Big Data for Early Childhood Development Outcomes

What will it Take?

Egbert Benstrong

Context

- From its inception the IECD has been a proactive and a firm believer in action – oriented decision making processes.
- For instance, in its first National Action Plan in 2013 , a set of nine core principles was articulated.

Principle 8: Evaluation and data collection

- Data and evidence are needed to monitor outcomes, advance understanding of local situations and guide effective policy response.
- Data on early childhood are gathered at different sources using diverse indicators.
- Many gaps and inconsistencies in data gathering across the different services were identified, and consequently
- Research on their impact was limited.

What has been happening since?

- IECD has initiated an number of audits and one in particular, targeted the child minding service providers (usually care for children aged from six months to 3 ½ years) in 2013
- The results from this survey assisted in the development and implementation of a Standards Framework for child minders

- A second example, has been the collaboration between the IECD and the Ministry of Health on two important projects:

(1) Monitoring Child Development Outcome in 2016

(2) Data Management in Early Childhood Care and Education, piloted in 2017

(1) DDST/Milestone

- Data collected on 959 and 1199 children (aged 3 ½ years) for the Milestone and DDST study respectively and also involved 1229 parents
- In this exercise it was possible to look at some milestone indicators such as birth weight, length, head circumference, and breastfeeding trends
- It was also possible to gauge parental perspectives not only about the child's development but also about the quality of the service provided by the Maternal and Child Health service providers.

(2) Data Management

- Data Management in Early Childhood Care and Education was piloted in 2017 and the collection of data for 2018 has just been completed
- the pilot was conducted with technical assistance from the World Bank Group

- Using WHO (2006) references as benchmarks, a set of 9 indicators were developed and used in the pilot. These were:

(1) Prevalence of low height – for – age in children under five years of age (stunting).

(2) Prevalence of low weight-for-height in children under five years of age (wasting).

(3) Prevalence of high weight-for-height in children under five years of age (overweight in young children).

(4) Prevalence of haemoglobin <11 g/dL in pregnant women is defined as the prevalence of a hemoglobin concentration of 11< g/dL in the blood of pregnant women.

(5) Anaemia among young children is defined as the prevalence of a haemoglobin concentration of < 11g/dL in the blood of children 6 through 59 months of age.

(6) Prevalence of infants born < 2500 g (low birth weight) is defined as babies born weighing less than 2500 g

(7) Proportion of pregnant women with gestational diabetes mellitus (gestational diabetes).

(8) Exclusive breastfeeding is defined as no other food or drink, not even water, except breast milk for the first six months of life.

(9) Prevalence of partial breastfeeding at 12 months

- For the pilot a total of 6848 infants and young children aged from 0 – 60 months were covered. It was found for example, that:
  - About 1 in 10 young children was stunted (low height-for-age).
  - About 1 in 10 young children was over weight (high weight – for-height)
  - About 43 percent of pregnant women were positive for anaemia
  - A similar percentage was also recorded for young children
  - About a quarter of infants aged 6 months or less were exclusively breastfed

- In education, as a third example, was the Assessment of the Early Learning Environment in Pre School (Creche) Setting conducted by IECD and the Department of Education in 2018
- One of the conclusions of the report, for instance, was that resource provisions were below the expected standard which could have an impact on early learning experience of the child

