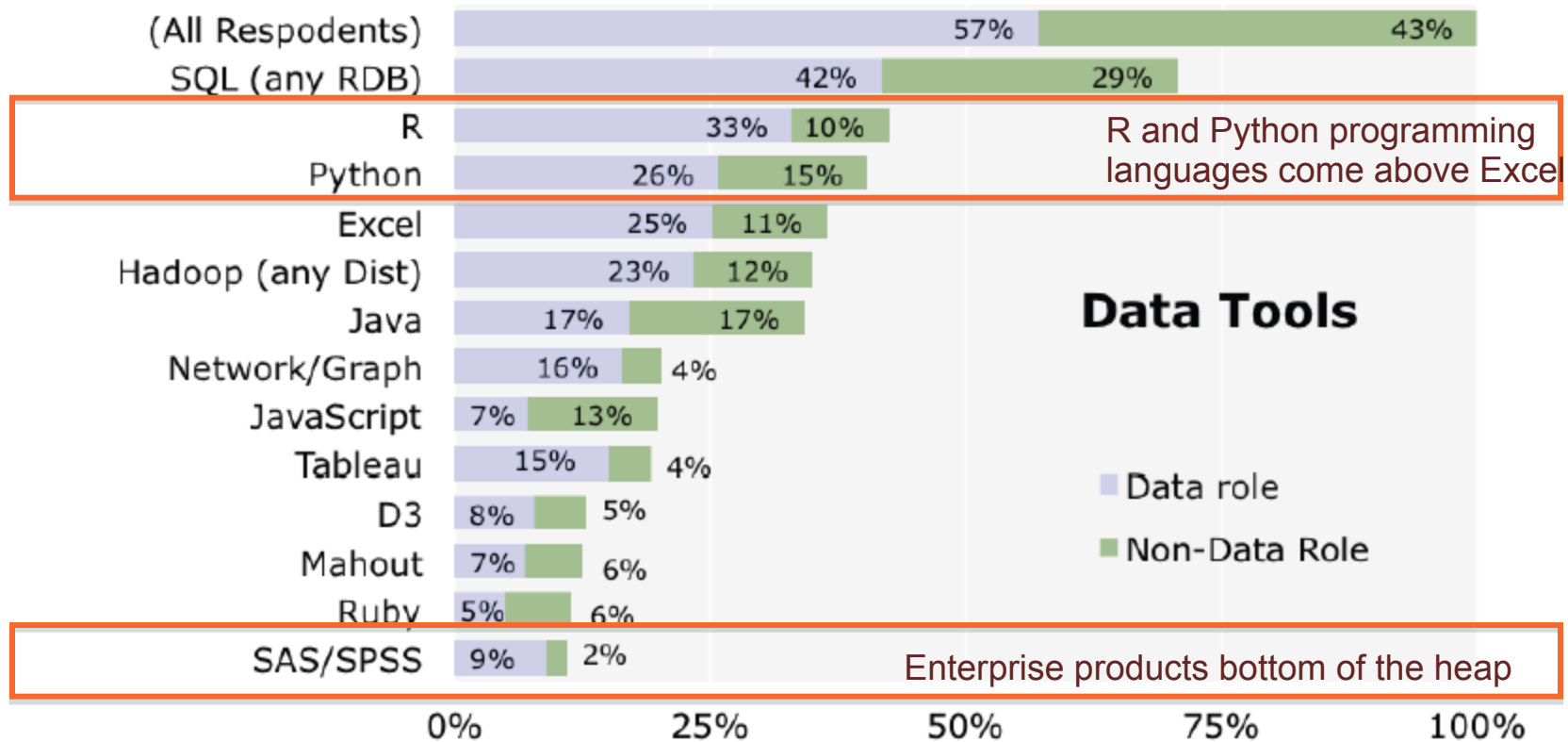
Has Domain Knowledge

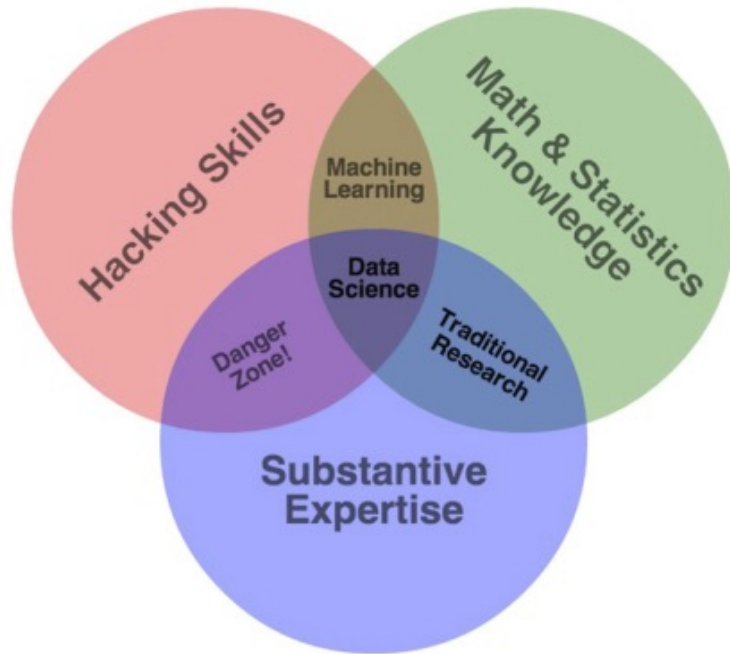
}

Some of the techniques data scientists use

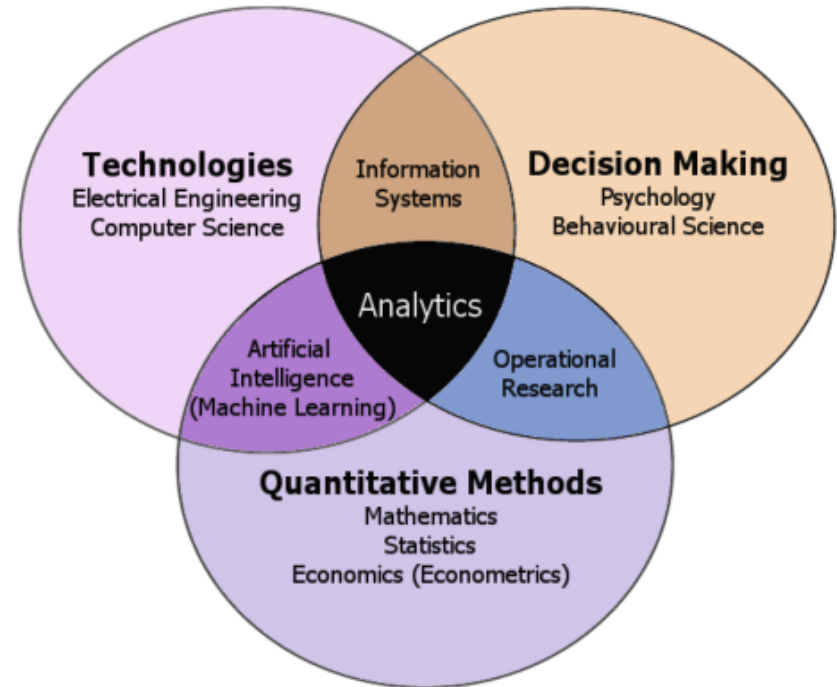
- Classification
 - Clustering
 - Association rules
 - Decision trees
 - Regression
 - Genetic algorithms
 - Neural networks and support vector machines
 - Machine learning
 - Natural language processing
 - Sentiment analysis
 - Artificial intelligence
 - Time series analysis
 - Simulations
 - Social network analysis
-

Technologies for data analysis: usage rates





(a)



(b)

Figure 1: (a) the data scientist (Conway, 2011); (b) business analytics (Robinson, 2014)

