



5th Malaysia Statistics Conference

29 November 2017

Sasana Kijang, Bank Negara Malaysia

2017

From Data to Knowledge : The Journey

Statistical Standards, Methodologies, and Applications in Data Management and Usage

Data To Insights

Julia Goh
UOB



5th Malaysia Statistics Conference

Characteristics Of Good Data

Analysis

- Quality
- Accuracy
- Unbiased
- Comprehensive
- Timeliness
- Availability and accessibility



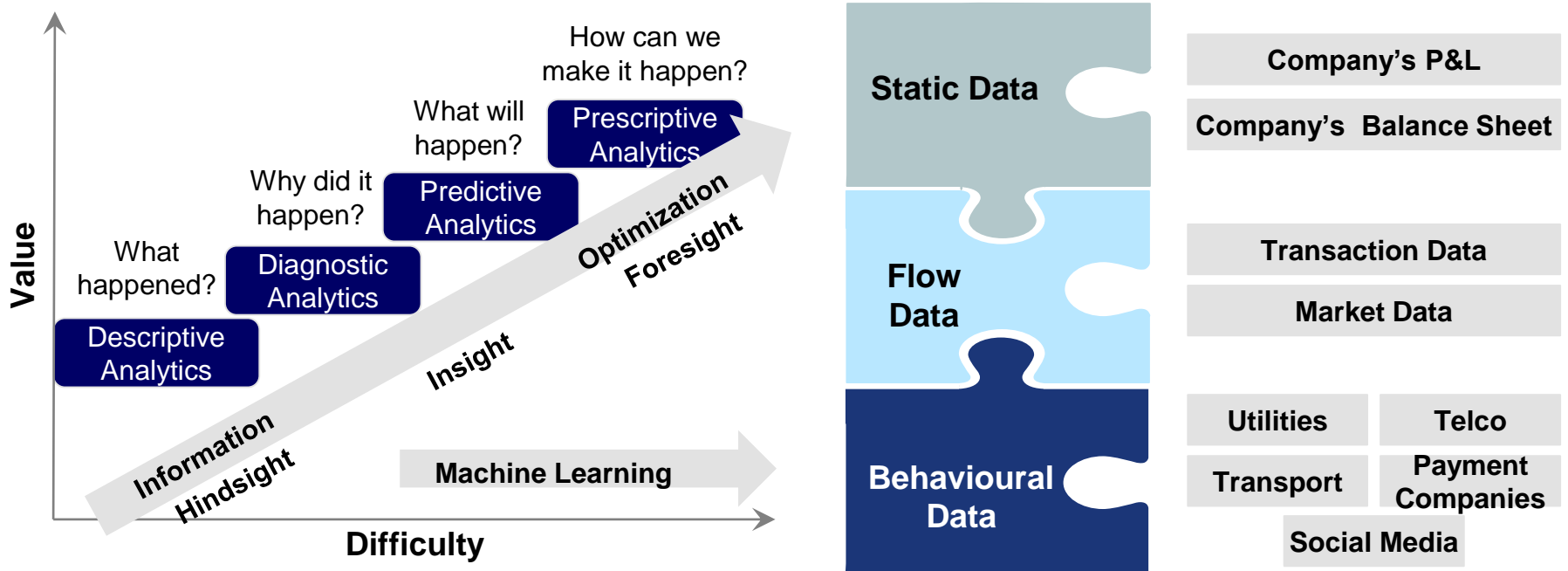
Forecasting

- Completeness
- Comparability
- Consistency



Data Is The New Currency

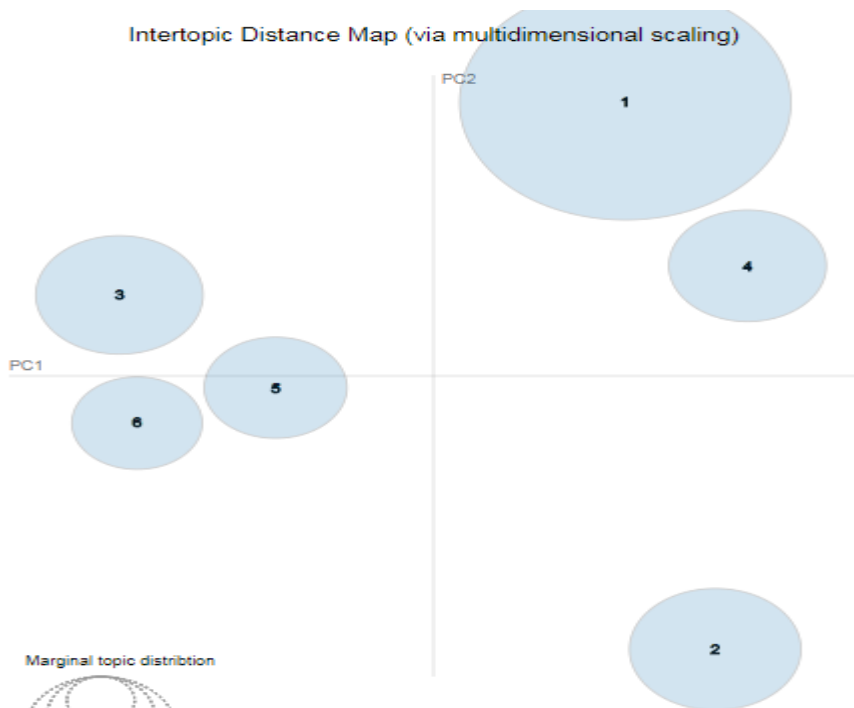
The Evolution of Data Analytics



Source: Gartner, UOB

MY Budget Speech - Topic Modelling

Intertopic Distance Map (via multidimensional scaling)