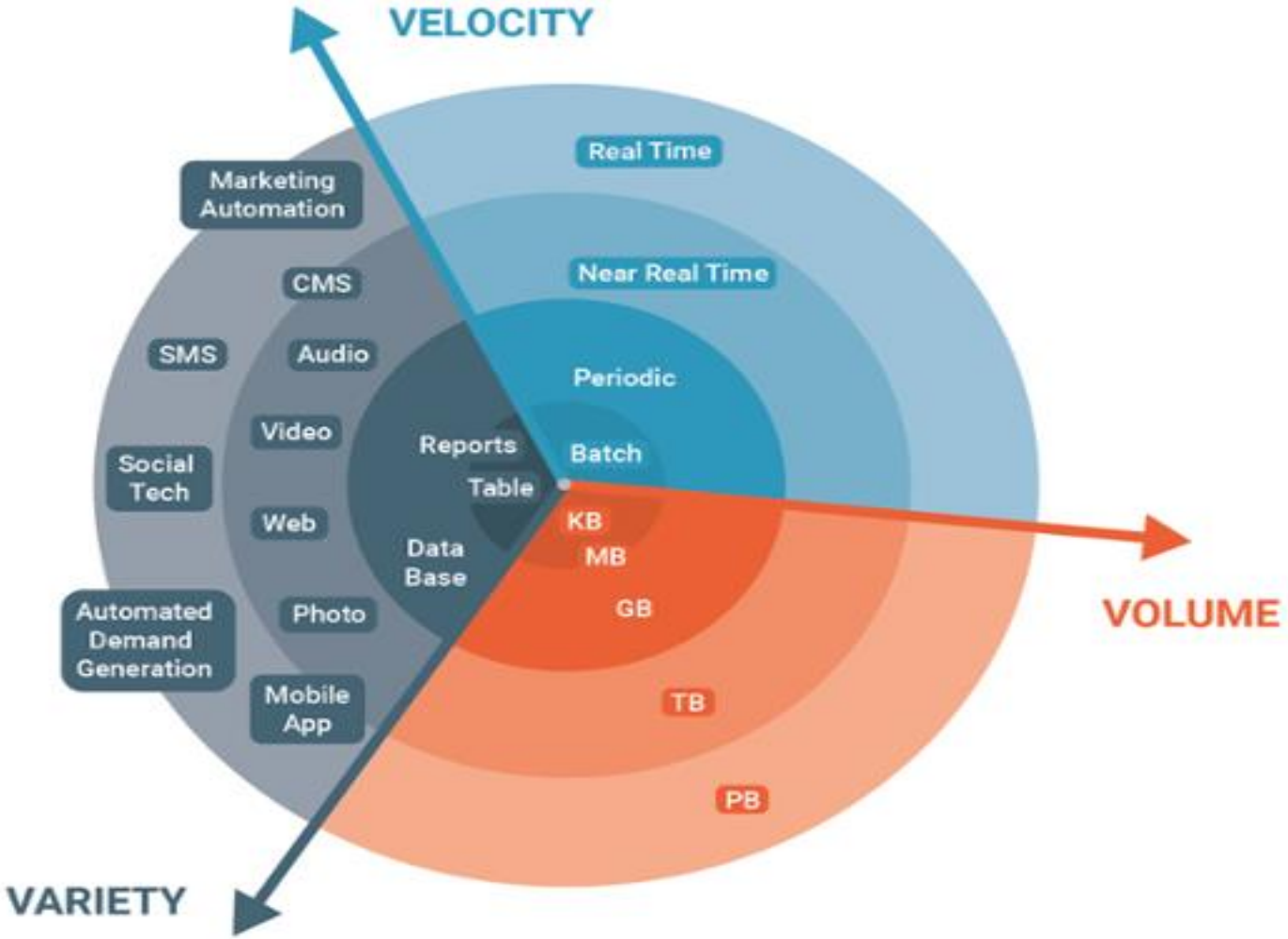
- Finally, and still in education, we have been assessing the level of Readiness for Creche or Pre-school for children who had experienced different early care provisions
- In Seychelles, some are cared for at home, some in child minding facilities and the rest in day care centres. All three have different care environments.
- This assessment has tremendous predictability analytics potential if linked to later national key stage testing

What will it take to improve development outcomes?

- Before addressing this question, it is perhaps important to understand the term “Big Data”
- Big data is gaining much prominence due to its ground-breaking contribution to decision making in many organisations

