



Augmented analytics

An approach to develop
better insight from data

theiet.org/augmented-analytics

Contents

1. About this fact file	3
2. Understanding business intelligence, management information and analytics	4
2.1. Business intelligence (BI)	4
2.2. Management information (MI)	4
2.3. Analytics	5
3. Augmented analytics	6
3.1. The change from purely specialist expertise	6
3.2. Data wrangling	7
3.3. Citizen data scientists	7
4. Augmented analytics – a crossover of technologies and trends	8
5. About the IET	11

Augmented analytics is published by the Institution of Engineering and Technology.

Please note that the views expressed in this publication are not necessarily those of the IET. It is not intended to be a guidance note with a specified set of recommendations or actions but rather seeks to add understanding and debate around the topic.



The Institution of Engineering and Technology (IET) is working to engineer a better world. We inspire, inform and influence the global engineering community, supporting technology innovation to meet the needs of society. The Institution of Engineering and Technology is registered as a Charity in England and Wales (no. 211014) and Scotland (no. SC038698).

© The Institution of Engineering and Technology 2020.

1. About this fact file

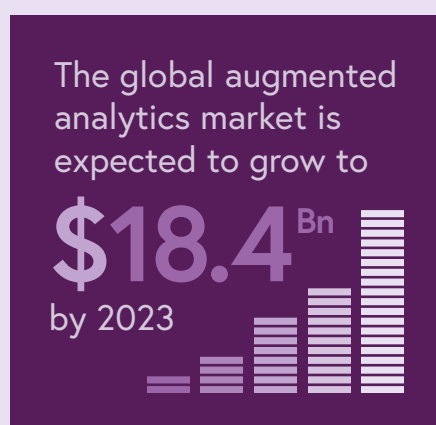
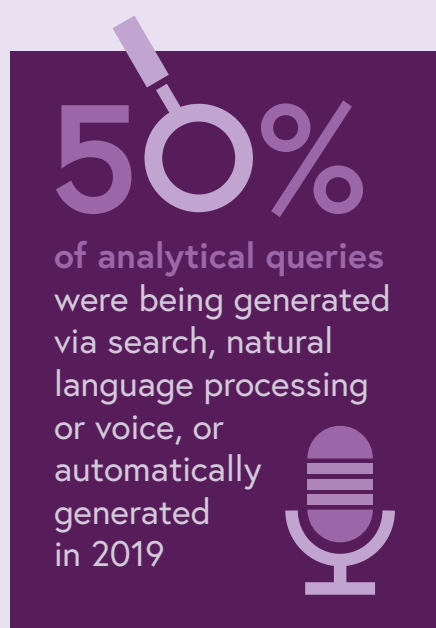
Augmented analytics is an analysis approach that uses existing and emerging software technologies such as advanced statistics, machine learning (ML) and natural language processing (NLP) to develop sophisticated insights from data.

Business intelligence (BI), management information (MI) and analytics are three separate though interlocking domains within information technology. BI is a way to reveal actionable insights in your data, find out what happened, and then explore why it happened. MI allows you analyse trends, forecast the future and resolve problems you identify. Analytics enables you to assess your organisation using advanced statistical and modelling techniques. These insights can be used to chart or change the path of your business.

This fact file reviews the terminology around the business intelligence, management information and analytics domains and sets a context for augmented analytics. It introduces the knowledge domain, defines augmented analytics, and discusses the challenges and benefits that it can bring. The intended audience of this fact file includes technology leaders and decision makers, practitioners, business intelligence and analytics communities, and data architects and specialists.

In 2019, approximately 50% of analytical queries were being generated via search, natural language processing or voice, or automatically generated¹. According to industry analyst Gartner, augmented analytics would by 2020 already be a dominant driver of new purchases of analytics and business intelligence, data science and ML platforms, and embedded analytics. Gartner has refined this view as the "democratisation of technology... providing people with easy access to technical or business expertise without extensive or expensive training"². Another analyst firm, Research and Markets, has predicted³ that the global augmented analytics market is expected to grow to \$18.4Bn by 2023.

The author of this fact file is Alastair McCullough BA (Hons) MSc FIET FIOD FRSA FBSC CITP MCMI ChMC. He has helped organisations with leading technology strategies and solutions for nearly thirty years. Since founding the UK office of Seattle-based metadata software company Noetix Corporation in 1997, he has specialised in data, analytics, business intelligence and management. Today he is the Data and Analytics Competence Strategy Leader, for Europe, Middle East and Africa, and an Enterprise Information Architect with IBM Services in London, UK.



¹ *Gartner Identifies Top 10 Data and Analytics Technology Trends for 2019* –

<https://www.gartner.com/en/information-technology/glossary/advanced-analytics>.