Are OR and data science practitioners the same types of people?

“... data science is not merely statistics, because when statisticians finish theorizing the perfect model, few could read a tab-delimited file into R if their job depended on it.” (Mike Driscoll, quoted by O’Neill and Schutt, 2014, p. 7).

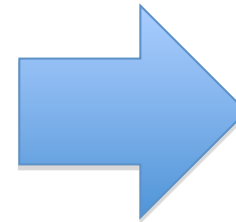
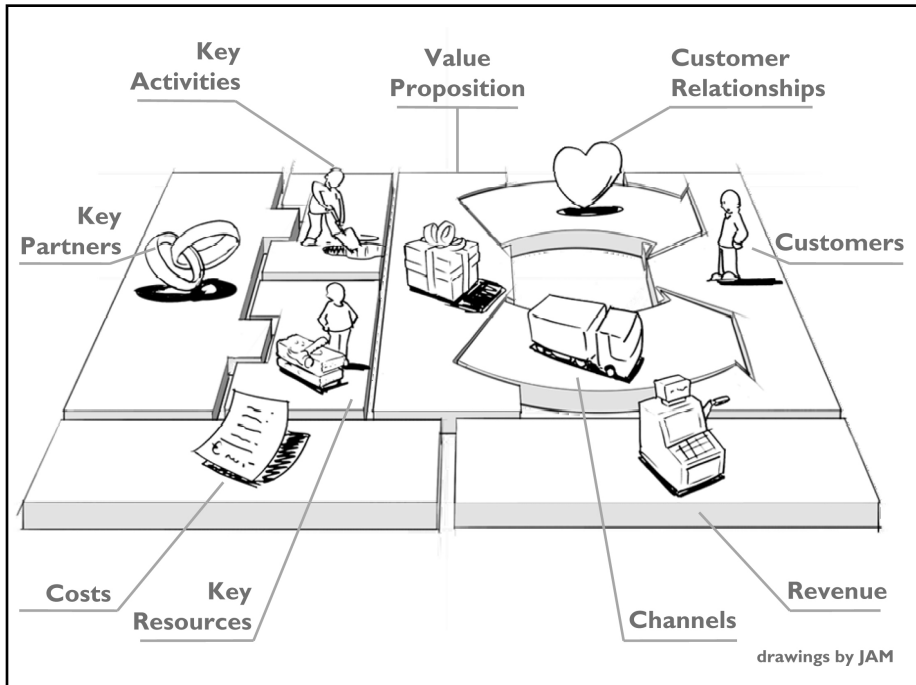
Or, put another way, Josh Wills’ definition: “data scientist (noun): Person who is better at statistics than any software engineer and better at software engineering than any statistician”.

