A Journey from Big Data to Smart Data

Fernando Iafrate

Abstract. Nowadays Decision is a matter of information (related to the data), knowledge (related to people & processes) and timing (capacity to decide, act and react in the right timing). The huge increase of the data volume, velocity but also format (unstructured data like: blogs, logs, video…) generated by the "digitalization" of our world modifies radically our relationship to the space (in motion) and time (all the time) dimension and by capillarity, the enterprise vision of the performances monitoring and optimization. This paper introduces via three use cases, how the closed loop between Big Data & Smart Data embedded into the business processes can help the enterprises facing this new challenge.

Keywords: Big Data, Smart Data, Smart Data, Operational Performance Management (OPM), Operational Control Center (OCC), Yield Management, Bid Price, Business Process, Zero Latency Organization, Key Performance Indicator (KPI), Business Intelligence (BI).

1   Introduction

The concept of Big Data usually covers four notions (called the 4V): the data Volume, the data Velocity, the data Veracity, and the data Value, we decided to enhance this approach by splitting the 4V in two subset: Data Volume & Velocity for the Big Data (where we manage the real time information as the "actual" state), and data Veracity & Value for the Smart Data (where we manage the cross functional analytical information as the "future" state), the merge of these two concept called "closed loop" in this paper is the corner stone of our three use cases. Most of the enterprises are struggling with the management of their data, every day; the Digital World is generating new sets of data coming from different sources (the web is the main one) in different formats structured (can be directly

Fernando Iafrate
Senior Manager of the Business Intelligence & Data
Architecture Domain Disneyland
Paris (France)

loaded in a relational database) and unstructured (cannot be directly loaded in a relational database, like images, transaction logs, blogs…), lot of these data are holding "noise" (information or meta data having low or no real value for the enterprise).The purpose of our Smart Data (thanks to the analytical skill & tools) is to hold the valuable data of the enterprise and use them to interact in "real time" (most of the time in the transactional process activity and timing) using Business Intelligence applications.

2   A Journey from Big Data to Smart Data

As a closed loop enabled by automated Business Intelligence apps and organization in order to manage & monitor the performances of the enterprise, enabling a Zero Latency Organization.

2.1   What Is?

Big Data (according to the Gartner): is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making.

Smart Data: is in fact a subset of data (Big or not) valuable for the enterprise and cross functional.

Zero Latency Organization: is an organization where the decision to action latency is directly linked to the process latency you are monitoring/optimizing, it will imply the right "decision" KPI's (key performance indicators related to the business process), Business Intelligence tools (including the capacity to manage the right data velocity) and an organization in a decision for action processes.

Fig. 1 from Big Data to Smart Data

2.2   What For

The Digital World, is in reality more or less about managing "Data" (all kind of data), if you really want to have a serious advantage in this arena, you will need to have a strong skill on Data Management but also the right organization, processes and associated Business Intelligence apps in order to leverage the full value behind those data.

   The main goal is to move from a data organization (struggling with the data management) to a learning organization (leveraging all the value behind the data, with the right processes and organization)

2.3   Key Issues

What are the factors that lead to successful strategic deployment of information management in a context where we generate every day more data than what we can really manage/analyze without a strong Business Intelligence organization, automated processes & Business Intelligence apps?

2.4   What You Need to Know

We built a world-class Business Intelligence & Data Architecture environment, which provides customer-focused (before, during and after the visit) strategic and "real-time" (align with the business processes)   insight to a broad set of users, but also automated interaction driven by Business Intelligence apps connected to the "Big and Smart Data" analytical processes & tools. Business activity is predicted and continuously monitored against key performance indicators (KPI's).
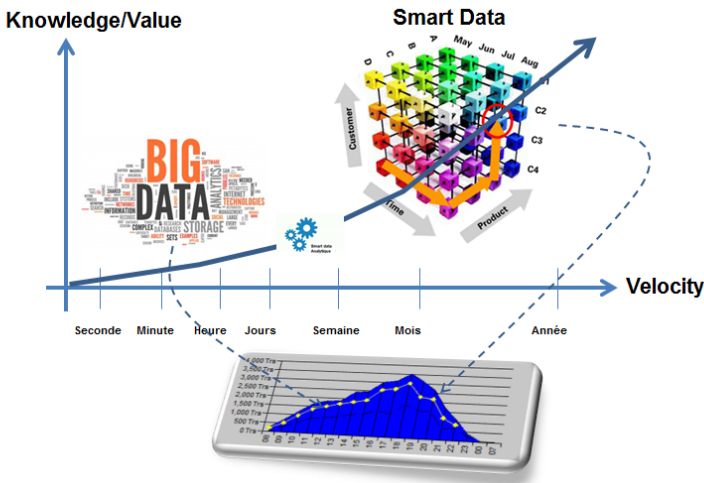