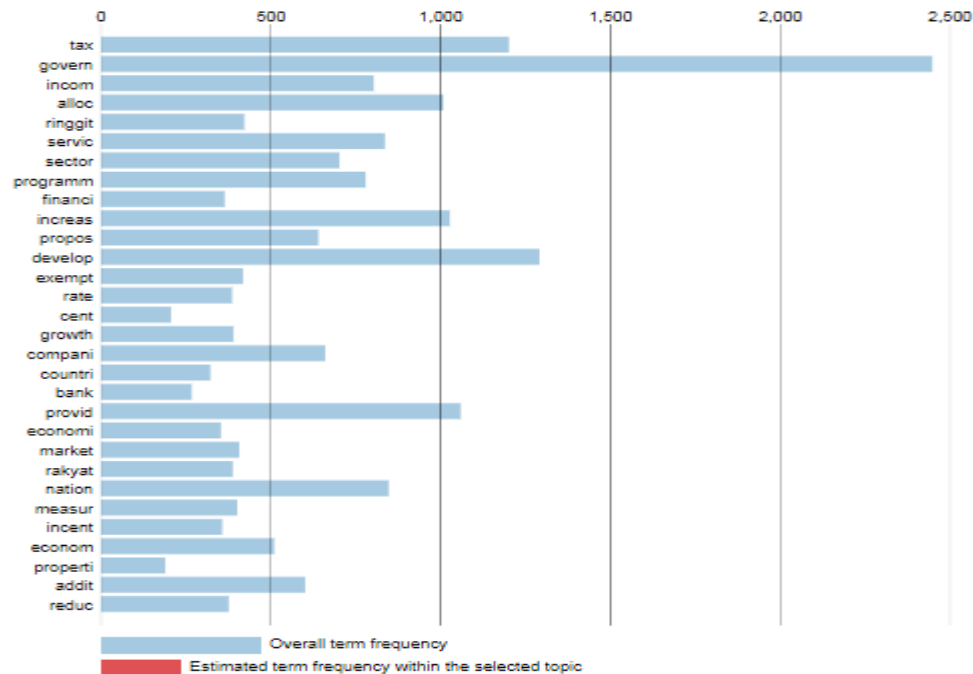


Marginal topic distribution



Source: UOB DMO

Top-30 Most Salient Terms¹



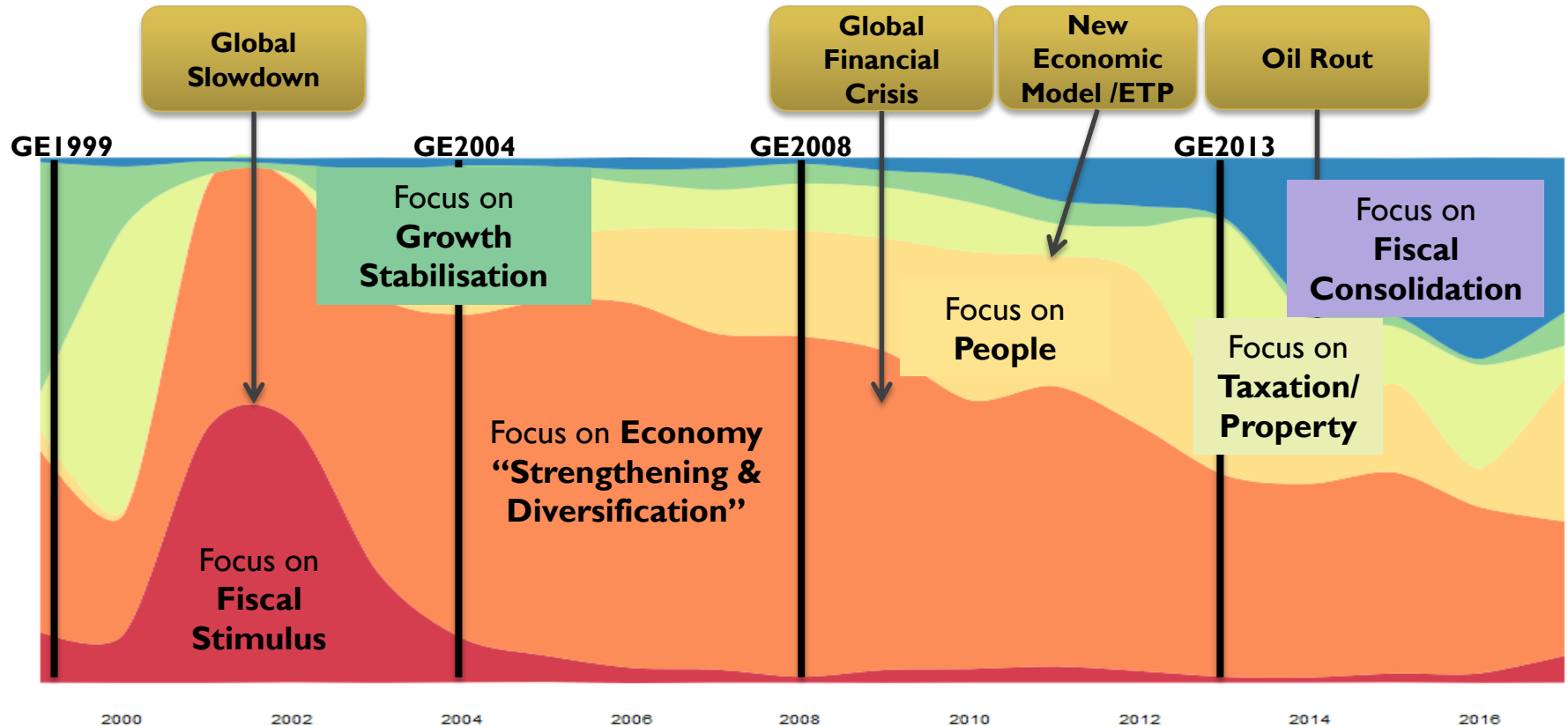
1. saliency(term w) = frequency(w) * [sum_t p(t | w) * log(p(t | w)/p(t))] for topics t; see Chuang et. al (2012)
 2. relevance(term w | topic t) = λ * p(w | t) + (1 - λ) * p(w | t)/p(w); see Sievert & Shirley (2014)

Words Describing Topic

1	2	3	4	5	6
RINGGIT	GOVERN	ALLOCATE	TAX	COUNTRIES	ALLOCATE
CENT	DEVELOP	GOVERN	INCOME	FINANCIAL	SERVICE
INCOME	PROVIDE	PROGRAMME	RATE	BANK	PROGRAMME
TAX	INCREASE	RAKYAT	PROPOSAL	ECONOMY	GST
ALLOCATE	SERVICE	ECONOMIC	EXEMPT	TRADE	BUDGET
ECONOMIC	NATION	SCHOOL	COMPANIES	LOAN	ASSESS
INCREASE	SECTOR	ASSIST	CURRENT	ECONOMIC	GOVERN
