ERICSSON WHITE PAPER
Uen 288 23-3211 Rev B | October 2015

ERICSSON

BIG DATA ANALYTICS

ACTIONABLE INSIGHTS FOR THE COMMUNICATION SERVICE PROVIDER

Analytics-generated insights are increasingly driving successful decision-making for communication service providers (CSPs). From network enablers to business support systems, there are opportunities to utilize insights gathered from careful analysis of network data.

Big data is at the core of this opportunity. But the traditional technology view of big data is not enough. Rather than focusing on big data technology, CSPs should start from the business value they want to create and apply extensive telecom competence to understand how the relevant insights can be extracted from raw data before applying big data techniques as needed.

INTRODUCTION

People, devices and networks are constantly generating data. When users stream videos, play the latest game with friends, or make in-app purchases, their activity generates data about their needs and preferences, as well as their QoE. Even when users put their devices in their pockets, the network is generating location and other data that keeps services running and ready to use.

As a result, the rate of mobile network data traffic growth is increasing rapidly. It is estimated that by 2020, the number of smartphone subscriptions will have increased from today's 2.7 billion to 6.1 billion, and the total amount of mobile traffic generated by smartphones will be five times that of today [1].

The big-data-driven telecom analytics market alone is expected to have a compound annual growth rate of nearly 50 percent – with annual revenues expected to reach USD 5.4 billion at the end of 2019 [2].

Communication service providers (CSPs) can make use of this big data to drive a wide range of important decisions and activities. These include: designing more competitive offers and packages; recommending the most attractive offers to subscribers during the shopping and ordering process; communicating with subscribers about their usage, spending and purchase options; configuring the network to deliver more reliable services; and monitoring QoE to proactively correct any potential problems. All these activities enable improved user experience, increased customer satisfaction, smarter networks and extended network functionality to facilitate progress into the Networked Society.

The profound impact that increased broadband networking will have on society will also create business opportunities in new areas for CSPs. Improved real-time connectivity and data management enables the creation of tailored data sets, readily available for analysis and machine learning. This enables data-driven efficiency improvements in several business areas – for example, transport, logistics, energy, agriculture and environmental protection. Furthermore, decision-making in business and society will be facilitated by access to insights based on more accurate and up-to-date data.

In the past, CSPs were prevented from benefiting from the value of big data on account of its sheer weight. The volume, velocity and variety – or the three Vs – of big data were simply overwhelming. These data-handling challenges have now largely been met by a variety of easily obtained tools. Distributed databases, complex event-processing frameworks, analytics libraries and so on have been developed in the open-source community and are readily available to CSPs.

But like a blank spreadsheet, these tools are simply a platform for data handling. The real value comes from knowing which combination of the vast array of data elements reveals the desired insights. That is where deep network and operational expertise are of paramount value. Only when these key relationships are understood can the necessary insights – such as user behavior, network performance and causes of experience issues – be gained.

BIG DATA FOR THE CSP

For the CSP, big data not only means a fundamental shift in the way data is stored and managed – it also entails deploying powerful real-time analytics and visualization tools, collaboration platforms and the ability to automatically create links with existing applications vital to the CSP's business. These include operations and business support systems (OSS/BSS) and customer relationship management (CRM).

Insights create the connection between the CSP's data assets and the desired business results. The first step is to connect data collected from a variety of sources: network and non-network, structured and unstructured. Insights may be derived from properly correlated data. For example, it is possible to create extremely rich subscriber profiles that can be useful for customer care, campaign management, customer intelligence and so on. Finally, insights can dynamically change network behavior to ensure an optimal subscriber experience.