Fig. 2 from Big Data to Smart Data a "closed loop"

We have created a "closed loop" between the "Big Data" (for the real time information showing the current state) and the "Smart Data" (analytical model showing the expected state) in order to monitor our activities. By linking the two data concept (Big Data & Smart Data) we moved form a probabilistic forecast (issue from the Smart Data) having a certain level of probability to a more deterministic approach where the "current state" (extracted from the Big Data) is compared in "real time" (timing linked to the target process) to the "expected state" (forecast generated by the Smart Data) in order to monitor and optimize the current activity and the impacts of the decisions/actions. At the end of the day, the Big Data are then filters and sorted… (The map reduce model), included in the Smart Data (models) in order to enrich them, this closed loop generates an auto learning process, which is the basement of our marketing and sales knowledge management solution.

3   Use Cases

As the context has been defined, I would now like to introduce three cases of application in the enterprise of Business Intelligence applications linked to BIG & Smart Data.

- The customer relational process (discovery phase)
- The selling processes (price & availability phase)
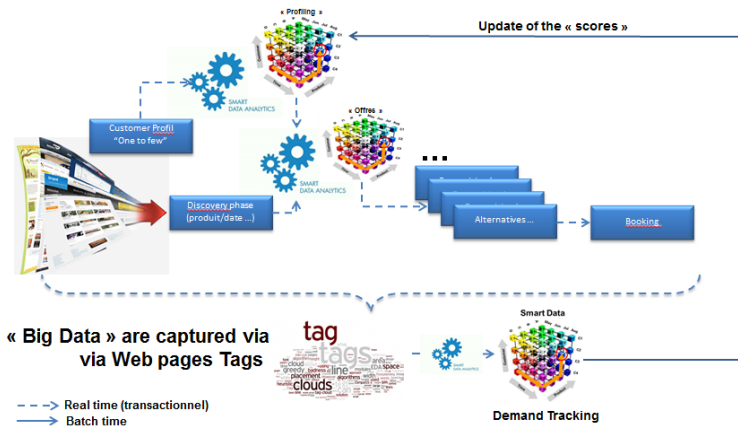- The operational processes (customer experience in the parks)

Fig. 3 Use case #1: Demand Tracking System

In this case, we have implemented an "auto learning" solution enabled by a Business Intelligence applications in order to propose on our E-commerce web site, the right product to the right customer at the right time.

This solution is dealing with:

–   Real time information coming from the web site (feeding the "on the fly" scoring process)
–   Predefine scores and products list modelized by customer segment enabling the "eligibility" (what product fits the customer request) product list for the scored segment.
–   The solution generate thousands of price and availability check in order to propose best offers & alternatives to the customer based on the current Bid Price (the bid price, is the minimum expected revenue in the request context: Arrival date/ Length of the stay/ Patry Mix / Hotel Inventory, ~2,8 Millions of bid price point are indexed in real time by using the remaining hotel capacity, more an inventory is constrainted, higher is the bid price)
–   The solutions are in an auto-learning mode enabled by the demand tracking "Smart Data" wich recalculate every day the score variance… based on the daily data captured from the web site (all the traffic is modelized, even if it generates no bookings)

Fig. 4 Use case #2: Bid Price concept for Dynamic Yield Management

The bid price concept aim to validate the product availability based on minimum expected revenue (the bid price) directly indexed with the inventory (room in this case) left to sell.

The expected product value (link to the customer request for an Arrival date/ Length of the stay/ Party mix/ Hotel inventory, is compared to the bid price (minimum expected revenue for this context), if the product value is >= to the bid price, the product will be available for sales, else the product will be consider as unavailable (even if we still have free inventories).

