

Big Data for Measuring the Information Society

Mohammad Ahli
Executive Director – Statistic Sector



Federal Authority | هيئة اتحادية

Big Data for Measuring the Information Society

Project scope

Utilizing big data from telecom industry (MNO & ISP) to improve and complement existing statistics and methodologies to measure the information society

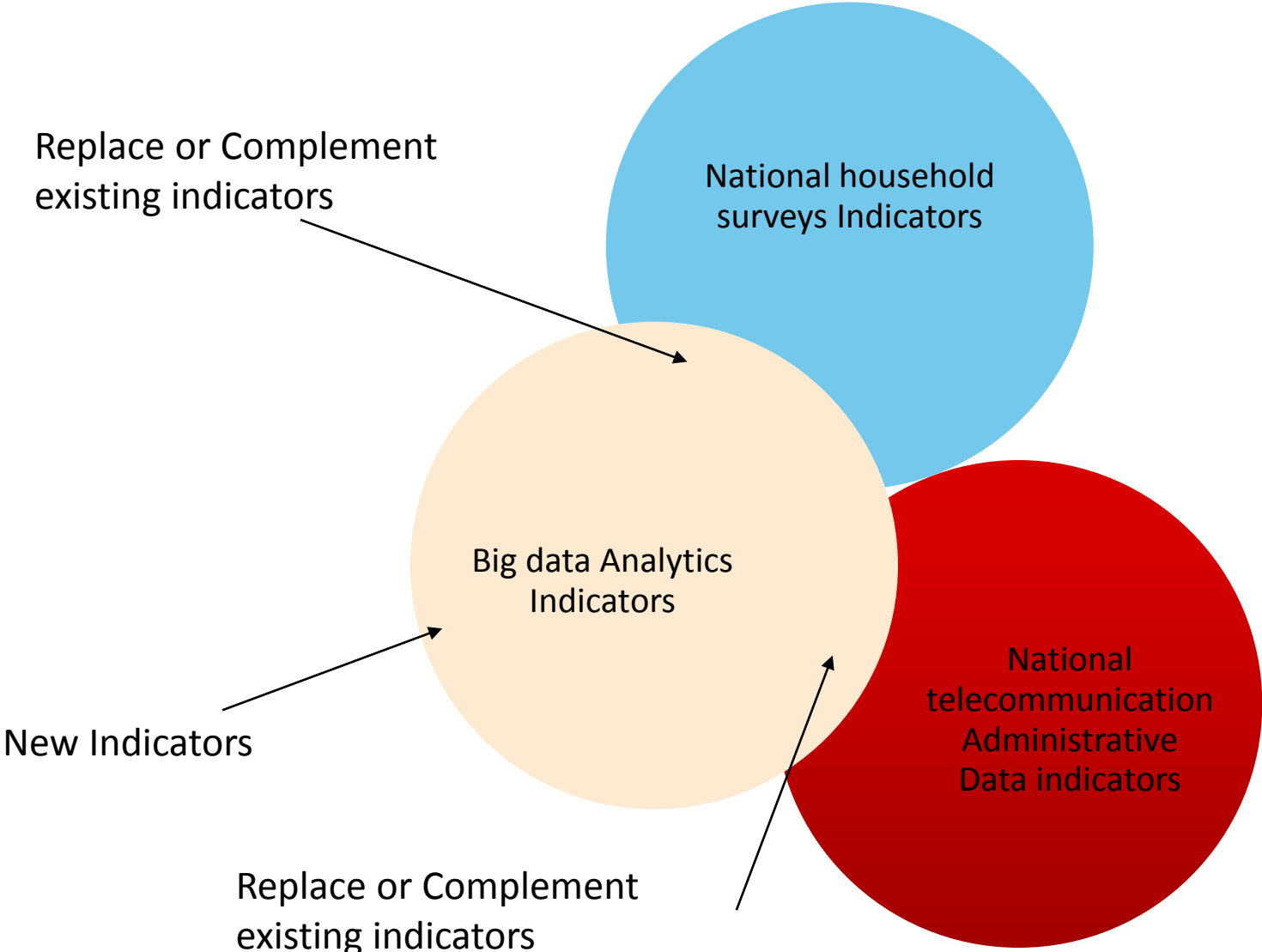
Project objective

Using Big data to produce new and existing official statistics ICT indicators to enhance data collections, benchmarks and methodologies to measure the information society.