Soft OR and Problem Structuring Methods (PSMs)

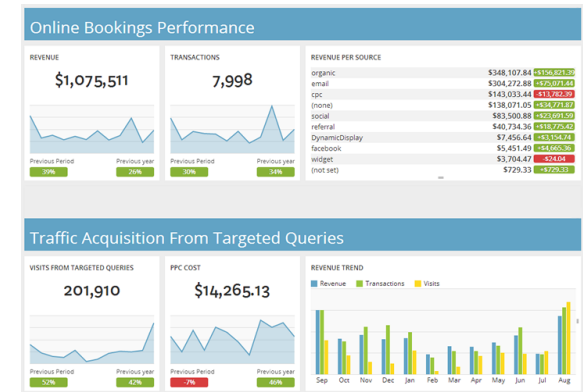
- PSMs offer “*a way of representing the situation...that will enable participants to clarify their predicaments, converge on a potentially actionable mutual problem or issue within it, and agree commitments that will at least partially resolve it*” (Mingers and Rosenhead 2004).
 - Soft Systems Methodology (SSM) used by the IS community to help appreciate the context of and need for IT
 - SSM helps define the business model which defines business valued within a situation
-

Business Logic: Value driven by Business Model

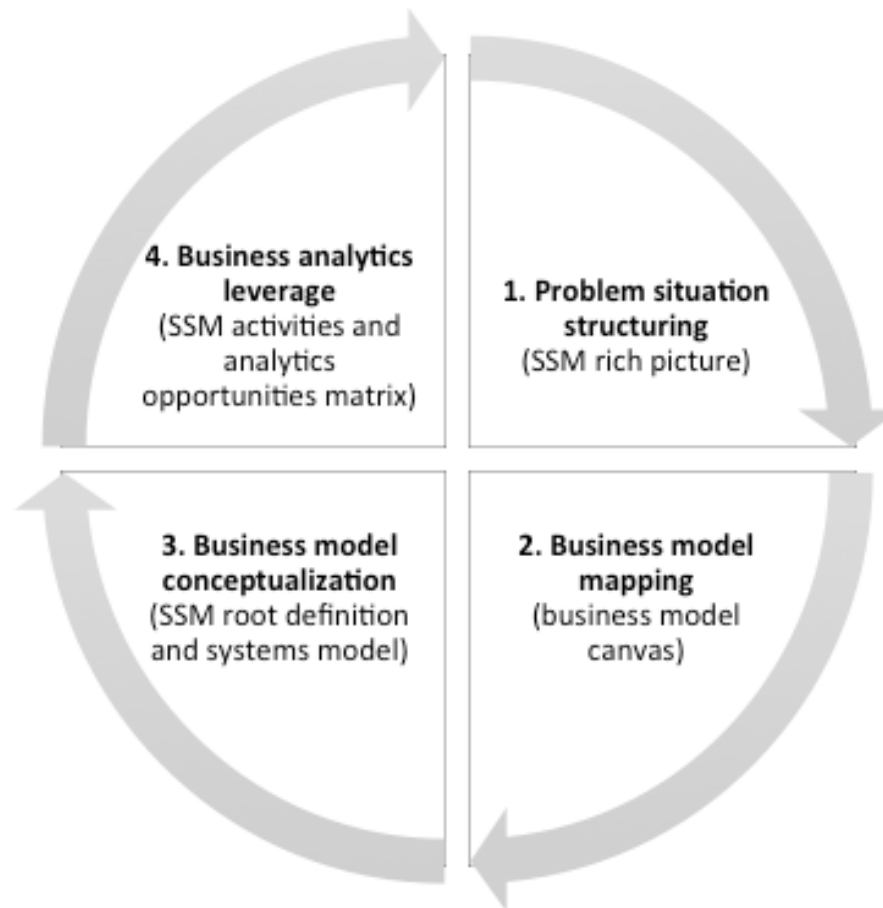