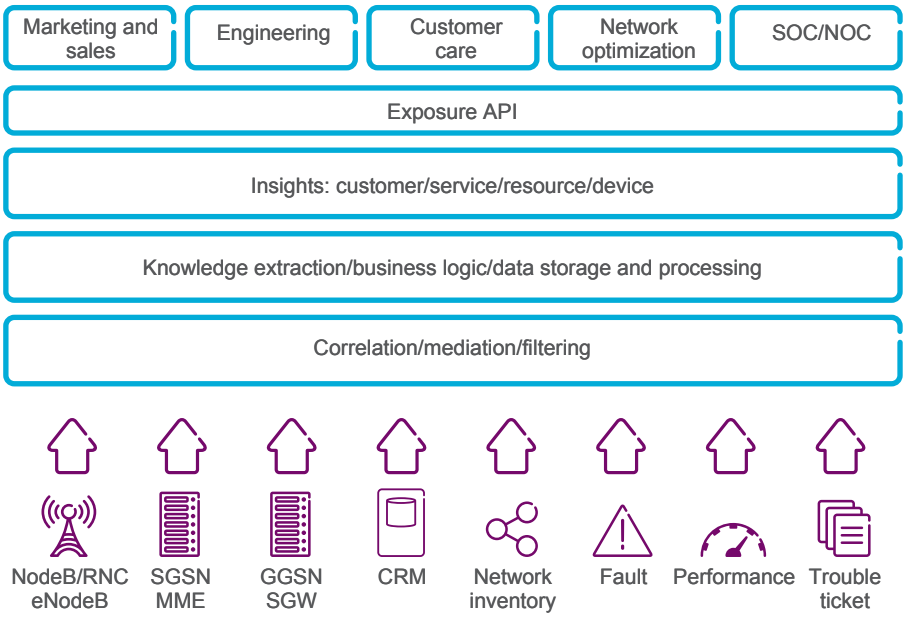

Figure 1: The big data analytics stack.

FROM DATA TO INSIGHT

Successful decision-making will increasingly be driven by analytics-generated insights. And the more accurate and timely these are, the better chance a (human or automated) decision-maker has to anticipate and profit from change.

The keys to effective data-driven decision-making are: the ability to sift through large amounts of data; and the ability to combine data from several sources to gain a more comprehensive view of the business. Seemingly unimportant data can become crucial when combined with other sources to reveal new insights. For example, combining what people write in social media with network data can play a vital role in fully understanding the way users experience a service, or in knowing whether they are likely to churn.

In practical terms, there are several important factors to be considered in managing the expansion of data, devices and network complexity. These are:

- the growth of mobile broadband
- enhanced connectivity and the adoption of machine-to-machine (M2M) technology
- the power of cloud computing
- service expansion and the shift of scale and context with relation to such services
- easy and cheaper access to big data technologies.

The complexity of handling this expanded universe of data sources is compounded by the need to link, match and transform data across heterogeneous entities and systems, while managing scale and ensuring timeliness. Organizations need to understand key data relationships, such as complex hierarchies and links between data types and sources.

A NEW VIEW OF BIG DATA

It is clear that CSPs need to revise their thinking when it comes to big data. The focus should not be on storing and processing large amounts of information, but rather on the ability to make decisions in a timely and accurate way. Analytics – guided by domain expertise – allows CSPs to gain more precise and timely consumer and network insights, which in turn allow for improved operational efficiency and open up new opportunities for revenue generation [3].

For CSPs, a flexible and layered analytics platform allows for the simple creation of new use cases, and such an approach facilitates insights into their existing business processes and tools. For instance, CSPs will be able to prioritize VIP subscribers, provide premium service in a specific region and focus on user experience.

MOVING FROM DATA-HANDLING CAPABILITIES TO VALUE-CREATING SERVICES

Today, big data is understood more or less from a technology perspective: the possibility of better storage (volume), the ability to process information and make it available in real time (velocity) and the ability to deal with various kinds of data sources, including structured, semi-structured and unstructured ones (variety). The technology exists, so the essential issue is how CSPs can make sense of the massive volumes of data and deliver value to subscribers and businesses.

For the CSP performing network-near analytics, extracting value is a domain-specific task, and off-the-shelf IT components do not provide the solution. CSPs are therefore starting to look beyond traditional big data techniques and focus on the analytic value that can be gained from transforming huge volumes of complex and high-velocity data into business insights. The focus is increasingly shifting to three As: adequate, accurate and actionable insights that will enable CSPs to improve their existing processes and drive better decision-making.