Every inventory update (hotel room in this case) by a sale (reduce availability) or a cancel (increase availability) will impact the bid price index, and by

capillarity will reference a new bid price dynamically, this new bid price will then be compare with the requested product value, in order to enable of disable the product availability. This mechanism will secure the intraday optimization by reacting in "real time" to the customer behavior (not only to what was forecasted the previous night in term of customer demand)

The bid price is generated by our yield management solution every day, we generate more than 2.8 Millions of bid price points (number of inventories x forecast horizon), see next figure (5) for more details.
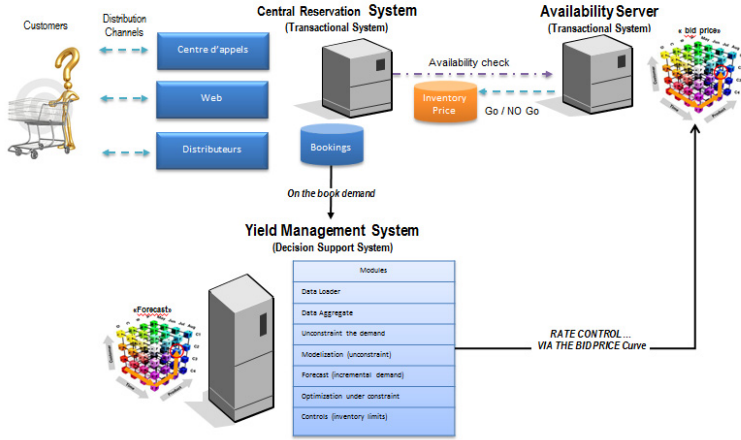

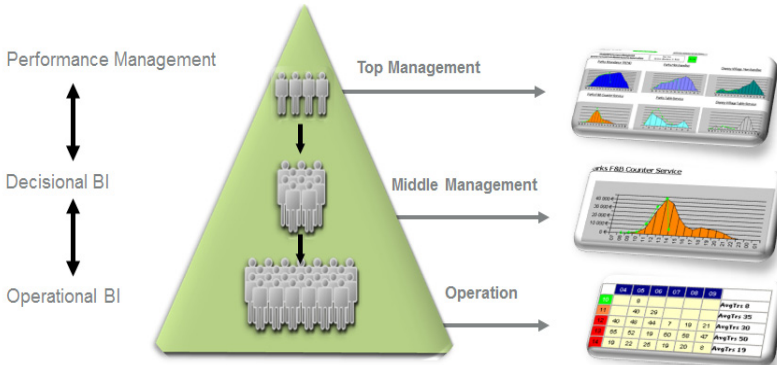**Fig. 5** Use case #2: what is behind the magic?

Fig. 6 Use case #3: the "Data" align connect all the level of the enterprise

In order to leverage the bid price table produced (every night) by the Yield Management solution (as an output of the demand forecast by product value), we had to implement an Availability Server. The Availability Server is embedded in the selling transactional process and activated during the availability and price check, the architecture is scalable and can hold thousands of availability & price request per minute, in order to support the organic growth of the digital activity on our E-Commerce web site.

The OPM (Operational Performance Management) aim to align and connect all the level of the enterprise by managing in "real time" (align with the timing of each one of the target processes) the operational performances of the enterprise.

The onsite operational performance of the enterprise will have a direct effect on the customer experience, and our company is the place where the "dreams comes true". In order to monitor and manage this operational performance, we have implemented an OPM (Operational Performance Management Solution).
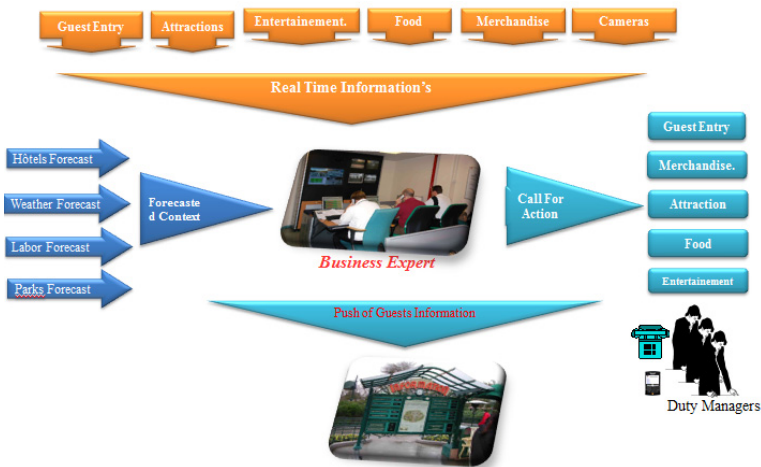

Fig. 7 Use case #3: the OCC (Operational Control Center) [Grundstein & Rosenthal-Sabroux, 02]

The OPM (Operational Performance Management) is a set of Business Applications publishing "real time" (link to the target process timing) information and alerts as KPI's for the operational processes, but also the right organization to leverage that information.