Business Model



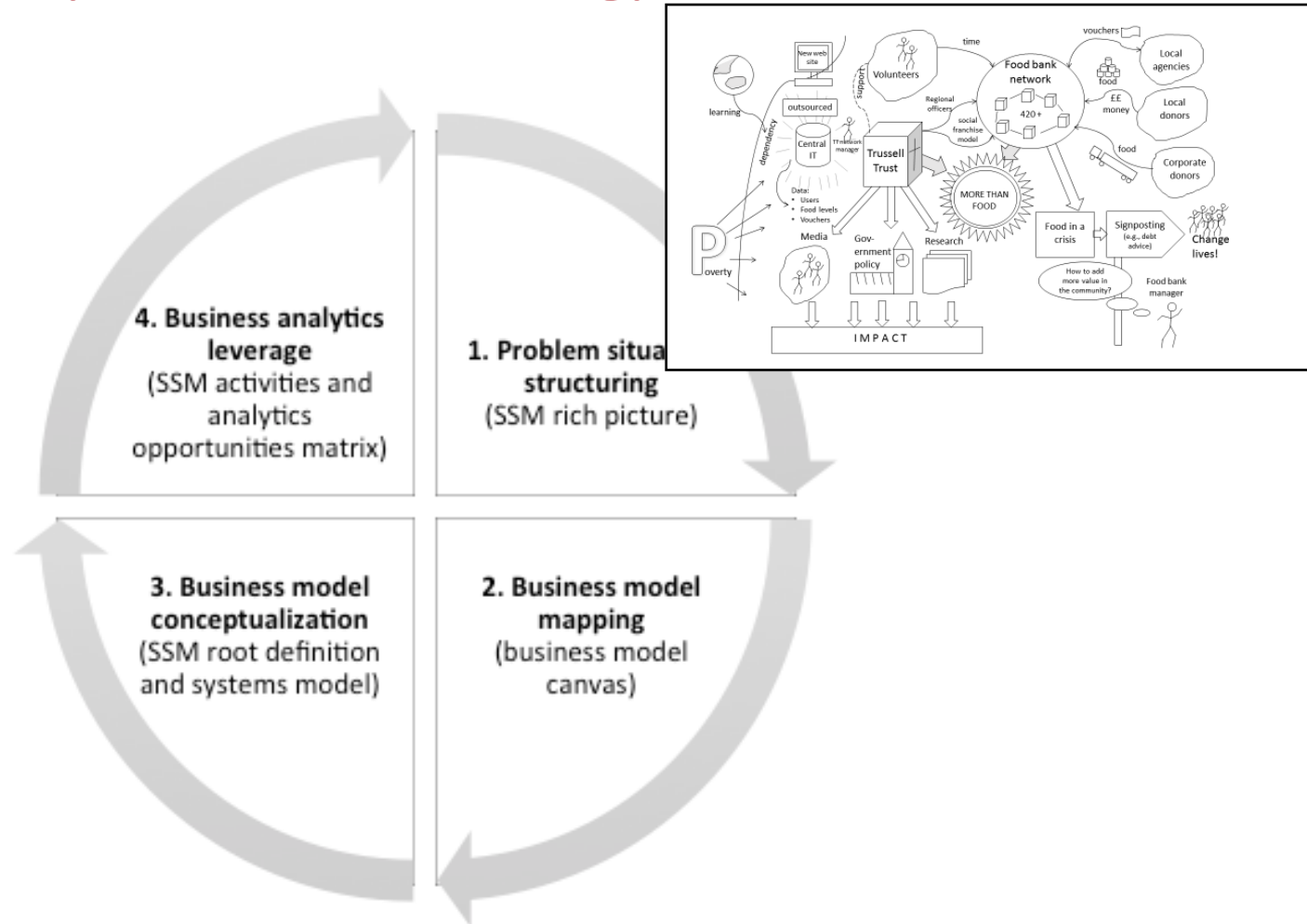
Analytical Work



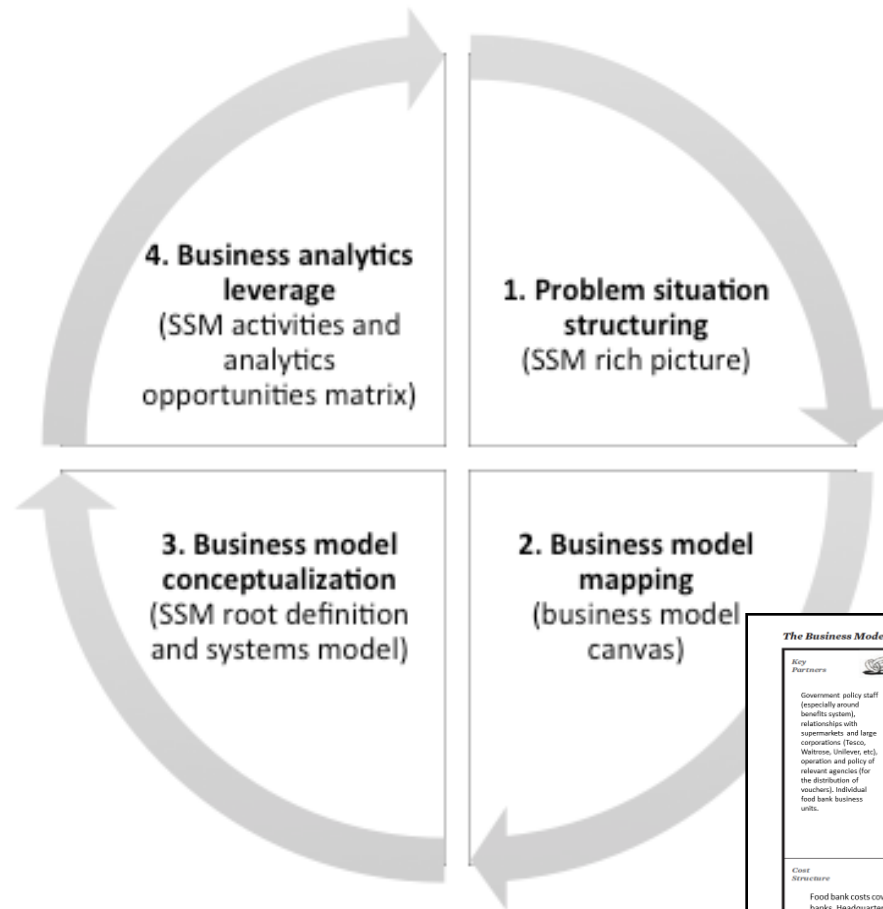
Business Analytics Methodology



Business Analytics Methodology



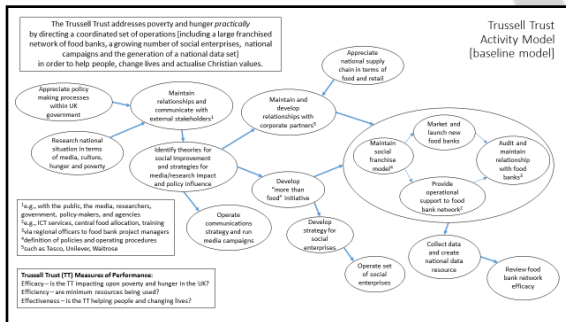
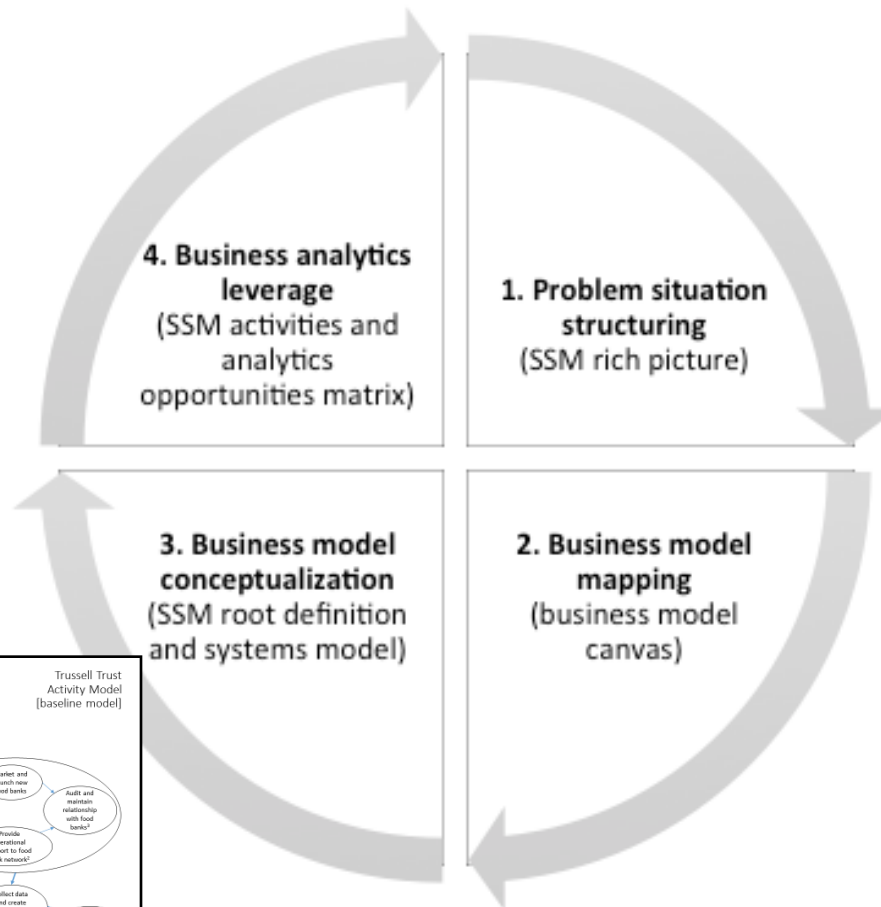
Business Analytics Methodology