CSPs need to shift from analytics application silos to more generic, horizontal analytics environments that take in a wide array of data sources, while supporting a variety of applications and services. Cutting-edge IT components like data storage, data management and network resources are the basis, on top of which CSPs need to apply telecom-specific analytic logic (data models, rules sets and so on) to bring business benefits and maximize capital investments.

To exploit the full potential of OSS/BSS, the CSP needs to enhance these systems with real-time and predictive analytics functionality to support automated decisions, such as continuous updates of network caches based on where a user is expected to move to next, or the offer of dynamic subscriptions based on predicted service usage. The more real-time efficiency an analytics function provides, the more value it can bring to operations. Storing all data in a large database and then running queries on top is a naive approach that will neither scale nor provide the analytics capabilities needed.

OSS/BSS management functionality based on analytics, such as customer experience management (CEM), end-to-end performance monitoring and device management, allows the CSP to more effectively handle the increased complexity of network management, M2M and ever-increasing data volumes.

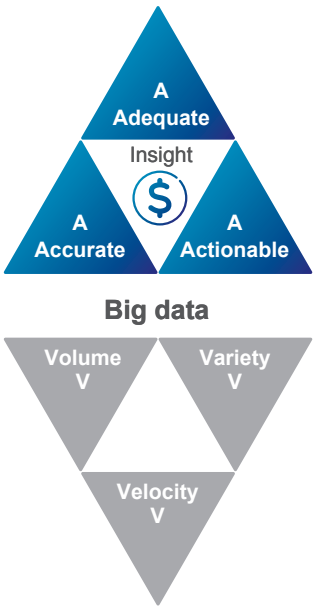

Figure 2: The three Vs and three As of big data.

OPTIMIZE, ENRICH AND CREATE

To derive value from big data, CSPs must know what data to use, how to process, correlate and limit the data, and how

to interpret and apply the resulting insights to different processes (relating to systems, operations, business, consumers and so on). Network data can either be user data transported by the network or data generated by the user device or network element including the core, RAN, transport, terminal, CRM, Home Subscriber Server, charging and billing, service proxy, policy nodes, self-care application, the CSP's web portal and so on.

Three core areas have been identified in which CSPs can maximize the value of their data assets: optimize operations, enrich operations and create new business opportunities.

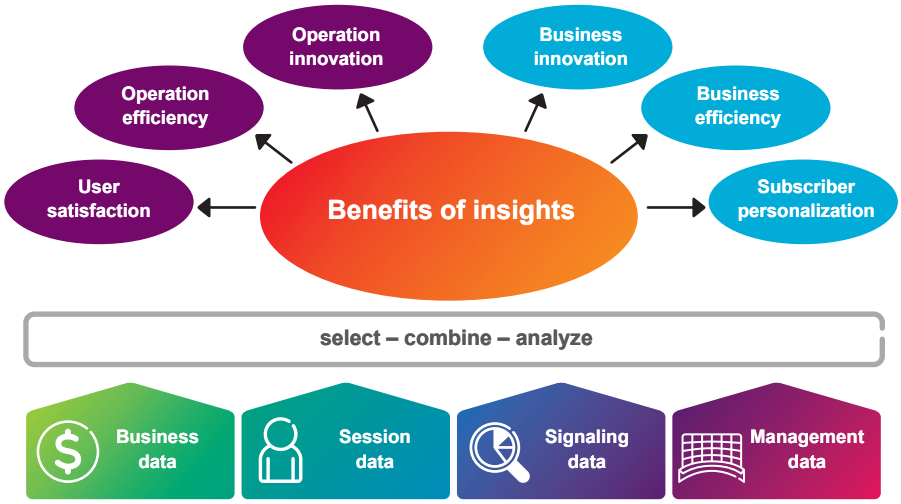

Figure 3: Creating value from network data.