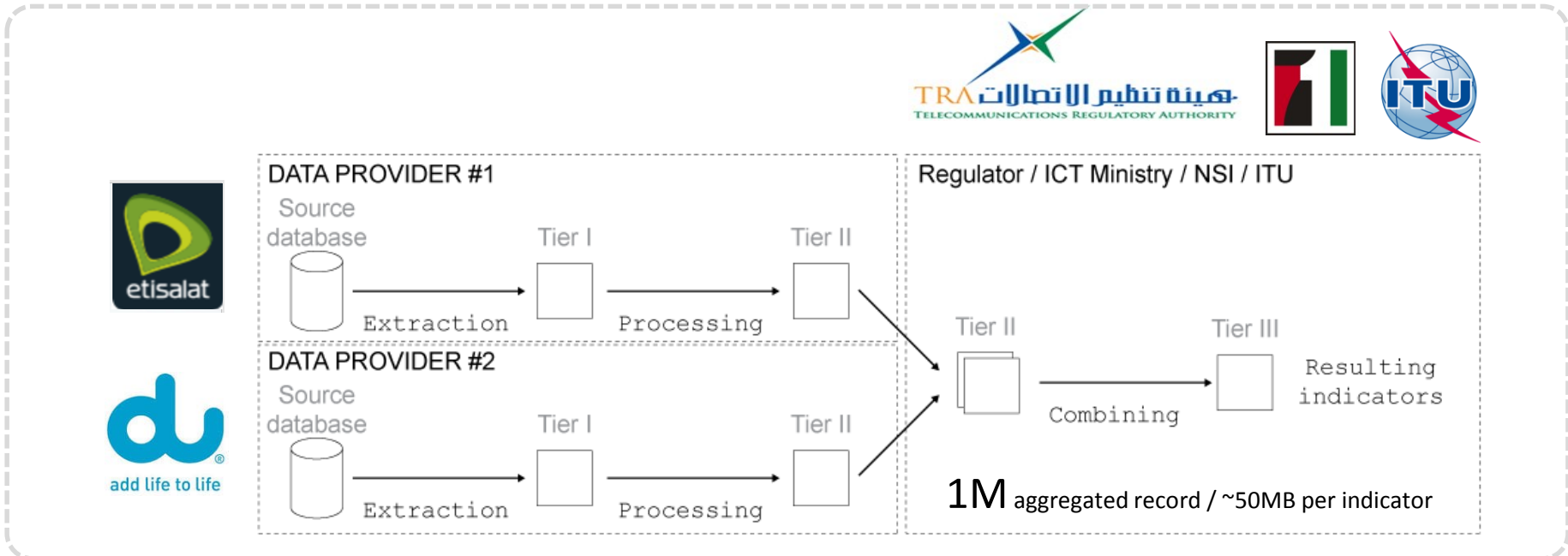
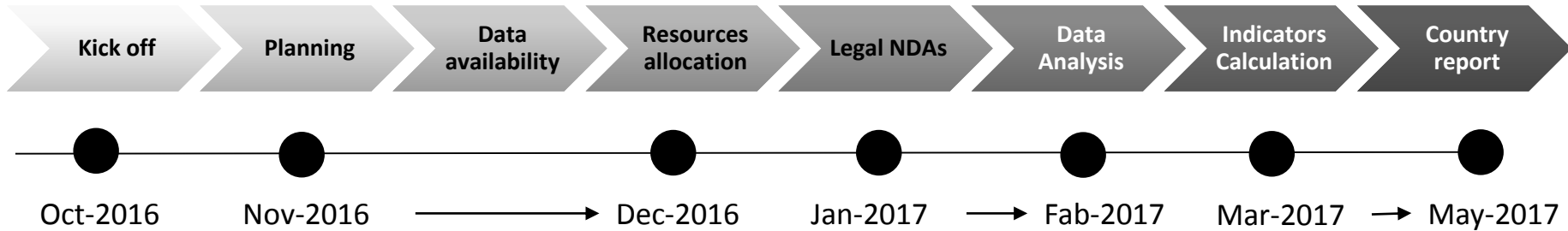
Stakeholders

- Telecommunication Regulatory
- National Statistical Office
- Telecommunications Service Providers

The role of Big data in the development of ICT access and use indicators



UAE participation in ITU Big data pilot project



United Arab Emirate Big Data for Measuring the Information Society Pilot Project Summery

5 Partners



44+ Trillion

Event as initial raw data by both data providers combined.