² *Gartner: Top 10 strategic technology trends in 2020*; Burke, Brian; *Computer Weekly*; 2 January 2020:

<https://www.computerweekly.com/opinion/Gartner-Top-10-strategic-technology-trends-in-2020>.

³ *Augmented Analytics Market by Software, Service (Training and Consulting, Deployment and Integration, and Support and Maintenance), Organisation Size (SMEs and Large Enterprises), Deployment Type, Vertical, and Region – Global Forecast to 2023*; February 2019 - https://www.researchandmarkets.com/research/lshr95/18_4_bn?w=5.

2. Understanding business intelligence, management information and analytics

Business intelligence (BI), management information (MI) and analytics are three separate though interlocking domains within information technology. It's worth noting, however, that they differ fundamentally from 'pure' information technology. Information technology⁴ is the enabler for the delivery of business intelligence and analytics solutions that offer critical *insights* to the organisation, and from which new insights may be derived. Likewise, management information is not the same as business intelligence, and analytics differs again.

2.1 Business intelligence (BI)



The term BI describes and characterises a swathe of software, application, business and technical processes designed for turning critical business knowledge into actionable information. BI is often wrapped into

the concept of business performance management, which can be framed as 'the analytics of business'. Components include metrics, measures, dashboards, scorecards, planning, budgeting, forecasting and visualisation technologies such as graphs, charts, representations and advanced graphics.

Examples of modern BI software solutions include QlikView®, QlikSense®, MicroStrategy®, SAP®'s Business Objects™, IBM®'s Cognos™ Analytics and Cognos™ Planning Analytics, and Microsoft®'s Power BI™. Reports are developed by business end users – rather than programmers or developers – to enable analysis, dashboards and scorecards that monitor business performance, analyse trends and measure results. Typically, these tools support classic intelligence functions such as data slicing, dicing, drilling⁵, collapsing⁶, pivoting, representation and visualisation of data. However, they also include trend and seasonal analysis, scorecards and dashboards.

BI software solutions combine rich statistical software tooling with the ability to test and answer different questions and hypotheses and to project futures based upon simple statistics (average, mean, count, sum etc.). They also feature data manipulations including set theory, tests, rules and logic such as AND, OR, NAND, intersect, union and minus functions.

The key feature of BI is its ability to support close interaction with users. Such users are free to revise their models and ideas based on insights, adding graphics and charts, embedding logic, mathematical and statistical operators and – sometimes sophisticated – rules for assessing data or producing programmed insights.

2.2 Management information (MI)



MI typically involves the delivery of **predominantly fixed format reports** that record activities. These commonly include analyses of volumes (how much, how many, how few, how often), summaries (such as where and what type) and pure statements of recorded facts. Statements of facts are usually taken from a non-analytical data store such as those relating to financial

⁴ Information technology in the form of technical infrastructure, support services and data sources such as 'big data' solutions like Hadoop, open-source stream processing tools such as Kafka, data warehouses, databases and files.

⁵ Drilling is where an end user performs increasing refinements to report details, analysing to ever more detailed knowledge scope or grain, moving down through data. For example, from country statistics to county to town to area to street to house.

⁶ Collapsing is where an end user performs increasingly high-level report refinements, analysing broader scopes of knowledge, moving up through data. For example, from company salesperson statistics to sales team to town to region to country to international region to overall company.

systems including general ledger, accounts payable, purchasing, payroll and human resources. MI reports include business-related features such as totals, summaries and simple statistics. Typically, the reports come in the form of spreadsheets or fixed format printed or electronic outputs, such as a summary of monthly payslip totals in a PDF file.

MI tends to provide information more than generate insights. Where insights do arise, the end user must provide more intelligence and additional analysis themselves, building on the traditional MI.

2.3 Analytics



Traditionally, the analytics domain has focused on 'business assessment', namely data gathering, questioning, modelling, predictions/predictive analytics, data mining, simulation, visualisation and optimisation. In recent years, analytics has extended to include the 'cognitive' knowledge domain, which features ML and aspects of artificial intelligence (AI).

Gartner defines advanced analytics as:

"The autonomous or semi-autonomous examination of data or content using sophisticated techniques and tools, typically beyond those of traditional BI, to discover deeper insights, make predictions, or generate recommendations⁷."

The analyst firm notes that advanced analytic techniques include data/text mining, ML, pattern matching, forecasting, visualisation, semantic analysis, sentiment analysis, network and cluster analysis, multivariate statistics, graph analysis, simulation, complex event processing and neural networks.