Operations can be optimized using insights gleaned from network data to ensure network resources are used where they generate most value. For example, connecting the analytics function close to the data source decreases the insight delay. The network-near data analytics includes real-time distributed processing to collect, correlate and analyze data from all parts of the network, as well as from probes.

To enrich operations, a CSP can analyze user behavior and identify new product offerings that will meet their needs more effectively. This type of insight will become more important as the Networked Society takes shape and communication involves more new device types.

With new insights, value can be created both for existing and new subscribers. Examples of insights that can be gleaned automatically include mobility patterns that can inform municipal transportation agencies how best to plan their bus routes. In several use cases, the revenue chain is based on a B2B model (where the additional revenues come from a third party, while the subscriber has the benefit of the service).

PRIVACY

Enhanced connectivity and computing power, as well as increased uptake of big data technologies, not only provide opportunities for CSPs but also give rise to new challenges in terms of security and especially privacy. The degree to which big data analytics introduces privacy challenges varies, depending on the kind of data that is processed and on the unforeseen use cases of the data itself.

While security challenges can often be addressed with established techniques, the challenges in terms of privacy often have different aspects that call for new ways of using the established techniques or even new techniques. New challenges include the protection against misuse of personal data generated by linking different data sources, as well as the necessary changes to privacy-enhancing technologies due to the increasing amount and rapidly changing nature of the data. Following existing legislation may not be sufficient in terms of user privacy, as regulations quickly become outdated in the rapidly evolving technology environment. Regulations are also country-specific and may vary in their approach to developments in big data technology.

Privacy challenges can be addressed by firmly integrating privacy awareness into the ways of working and taking privacy considerations into every step of the design process of a new product or service (privacy by design [4]). This includes considering the whole data life cycle management, from collection to processing, storage, and finally erasure. Internal processes can be adjusted and their quality communicated by performing privacy audits. Standardization activities such as the introduction of the ISO 29100 standard [5] can also support internal work.

Technical approaches to address security and privacy include established security techniques such as encryption for secure communication and secure data storage, and refined use of established techniques such as stringent and granular access control. However, they also include specialized techniques such as adjusted data collection, automated data life cycle management (data quality and data retention), audits and logging, and anonymization methods.

THE EXPANDED BIG DATA SYSTEM

Analytics enables CSPs to turn big data into actionable insights. And ultimately, analytics will help the Networked Society evolve by addressing devices, network-connectivity growth, adoption of M2M, growth of the cloud and new business opportunities.

A significant barrier to CSPs applying big data analytics to their solutions is that the data available from network elements, BSS systems, probes and so on is typically elementary. The processing, combination and proper interpretation of this data requires wide and deep protocol and system knowledge. On the other hand, the data processing must be in harmony with the application to business use cases, which requires careful selection of those details that must be transported higher up in the processing chain.

To lift these barriers, telecom-specific analytics platforms are needed that, on the one hand, have built-in domain expertise to translate raw network data to more meaningful but still elementary insights. This expertise can include mobility, QoS, device capability and over-the-top experience. On the other hand, these platforms must provide the means (such as through application programming interfaces (APIs) and software development kits (SDKs)) to assist in the development of CSP analytics use cases.

The careful planning of such a platform also enables CSPs to incorporate additional data sources, even from other industries. The capability to contribute to industry and society-type use cases, combined with future 5G technologies, will definitely be an important part of CSPs' success.

To extract meaningful business value using analytics, the big data stack needs to be expanded with a number of components and technologies. In short, these are:

- the management of data assets, with real-time process control for analytics
- scaling up analytics to gain insights from big data, as well as the ability to solve increasingly complex problems using more data
- an analytics-as-a-service approach and the flexibility to choose between local and hosted deployment options
- securing consumer privacy, while still extracting relevant and timely information from data
- and most importantly, domain-specific analytics platforms and SDKs, and access to experts to ensure that true insights are being extracted.