1+ Million

Consolidated and aggregated data record /~50MB per indicator

100%

UAE Coverage

Call Detail Records (CDR) and Internet Protocol Data (IPDR)

13 Enhanced

ICT Indictors using Big data , one considered new for UAE (DB03)

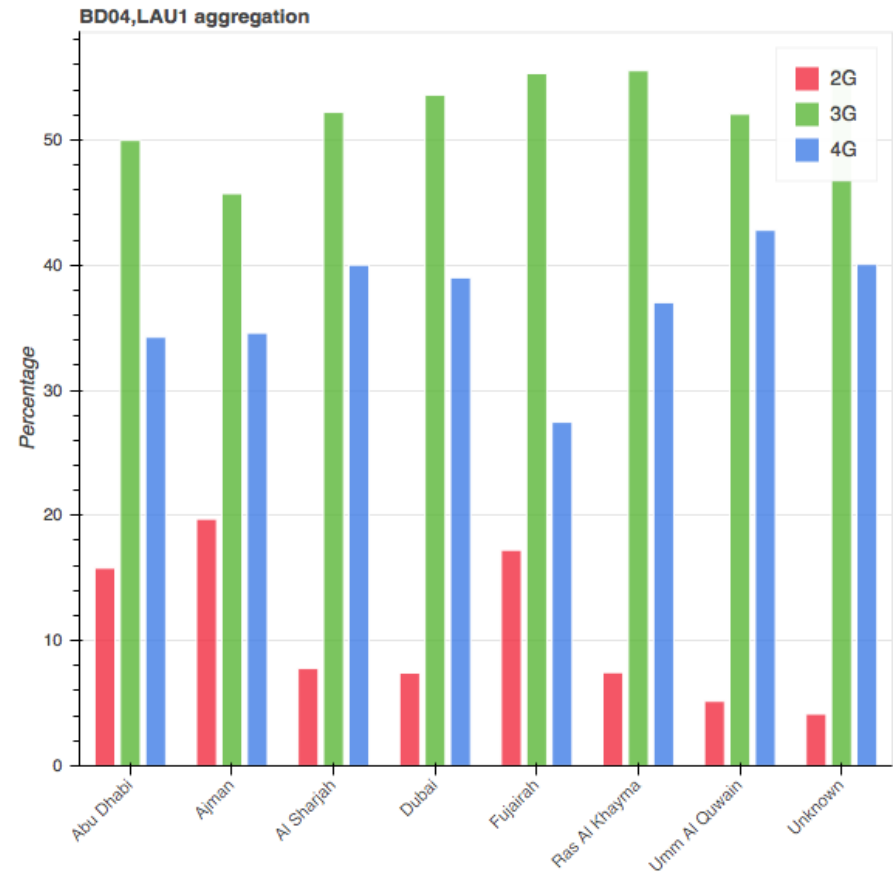
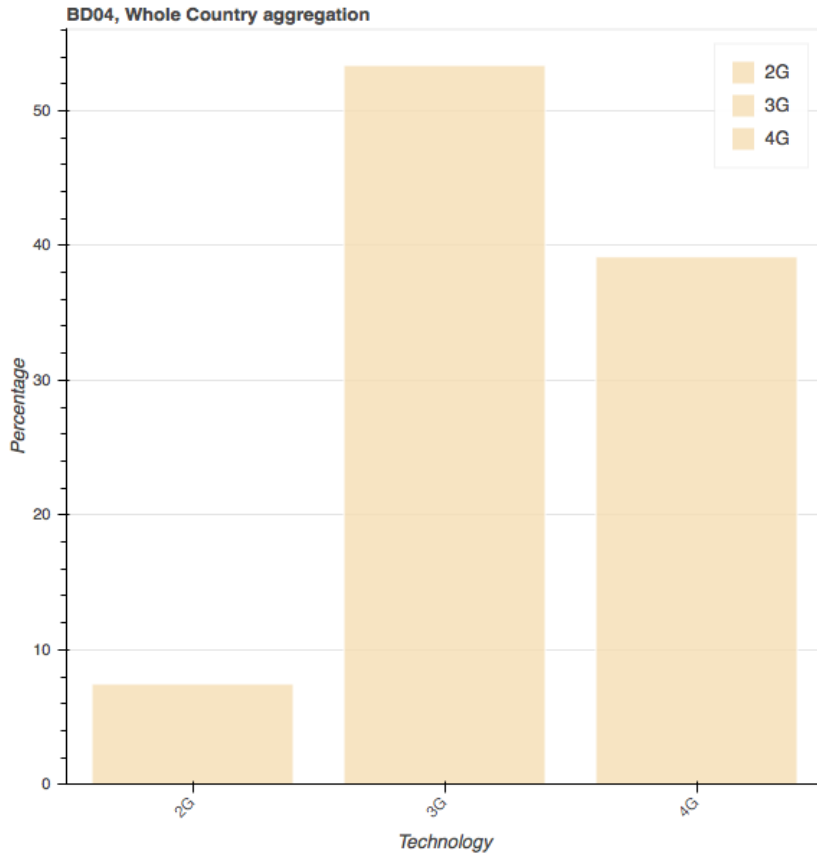
1 New