This definition highlights a crossover between analytics and BI. If the two knowledge domains were represented as a Venn diagram, there would be a noticeable

intersection around areas such as visualisation and the ability to offer accessibility in the form of data slicing, dicing and drilling. However, a significant difference between the two is the way that data can be prepared, transformed, loaded, cached and acted upon as part of deep statistical analysis.

BI tools tend to rely more on metadata⁸ layering via materialised views⁹, data structures and varietal metadata in logical 'layers', lower than that of presentation, and before analysis. BI tools present data more for manipulation and representation. In contrast, analytics tooling tends to focus more on the capabilities around the mining, decisioning and analytical deployment of data from varied sources, before deep statistical and probabilistic analysis.

Currently, there's confusion in the marketplace and amongst technical experts due to the misunderstanding that ML and AI algorithms are actually elements *within the advanced analytics domain*. This is because of the statistical and probabilistic nature of their features, their historical development and the nature of their commercial and academic uses in the area of analytics. However, from this perspective, these are not so much 'cognitive' as purely advanced forms of statistical tooling.

In technical terms, we might differentiate analytics from BI and MI by understanding its software focus. Analytics focuses on tools with a significant deep data analysis capability. These include **SPSS**[®], **Oracle™ Analytics**, **IBM Planning Analytics** (formerly IBM Cognos TM1[®]), advanced **Microsoft™ Excel™**, **Tableau™**, **Tibco® Spotfire**, **SAS®** and programming languages or packages such as open source **R**, **Shiny®**, **Python™** and **MATLAB®**, which also feature advanced analytics capabilities.

IBM's famous **Watson™** offering is often presented to aid understanding as if it were almost conscious and thinking as a human would¹⁰. Technologically it's a portfolio toolset, that includes ML capabilities that typify probabilistic statistics. In this respect, it fits directly into the advanced analytics domain and is synonymous with **data science**.

⁷ *IT Glossary: Advanced Analytics*; gartner.com – <https://www.gartner.com/en/information-technology/glossary/advanced-analytics>.

⁸ Metadata is a set of data that describes and gives information about other data. For example, in BI metadata layers, a BI developer might create structures for reporting that bear business names that are easily understood by a non-technical user. This is done rather than using the structure names in a Hadoop file system, data warehouse, database or mart, that are technically focused and lack everyday meaning for the user. Technically focused terms would prevent users from understanding how to report easily.

⁹ A database solution originally developed by Oracle Corporation, materialised views are database structures that typically store (or represent the storage of) rows of data, segmented into columns or fields. This is as a result of a piece of code that executes at a time or in response to a rule being initiated. Materialised views act conceptually as partial windows on one or more large bodies of data, held in a database, to enable and support more effective, simpler or more relevant analysis. They are largely synonymous with **Data Marts**, that is, subject oriented data stores typically for online analytical processing (OLAP) and used with BI, MI and analytics reporting software tools.

¹⁰ *Teaching IBM's Watson How to think like a Human: High, Rob, IBM Watson Solutions* – <https://www.forbes.com/sites/ibm/2013/11/04/teaching-ibms-watson-how-to-think-like-a-human>.

3. Augmented analytics



Augmented analytics builds on the concepts and solutions developed as part of BI and MI and upon their historical legacy since the late 1990s. It's software technology that automates the selection and processing of data and features embedded ML and AI within the platform, product or suite of tools.

Augmented analytics solutions come with prebuilt models and algorithms, meaning that they don't need to be developed by specialist data scientists. Model construction and numerical analyses in augmented analytics solutions may adopt the 'always on' approach, working in the background to support continuous learning and to assist users by developing more accurate insights and decisions. They may also include a workflow mechanism that automatically adjusts and refines ML algorithms that contribute to the analysis outputs without the traditional intervention of a specialist that BI, older advanced analytics and MI required.

3.1 The change from purely specialist expertise



A key feature in turning data into actionable information via BI, MI and analytics lies in having the right expertise in the workplace. Today, people who offer experience and skills in BI, MI and analytics are typically technical consultants and developers specialising in areas such as **R**, **Python**, **RStudio**'s **Shiny**[™] and various types of metadata and metadata structures. They may also be data scientists; statisticians with expertise in advanced mathematical, probabilistic and statistical analysis skills.

These specialists are rare and expensive to hire as they're highly skilled individuals who may have advanced qualifications in statistics and often possess significant industry expertise and insights as well. Anecdotal evidence highlights staff retention issues, as individuals with such highly sought-after skills have plenty of employers vying for their services.