NEW OPPORTUNITIES

Analytics is essential to the enablement of cost-efficiencies and business innovation across all CSP business segments. Accordingly, solutions will need to provide cross-domain support with a common architecture to maximize reusability, agility and cross-selling of analytics applications based on facilitated use cases.

Insights can be exposed through an integrated and interactive platform that implements or reuses existing APIs and SDKs, allowing multiple applications to be supported by the same underlying data infrastructure.

The underlying big data technology is an enabler, but it does not, in itself, generate value. This is generated by the faster and smarter decisions enabled by insights gleaned from newly discovered and defined data relationships that lead to successful and beneficial use cases. The following sections provide a sample of the kinds of use cases that can be realized using big data analytics.

SMARTER NETWORKS

Smarter networks optimize the use of network resources and management of network traffic to improve the consumer experience. This enables the delivery of more compelling service offerings.

The mix and behavior of services are not static. Every day, new services become available in app stores, and network traffic patterns for individual services change over time. It is therefore important to have solutions that capture relevant characteristics and excel in a changing environment. Big data analytics provides the means of such large-scale statistical analysis for observing and predicting traffic patterns, as well as quickly detecting trend shifts and anomalies.

An example application of big data analytics for smarter networks is the prediction of optimal network parameter settings based on the characteristics of the traffic, to reduce battery consumption on the user device while not increasing network delay.

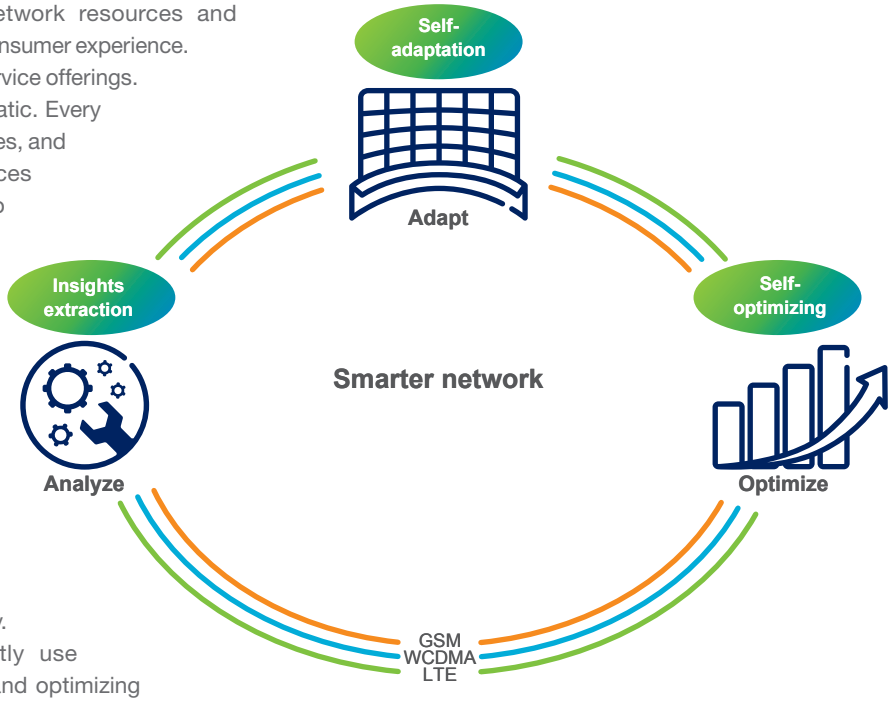

Figure 4: Smart and self-optimizing networks.

Self-organizing networks (SON) efficiently use network-near analytics in planning, building and optimizing network resources. All SON automation – such as provisioning, configuration and commissioning – can be adapted to changes in the environment and traffic demand based on the insights gained from big data analytics.

Finally, next-generation networks will provide ubiquitous mobile communication not only for people but also for connected things. This will add new challenges, such as the ability for the network to handle a massive number of connected devices at low cost and the need for increased energy performance on both the client and network side – challenges where big data analytics will play an important role.