ICT Indictor (Origin/Destination Matrix)

- Both service providers extracted the initial raw data from their data warehouses, relying on big data processing technologies using different tools (IBM , Informatica) and Data warehouse appliances (Teradata, Netezza).
- ITU Data scientist worked with Two TSP technical teams (2-3 persons) each and statistician.

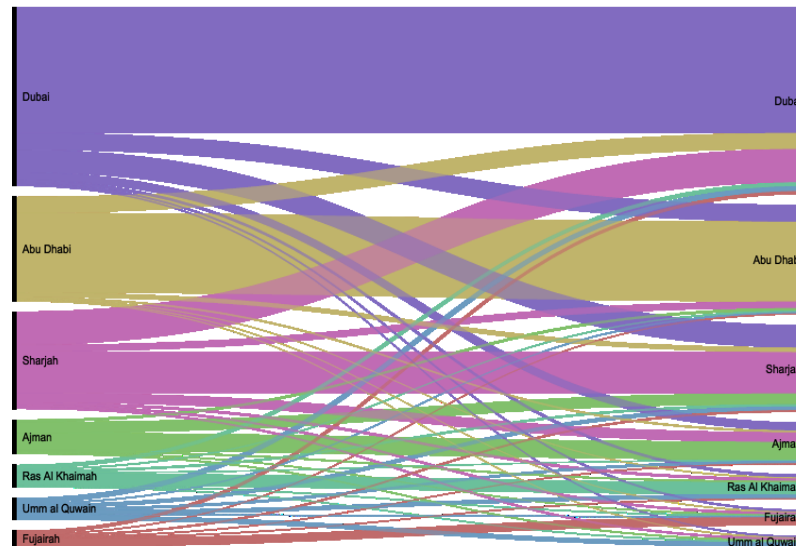
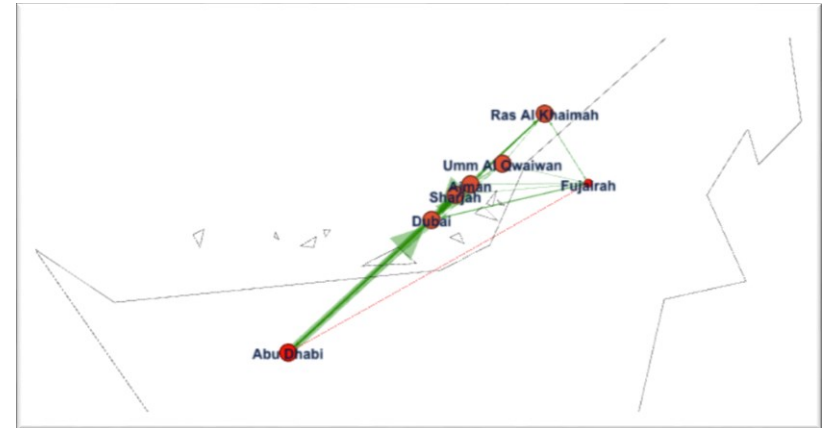
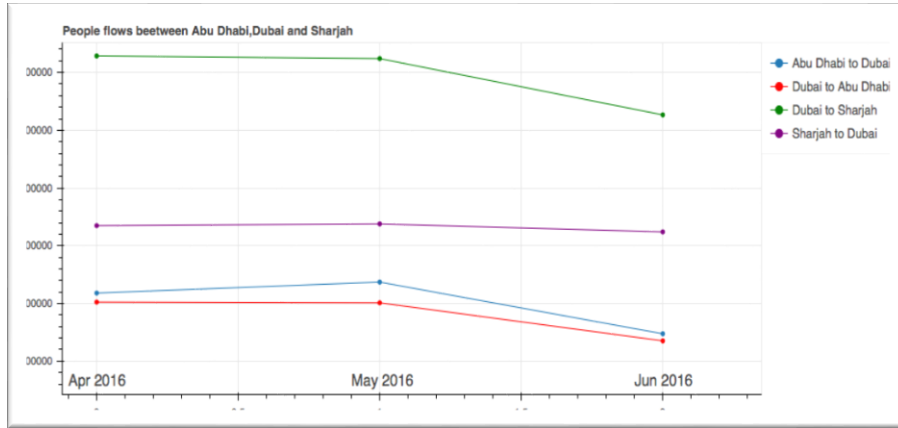
Results