The OCC (Operational Control Center) is an example of an organization leveraging (via business experts of each one of the operational domains) the Big to Smart data information for the monitoring and the optimization of the operational performances of the enterprise.

4   Conclusion
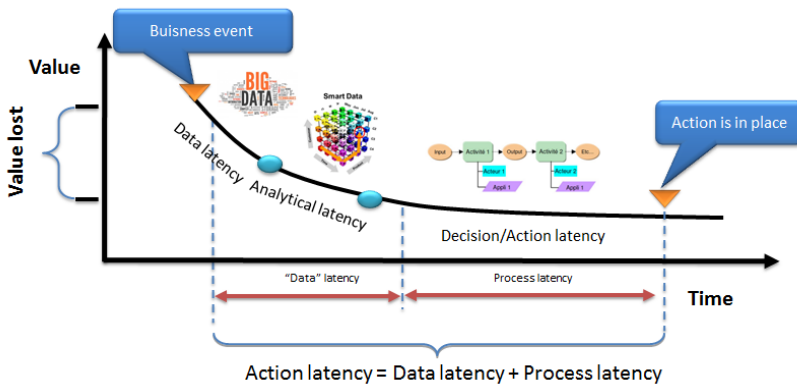
Fig. 8 Data value is a matter of timing and alignment with the business processes [Hacka-thorn, 01]

4.1   What Are the Main Features Links to the Big Data to Smart Data Closed Loop

- **Automate analysis.** In order to automate day-to-day operational deci-sion-making, organizations need to be able to do more than simply present data on a dashboard or in a report. The challenge is turning real-time data into something actionable. In short, businesses need to be able to automatically interpret data, dynamically, in real time. What this means in practice is the ability to compare each individual event with what would normally be expected based on past or predicted future per-formance. Smart Data must understand what normal looks like at both individual and aggregate levels and be able to compare individual events to this automatically.

- **Forward looking.** Understanding the impact of any given event on an organization needs to be forward looking (predictive BI)

- **Process oriented.** To be embedded within a process in order to make the process inherently smarter requires Smart Data need to be process-oriented. This doesn't mean that the process has been modeled with a business process management tool. Actions can be optimized based on the outcome of a particular process, but the process itself may or may not be explicitly defined.

- **Scalable solutions.** Scalability is naturally a cornerstone because it is based on event-driven architectures. This is critical because event streams can be unpredictable and occur in very high volumes.

- **An intelligent organization.** To give to the enterprise a tactical advantage in the business arena. Concern with transforming your data into an accessible store of high-value information and distributing the right information at the right time to the right person for the right decision. It's easy to see the correlation between the quality of the Business Intelligence and the success of the operation (those who comprehend and act quickly upon relevant facts have advantages over those who do not)

4.2   What Are the Main Risks of Not Implementing a Strong Business Intelligence Practice for the Enterprise?

As the world becomes more digital every day, which, by extension, creating a new stream of data types and associated speed and volume, the company that currently struggling to manage their data will be in a worse situation in the near future

References

Warden, P.: Big Data Glossary, Editions O'Relly

Smolan, R., Erwitt, J.: The Human Face Of Big Data

Business Intelligence for the Enterprise, Mike Biere, Edition IBM presse

Big Data: A Revolution That Will Transform How We Live, Work, and Think; Viktor Mayer-Schonberger and Kenneth Niel Cukier, Hmhbooks (2013)

Learning From Data by Yaser S. Abu-Mostafa, Malik Magdon-Ismail and Hsuan-Tien Lin, AMI Books (2012)

Yield et Revenue Management, Jean-Paul Sinsou, Edition ITA presse

Cross, R.G.: Yield Management. Broadway Books (1997)

Talluri, A.T., Ryzin, G.J.: The Theory and Practice of Revenue Management. International Series in Operations Research & Management Science. Springer (2005)

Hackathorn, R.: The BI Watch: real-Time to Real-Value. DM Review 14(1) (January 2004)

Grundstein, M., Rosenthal-Sabroux, C., Pachulski, A.: Reinforcing Decision Aid by Capitalizing on Company's Knowledge: Future Prospects. In: EJOR 2002 (2002) (to appear)