USER EXPERIENCE

Big data analytics reveals what the CSP needs to know to be able to take timely action to resolve issues that impact the user across devices, subscriptions, services and network resources. For

instance, CEM systems monitor and analyze the individual user experience of mobile broadband services and voice [6].

Measuring QoE through well-selected system service key performance indicators (S-KPIs) requires the analysis of distributed network data. Analysis of large data sets also enables precise analysis of the user experience and leads to insights into subscribers' real experience, which could incorporate social data, for example.

A CEM platform collects data down to an individual subscriber's sessions, and can automatically pinpoint the root cause of problems affecting consumer experience, either reactively (within customer care) or proactively (within the Service Operation Center (SOC)/Network Operation Center (NOC)). In addition, communication between customer care and the NOC/SOC is simplified by creating a common interface and language, easily linking QoE to the underlying network. Off-the-shelf data-processing technologies are not enough because root-cause insights are highly network-specific. In-depth network knowledge is required to work out what network data must be analyzed – at the user-session level and in relation to the detailed consumer-experienced service events (user triggers) – to understand both the perceived user experience and the directly associated root cause (in real time).

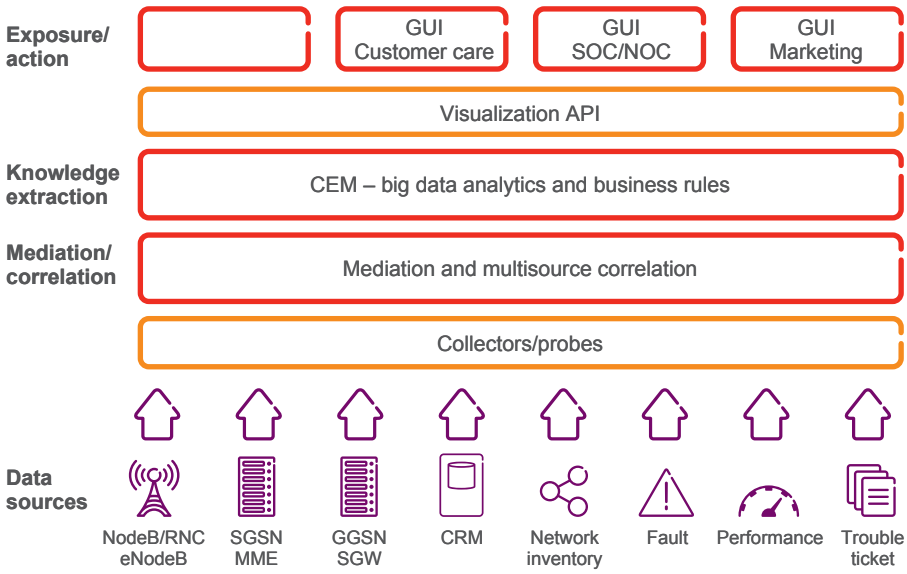

Figure 5: Customer experience assurance.

The user experience reported from the S-KPI measurements is not sufficient to understand the impact on the Net Promoter Score (NPS). The experiences must be evaluated using psychophysical methodologies, and the resulting perception understanding per each user service level index is a step closer to the NPS.

The consumer gets in touch with a CSP on several occasions, defined as "touch points." The touch points in question are Find, Buy (pay for), Receive (get, install), Use, Get Help, Modify and End. "Use" is (hopefully) the touch point to which the user has the longest exposure. It therefore becomes one of the most important in terms of measuring quality.

The whole range of consumer experience journey data must be considered, covering all the user's touch point experiences with the network provider and how each positive or negative event impacts their perception. Applying analytics to such data enables a CSP to predict which subscribers are best for up-selling or re-profiling in order to target their needs better. Compared with most other industries, CSPs are in a unique position; few other businesses can potentially measure virtually every usage of their services in real time per consumer.