The Business Model Canvas

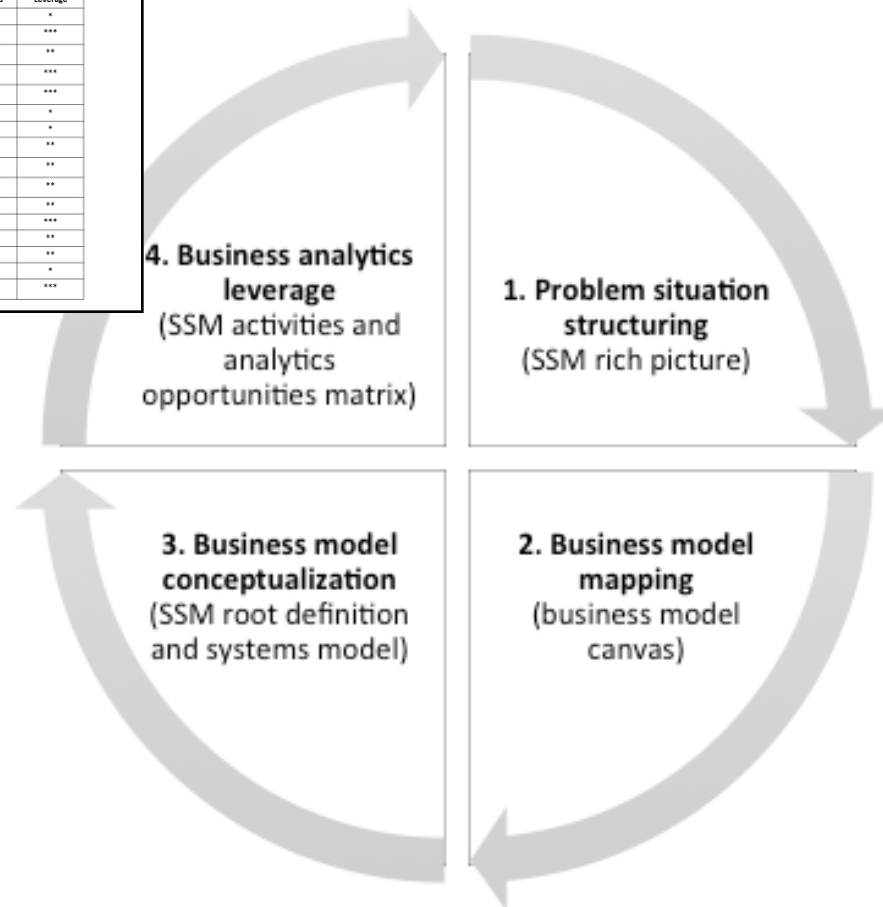
<p>Key Partners</p> <p>Government policy staff (especially around benefits system), relationships with supermarkets and large corporations (Tesco, Waitrose, Lidl, etc), operation and policy of relevant agencies for the distribution of vouchers, individual food bank business units.</p>	<p>Key Activities</p> <p>See Activity Model</p>	<p>Value Proposition</p> <p>The Trussell Trust network provides a social Enterprise model to its customers, which is similar to commercial franchises. This value proposition includes branding, business model and technical support for food bank boards and managers. All Trussell Trust food banks are evaluated annually by the Trussell Trust network to ensure compliance and good practice.</p>	<p>Customer Relationships</p> <p>Trussell Trust is a charity founded on Christian principles, with the foodbanks it supports primarily church-led. Foodbanks managed through regional officers. Annual conference run by the Trussell Trust for stakeholders, including volunteers.</p>	<p>Customer Segments</p> <p>Communities who wish to launch and operate food banks that provide three days of emergency food and supporting to individuals and families in crisis. The Trussell Trust customers will be a food bank board and food bank manager. TT have 1400 food banks, run by 420 organisations.</p>
<p>Key Resources</p> <p>Headquarters staff, premises, volunteer population, team of regional officers and food bank managers, donors and partners, branding, database, media relationships.</p>	<p>Channels</p> <p>The banks and food bank managers are managed through a set of regional officers, with a central IT service.</p>	<p>Cost Structure</p> <p>Food bank costs covered by individual food banks. Headquarters overhead (IT systems).</p>	<p>Revenue Streams</p> <p>Annual payments from food bank franchises, donations from individual donors and partners.</p>	

Business Analytics Methodology



Business Analytics Methodology

Activity from Baseline Activity Model	Potential Use of Analytics	Potential for Analytics	Business Model Leverage
Appreciate policy making processes within UK government	Descriptive view of size and shape of policy areas.	**	*
Research national situation in terms of media, culture, hunger and poverty	Descriptive view of national situation. Causal links between poverty, hunger and other variables.	***	***
Maintain relationships with external stakeholders (public, media, researchers, government, and agencies)	Descriptive view of stakeholder metrics.	**	**
Identify theories for social improvement and strategies for media/research impact and policy influence	Descriptive view of impact. Causal links between social improvement and TT strategies. Theory testing experiments.	***	***
Operate communications strategy and run media campaigns	Descriptive view of communications activity. Causal links - communications activity and impact. Theory testing experiments.	***	***
Maintain and develop relationships with corporate partners	Descriptive view of relationship metrics.	**	*
Appreciate national supply chain in terms of food and retail	Descriptive view of supply chain.	**	*
Develop "more than food" initiative	Descriptive view of need for services. Causal links between social improvement and TT strategies. Theory testing experiments.	***	**
Develop strategy for social enterprises	Descriptive view of need for services. Causal links between social improvement and TT strategies. Theory testing experiments.	***	**
Operate set of social enterprises	Causal links between social improvement and TT strategies. Theory testing experiments.	**	**
Maintain social franchise model	Descriptive view of network performance.	**	**
Market and launch new food banks	Descriptive view of need for services and network coverage.	***	***
Provide operational support to food bank network	Operational analytics - logistics, inventory.	***	**
Audit and maintain relationship with food banks	Descriptive view of food bank performance.	**	**
Collect data and create national data resource	Descriptive view of data usage within TT.	*	*
Review food bank network efficacy	Descriptive view of network performance. Causal links between poverty, hunger and TT strategies. Theory testing experiments.	***	***