Big data analytics enables service providers and government to monitor the progress of development of the technologies and to support the decisions for investing in regions that are lacking behind



DB04: Distribution of the mobile Internet access technology for UAE

BD16: Human Mobility



Origin/Destination Matrix for UAE administrative regions

Main indicators for ICT Access and Use by Households and Individuals

● Included
 ● Excluded
 ● Big Data (BD)
 ● Administrative data(AD)
 ● Household Surveys (HH)

Ind.	Key ICT Performance Indicators (KPI)	HH	AD	BD
BD01 ●	Percentage of the land area covered by mobile-cellular network, by technology		●	
BD02 ●	Percentage of the population covered by a mobile-cellular network, by technology		●	
BD03 ●	Usage of mobile-cellular networks for non-ip related activities, by technology			●
BD04 ●	Usage of Mobile-Cellular Networks for Internet Access, by Technology			●
BD05 ●	Number of Subscriptions with Access to Technology			●
BD06 ●	Active Mobile Voice and Broadband Subscriptions, by Contract Type		●	●
BD07 ●	Average Number of Active Mobile Subscriptions per Day, by Contract Type			●
BD08 ●	Active Mobile Devices	●		●
BD09 ●	IMEI Conversion Rate			●
BD10 ●	Fixed Domestic Broadband Traffic, by Speed, Contract Type		●	●
BD11 ●	Mobile Domestic Broadband Traffic, by Contract Type, Technology		●	●
BD12 ●	Mobile International Broadband Traffic, by Contract Type		●	●

Main indicators for ICT Access and Use by Households and Individuals

● Included
● Not fully
● Big Data (BD)
● Administrative data(AD)
● Household Surveys (HH)

	Big data ICT Indicators	HH	AD	BD
BD13	Inbound Roaming Subscriptions per Foreign Tourist		●	●
BD14	Fixed Broadband Subscriptions, by Technology		●	●
BD15	Fixed Broadband Subscriptions, by Speed		●	●
BD16	Additional Indicator: Origin/Destination Matrix			●

Big data can replace or complement existing several National household surveys Indicators

	Big data ICT Indicators	Ex: National household surveys Indicators
BD08	Active Mobile Devices	HH03 :Proportion of households with telephone
BD14	Fixed Broadband Subscriptions, by Technology	HH06 :Proportion of households with Internet
BD15	Fixed Broadband Subscriptions, by Speed	HH07: Proportion of individuals using the Internet
		HH10: Proportion of individuals using a mobile cellular telephone
		HH11: Proportion of households with Internet, by type of service
		HH12: Proportion of individuals using the Internet, by frequency

Challenges

- **Administrative and legal**

- Absence of standard legal and administrative procedures in place regulate TRA & NSO access to TSP big data.
- The need for identifying and signing the non-disclosure agreement (NDA) for external data scientist
- Service providers confidentially

- **Technical and methodological**

- Limited or missing data for some indicators.
- Unification of data formats, standardization and preprocessing were a time consuming task.
- Developing methodology, algorithms and validation methods is a multiple iteration process.

Recommendations and Lessons Learned

- Usage of big data complement and enhance official statistics, in addition to enabling the delivery of insights for leadership.
- Engaging TSP team in the big data analytics stage reduced the complexity of data security and confidentiality concerns and procedures.
- TRA (Telecommunication Regulatory Authority) role as the guardian of the service provider information confidentiality is crucial for project success.
- Formulating and engaging national data scientist team has a great added value for future big data analytics projects.
- Developing automation tool to pre-process acquired data with a format check and a preliminary analysis will reduce time and efforts with great impact on outcomes.
- Development of new model which protect data personal confidentiality.