DATA BROKERING AND MULTISIDED BUSINESS MODELS

New revenue streams can be captured – monetizing the data that the CSP owns – by leveraging the valuable knowledge extracted from user data. This knowledge is often valuable to a variety of potential business partners. For instance, data about consumer value, customer satisfaction, or consumer segment can be overlaid onto geolocation data, creating a view of concentrations of various types of consumers. These "heatmaps" can show changes over time, driven by time of day, day of week or special events. Analyzing this data set can help with targeted advertising, business investment planning (where to locate) or event planning.

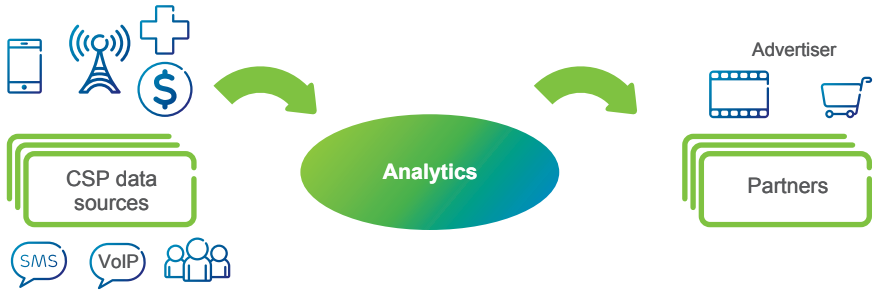

Figure 6: Brokering network data insights.

Typically, these opportunities are pursued through partners, and using aggregated – rather than individual – data becomes important for both practical and privacy reasons. Data monetization is the process of capturing, storing and managing appropriate data, performing analytics to identify key trends and patterns, and sharing the discovered insights with those who can create value from them. Sharing data in a per-subscriber-based manner is not well suited to monetized data services. Aggregating, summarizing, anonymizing, and packaging data in a way that is valuable to partners and third parties – such as advertisers, content publishers and social media – therefore requires infrastructure, systems and interfaces, as well as a clearing house and market for data exchange.

MARKETING

Effective marketing requires accurate insights into experiences and an evaluation of these, but with the introduction of big data analytics it is possible to do far more accurate customer satisfaction predictions, since it enables multifaceted and multisource analysis. CEM systems collect, aggregate and analyze events in the provider's network and other operational systems to provide insights about the consumer experience, including recommended actions.

The evaluated consumer-centric, real-time insights regarding devices, user behavior, contextual data and network data allow marketing departments to identify consumers with poor user experience. They also deepen the analysis with evaluation of the psychological perception of the services, unmet needs, and plans for appropriate targeted actions such as up-selling, cross-selling and retention campaigns.

The subscriber base is analyzed to identify subscriber user groups according to a broad range of parameters, in numerous combinations. It takes user data, such as charging and demographic information, from both network elements and related sources to provide a complete picture of the user in order to identify the most valuable and influential ones, and pinpoint which of them may be unsatisfied network detractors.

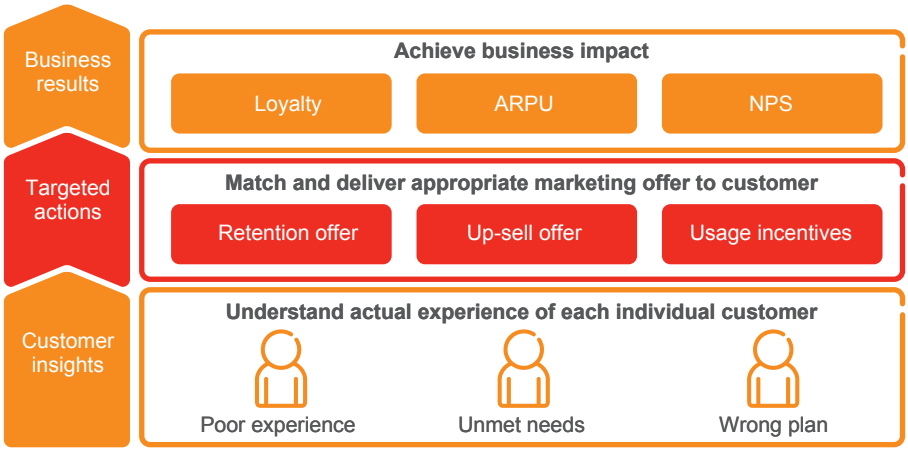

Figure 7: Customer insights, targeted actions and business results.