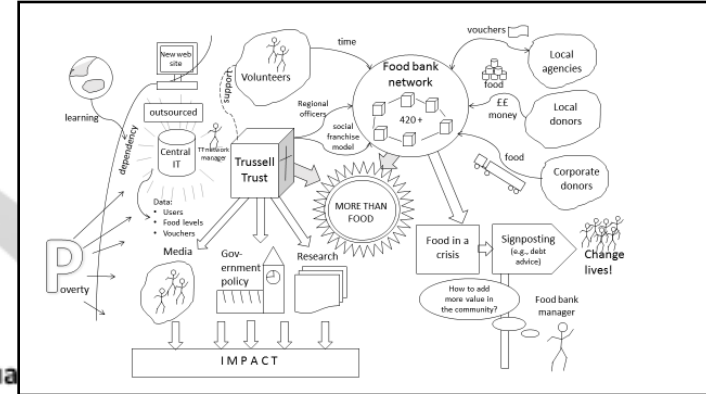


Business Analytics Methodology

Activity from Baseline Activity Model	Potential Use of Analytics	Potential for Analytics	Business Model Leverage
Appreciate policy making processes within UK government	Descriptive view of size and shape of policy areas.	**	*
Research national situation in terms of media, culture, hunger and poverty	Descriptive view of national situation. Causal links between poverty, hunger and other variables.	***	***
Maintain relationships with external stakeholders (public, media, researchers, government and agencies)	Descriptive view of stakeholder metrics.	**	**
Identify theories for social improvement and strategies for media/research impact and policy influence	Descriptive view of impact. Causal links between social improvement and TT strategies. Theory testing experiments.	***	***
Operate communications strategy and run media campaigns	Descriptive view of communications activity. Causal links - communications activity and impact. Theory testing experiments.	***	***
Maintain and develop relationships with corporate partners	Descriptive view of relationship metrics.	**	*
Appreciate national supply chain in terms of food and retail	Descriptive view of supply chain.	**	*
Develop "more than food" initiative	Descriptive view of need for services. Causal links between social improvement and TT strategies. Theory testing experiments.	***	**
Develop strategy for social enterprises	Descriptive view of need for services. Causal links between social improvement and TT strategies. Theory testing experiments.	***	**
Operate set of social enterprises	Causal links between social improvement and TT strategies. Theory testing experiments.	**	**
Maintain social franchise model	Descriptive view of network performance.	**	**
Market and launch new food banks	Descriptive view of need for services and network coverage.	***	***
Provide operational support to food bank network	Operational analytics - logistics, inventory.	***	**
Audit and maintain relationship with food banks	Descriptive view of food bank performance.	**	**
Collect data and create national data resource	Descriptive view of data usage within TT.	**	*
Review food bank network efficacy	Descriptive view of network performance. Causal links between poverty, hunger and TT strategies. Theory testing experiments.	***	***

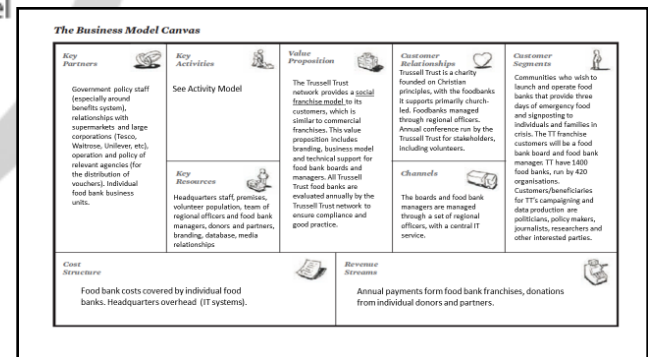
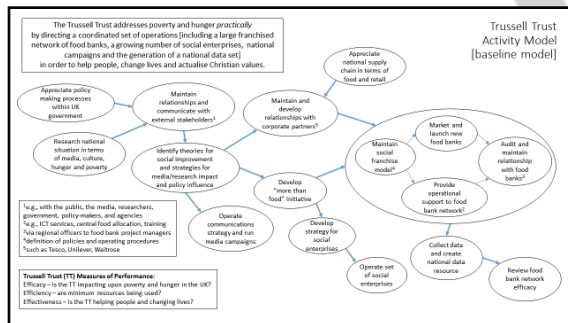
4. Business analytics leverage
(SSM activities and analytics opportunities matrix)