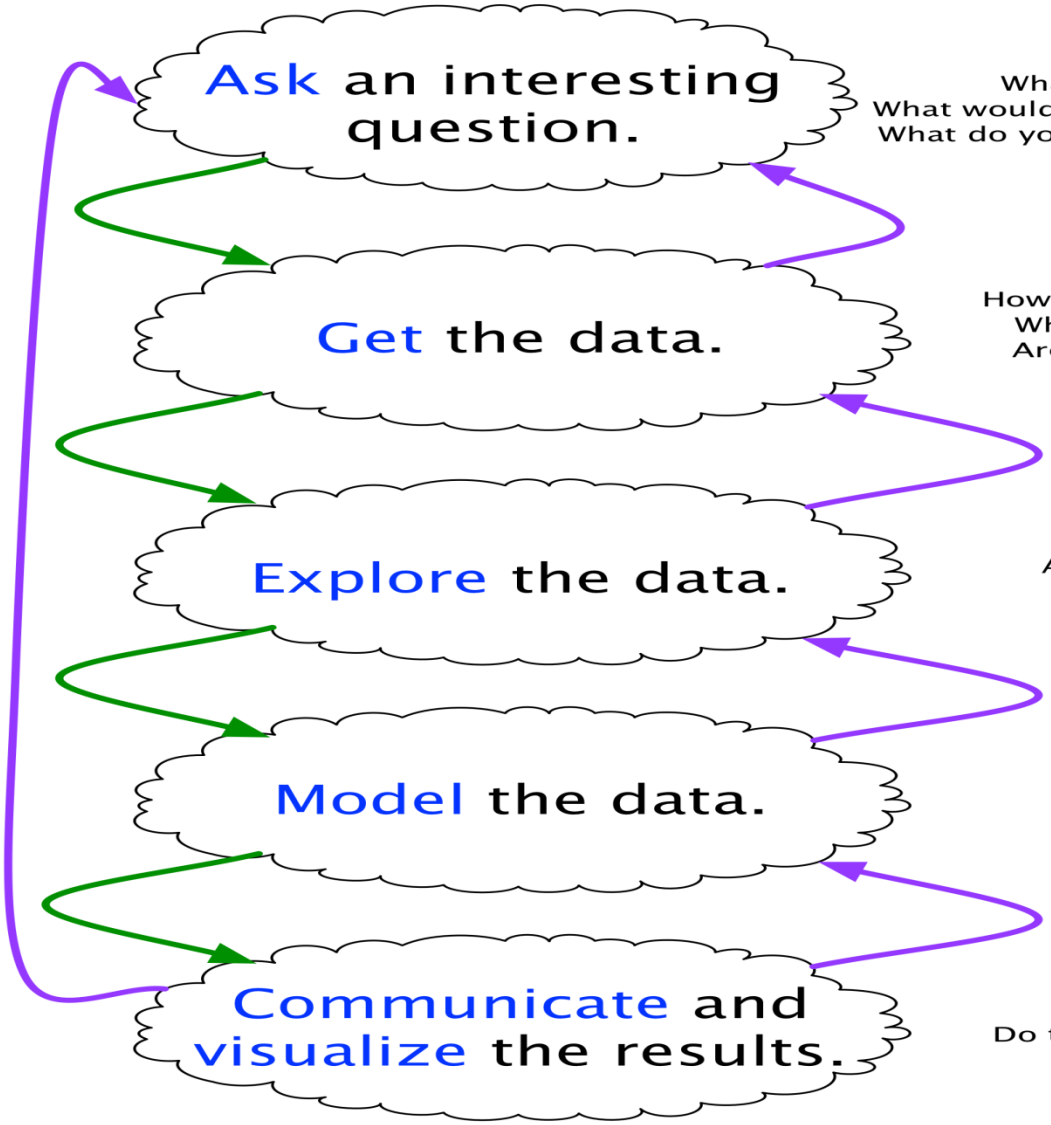
- Big data is larger, more complex data sets, especially from new data sources that we have (normally) been used to. These data sets are so voluminous that traditional data processing software just can't manage them (Oracle).
- Big data is sometimes described by the three Vs: **V**olume, **V**elocity and **V**ariety

The 3 V's of Big Data



Pickell, 2018

The Data Science Process



What is the scientific **goal**?
What would you do if you had all the **data**?
What do you want to **predict** or **estimate**?

How were the data **sampled**?
Which data are **relevant**?
Are there **privacy** issues?

Plot the data.
Are there **anomalies**?
Are there **patterns**?

Build a model.
Fit the model.
Validate the model.

What did we **learn**?
Do the results make **sense**?
Can we tell a **story**?

Joe Blitzstein and Hanspeter Pfister, created for the Harvard data science course http://cs109.org/.

Improving outcomes – what is required?

- The multi-sectoral nature of the early childhood care and development will potentially generate a huge volume of data
- To improve developmental outcomes therefore will require:

- Radical change in mind set - to accept that decision making needs to be data driven
- Procrastination and resistance to change
- Sing from the same page (cross- sectoral sharing of information guided by strong ethical principles)
- Training and professional development in big data analytics – we need a pool of trained data scientists!
- Investments in necessary technologies including information management systems, storage for big data
- Expansion into other EC sectors (Social and Community Development)

Conclusion

- Big data and data science are the new realities
- They call for a new thinking about the way we collect, manage, store and use data
- Leaving data on shelves or hard drives is no longer an option
- A winning start for all children should be achievable in the 21st century