In addition, by understanding the data in social networks (applying insights into the behavior of influential users – or leaders – and followers), it is possible to identify the expressed feelings about the service experience. Real-time automated analytics will enable closed-loop action, making decisions much quicker and more efficient.

The vast volumes of data in CSP networks across the consumer journey put high requirements on the IT infrastructure processing all this data. The solution to the required computational processing is based on big data technologies and analytics. Big data and analytics are therefore technologies that are used to realize a CEM system. Unlike point CEM tools, end-to-end CEM systems have a horizontal architecture; they collect data from multiple data sources across different network domains for all services and can be organically expanded with the evolution of technology and services.

There are many challenges; a typical example is that big data collections or streams may be optimized for scaling, for example by aggregating data close to the source. This, in turn, may impact the accuracy of the data analysis and the insights available for marketing. Data collection should therefore be optimized to as great a degree as possible to benefit the use cases and insights.

CONCLUSION

True data-driven insight calls for domain expertise. For the CSP, this means in-depth knowledge of how the network functions, what data to pull from the network's nodes and OSS/BSS systems, and an understanding of how to connect data from multiple sources end-to-end to yield an enriched set of information sources. This is what ultimately enables the creation of a range of services and user-centric applications. Smarter networks, customer experience management, data brokering and marketing are just some examples of what is possible.

A common, horizontal big data analytics platform is necessary to support a variety of analytics applications. Such a platform analyzes incoming data in real time, makes correlations (guided by domain expertise), produces insights and exposes those insights to various applications. This approach both enhances the performance of each application and leverages the big data investments across multiple applications. Storing and processing huge amounts of information is no longer the issue. The challenge now is to know what needs to be done within the big data analytics platform to create specific value.

While big data storage and processing techniques are necessary enablers, the goal must be the creation of the right use cases. The big data tools and technologies deployed have to support the process of finding insights that are adequate, accurate and actionable. Instead of talking about the three Vs of big data, CSPs should therefore focus their attention on these three As of big data.

GLOSSARY

API	application programming interface
BSS	business support systems
CEM	customer experience management
CRM	customer relationship management
CSP	communication service provider
GGSN	Gateway GPRS Support Node
GUI	graphical user interface
M2M	machine-to-machine
MME	Mobility Management Entity
NOC	Network Operation Center
NPS	Net Promoter Score
OSS	operations support systems
RNC	radio network controller
SDK	software development kit
SGSN	Serving GPRS Support Node
SGW	service gateway
S-KPI	system service key performance indicator
SOC	Service Operation Center
SON	self-organizing networks

REFERENCES

[1] Ericsson, Ericsson Mobility Report, February 2015, available at:
http://www.ericsson.com/res/docs/2015/ericsson-mobility-report-feb-2015-interim.pdf

[2] Research and Markets, Carrier B2B Data Revenue: Big Data, Analytics, Telecom APIs, and Data as a Service (DaaS) 2015-2020, July 2015, available at:
http://www.researchandmarkets.com/reports/3071341/carrier-b2b-data-revenue-big-data-analytics

[3] Analysys Mason, Big data analytics: how to generate revenue and customer loyalty using real-time network data, January 2013, available at:
http://www.analysysmason.com/Research/Content/Reports/big_data_analytics_Jan2013_RMA05/

[4] A. Cavoukian, Privacy by Design – The 7 Foundational Principles, available at:
https://www.privacybydesign.ca/index.php/about-pbd/7-foundational-principles/

[5] ISO/IEC, 29100:2011 Information technology – Security techniques – Privacy framework, 2011, available at: http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=45123

[6] Ericsson, White Paper: Translating user experience into KPIs, August 2015, available at:
http://www.ericsson.com/res/docs/whitepapers/wp-kpi-user-experience.pdf