GROWTH	IMPLEMENT	RURAL	PROPERTIES	SECTOR	GOOD
PROPOSAL	ADDITION	UPGRADE	HOUSE	SHARE	INCENTIVE
EXPORT	CONTINUALLY	INNOVATE	INCENTIVE	POLICIES	RELIEF
REDUCE	INDUSTRI	SERVANT	EFFECT	MARKET	NATION
NATION	PRIVATE	HOUSE	INDIVIDUAL	INTEREST	MEASURE
PROGRAMME	INVEST	DEVELOP	RAKYAT	CRISIS	COST
GRANT	ENSURE	STUDENT	SCHEME	SYSTEM	PROJECT
PRODUCT	MALAYSIAN	GRANT	STAMP	MEASURE	SABAH
ACTIVATE	COMPANIES	BENEFIT	BUSINESS	PRODUCE	SUPPLIED
RATE	PROJECT	PROJECT	FINANCE	INTERNAL	ADDITION
GOODS	MARKET	AREA	ASSESS	STRATEGIC	EDUCATE
YOUTH	PRODUCT	INTRODUCE	POSITION	FUND	IMPLEMENT
ECONOMY	ENHANCE	CIVIL	DUTIES	CURRENCIES	SARAWAK

Source: UOB DMO

Topic Modelling On Budget Speeches



Source: UOB DMO, UOB Global Economics & Markets Research