1. Problem situation structuring
(SSM rich picture)



3. Business model conceptualization
(SSM root definition and systems model)

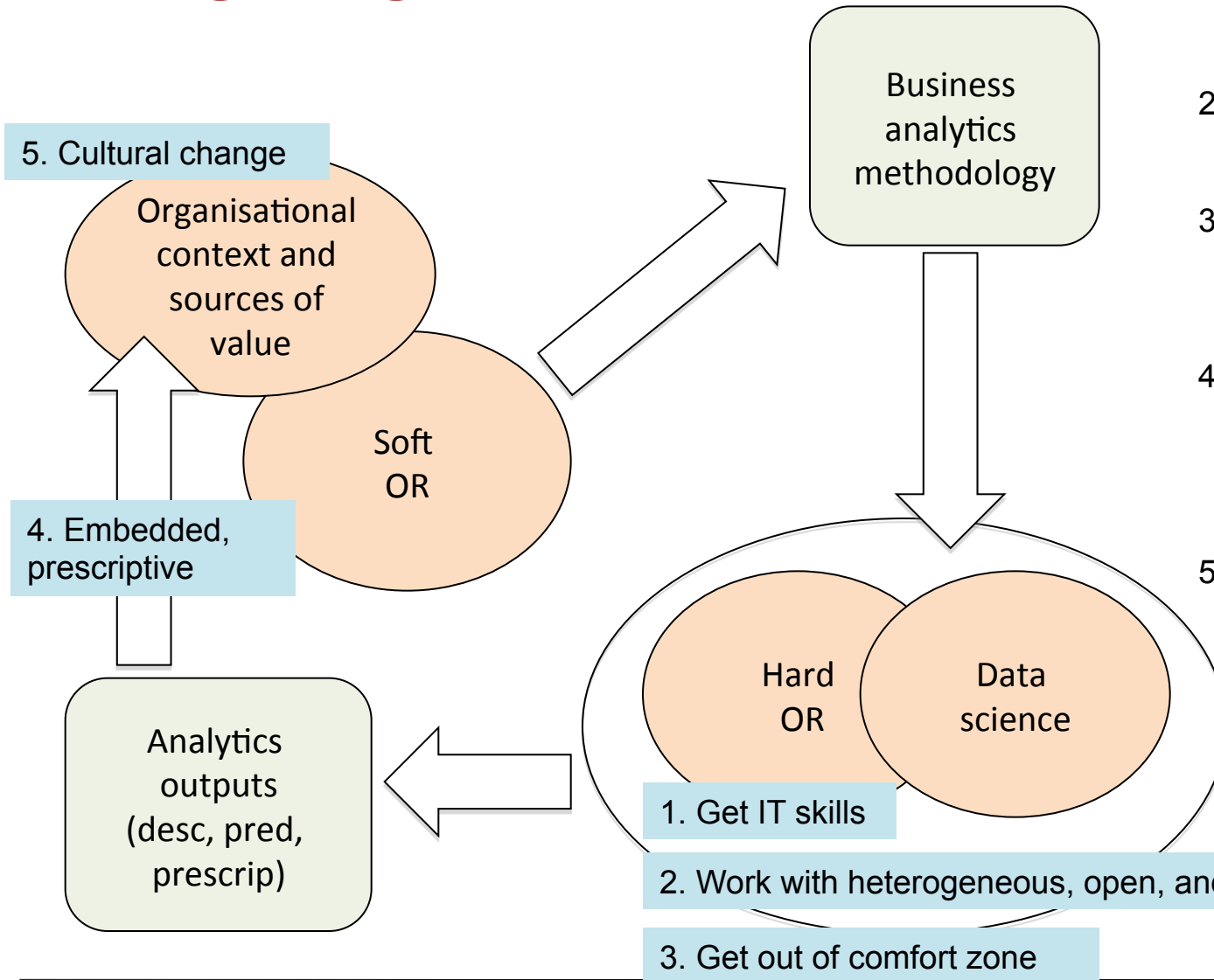
2. Business model mapping
(business model canvas)



Food bank app – multiple technologies

- To get the base data on food bank usage
 - SQL
 - To scrape Census data and get Google maps travel times
 - Python
 - To prepare data, create predictive models, and create data for visualization
 - R
 - To render the geospatial visualization as an interactive app for the user
 - D3
 - topoJSON
 - GDAL
 - Openstreetmap
 - Javascript
-

Putting it together



1. OR practitioners need IT skills – need to know where to start and where to stop
2. OR practitioners need to work with different data types, e.g., text
3. OR practitioners need to step out of specialist departmental niche (engineering mind set)
4. Analytics is embedded in the organization – not a separate function (prescriptive as well as predictive)
5. Analytics is about organizational transformation – culture change is needed to become data-driven

Questions and discussion