3.2 Data wrangling



The issue of data wrangling – sourcing, extracting, transforming, loading, preparing and using data – is important. There are several stages to go through before data is ready for statistical analysis. These include:

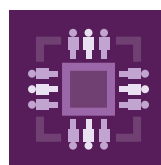
- extraction from a source system or data store
- transformation to ensure the data is usable and in an appropriate form
- loading into a target data store, cache, data sandpit, or other architectural component
- presentation via one or more layers (typically metadata) to the service or software solution that performs analytical functions upon it.

In concept, augmented analytics reduces the dependence on human skills in this area and simplifies the complexities of data wrangling. Augmented analytics is a data analysis approach that combines existing and emerging software technologies to provide, support and enable sophisticated insights into data. The augmented analytics domain features abilities that will identify hidden patterns in data whilst removing human bias from that data. This technology holds out the prospect of embedding **automated** insights into enterprise software applications in an actionable manner.

In theory, an augmented analytics engine can sift, cleanse, analyse and present the data ready for use by people or software mechanisms, transforming data **into actionable information** with minimal specialist human supervision.

A key advantage here is the potential to provide analytical insights to smaller corporate entities at lower costs. This means that capital outlay can be reduced, which is critical at a time of reduced margins and squeezed budgets.

3.3 Citizen data scientists



Gartner¹ identifies another aspect of this new knowledge domain, which it refers to as 'citizen data scientists'¹¹. The analyst firm says that throughout 2020, the number of citizen data scientists will grow five times faster than professional data scientists. Citizen data scientists use AI-powered augmented analytics tools that automate the data science function. The tools automatically identify data sets, develop hypotheses and identify data patterns. Businesses are looking to citizen data scientists to enable and scale data science capabilities. Gartner predicted in 2019 that by 2020, more than 40% of data science tasks would have been automated, resulting in increased productivity and broader use by citizen data scientists.

Throughout 2020, the number of citizen data scientists will grow

5X faster



than professional data scientists



By 2020, more than

40%

data science tasks would have been automated



¹¹ Gartner says that a citizen data scientist is someone "who creates or generates models that use advanced diagnostic analytics or predictive and prescriptive capabilities, but whose primary job function is outside the field of statistics and analytics". In other words, these are primarily enterprise corporate citizens. Source: *Gartner says more than 40 percent of data science tasks will be automated by 2020*; 16 January 2017.

4. Augmented analytics – a crossover of technologies and trends

Augmented analytics solutions are a development stemming from BI, MI and analytics. As a result, as a class of software solution, they typically include capabilities such as the following, which evidence both this historical background and the crossover of a number of different trends in the technology industry today.





ML – algorithms and tests related typically to data patterns

Natural language generation (NLG) – human language words and phrases, either generated by the software or instructions given to it by a user or operator

Automated generation of insightful information, narratives and actions

Cloud-based software provision such as software as a service (SaaS) or platform as a service (PaaS)

Software tooling that supports the design of visualisation dashboards

Broader graphical data visualisation facilities

Real-time data analysis

Provision of predictive and prescriptive analytics

Enrichment of data sources – merging third-party authoritative data with data in an existing operational data store

Metrics recommendations – key performance indicators, metrics and measures

Data flows and sources recommendations

Identification of drivers and performance influencers (a development of BI's performance management concepts)

What makes augmented analytics platforms unique is their ability to process data far faster than a human can and, importantly, without exhibiting the same amount of unintentional or intentional bias. With more traditional types of analytics and business intelligence approaches, data analysts test out their own theories and hypotheses, operating on the premise of knowledge. In the growth of augmented analytics, ML learns through developing models at a faster rate than a human can, and with minimal human interference. This, in theory, reduces the degree of human bias, inherent in such models in previous generations of software¹².

Although developing significantly, it's important to note that augmented analytics is still at an early stage compared to disciplines such as BI and MI.

Vendors offering augmented analytics solutions today¹³ include **IBM, SAP, Salesforce®, Qlik®, Microsoft, Oracle, SAS, TIBCO, dotData™, Aible™, AnswerRocket™, Big Squid™, DataRobot™, Outlier™, Prevedere™, Tellius™, Evolved Analytics' DataStories™, H2O.ai, and Sisense™.**

A development of augmented analytics is augmented intelligence, which supports a broader range of interactions with technology in smarter ways. It's a development of artificial intelligence emphasising the capability to assist humans rather than replace them. Gartner calls this "a design pattern for a human-centred partnership model of people and [AI] working together".¹⁴ Solutions such as Apple's Siri® and Amazon's Alexa™, are examples of augmented intelligence featuring both AI and NLG. They are of course widely used internationally in the general population and are developing increasingly useful capabilities and thought processes. The uptake by banking of a variety of bots evidences a corporate enthusiasm for 'business to consumer' applications¹⁵ in this domain.



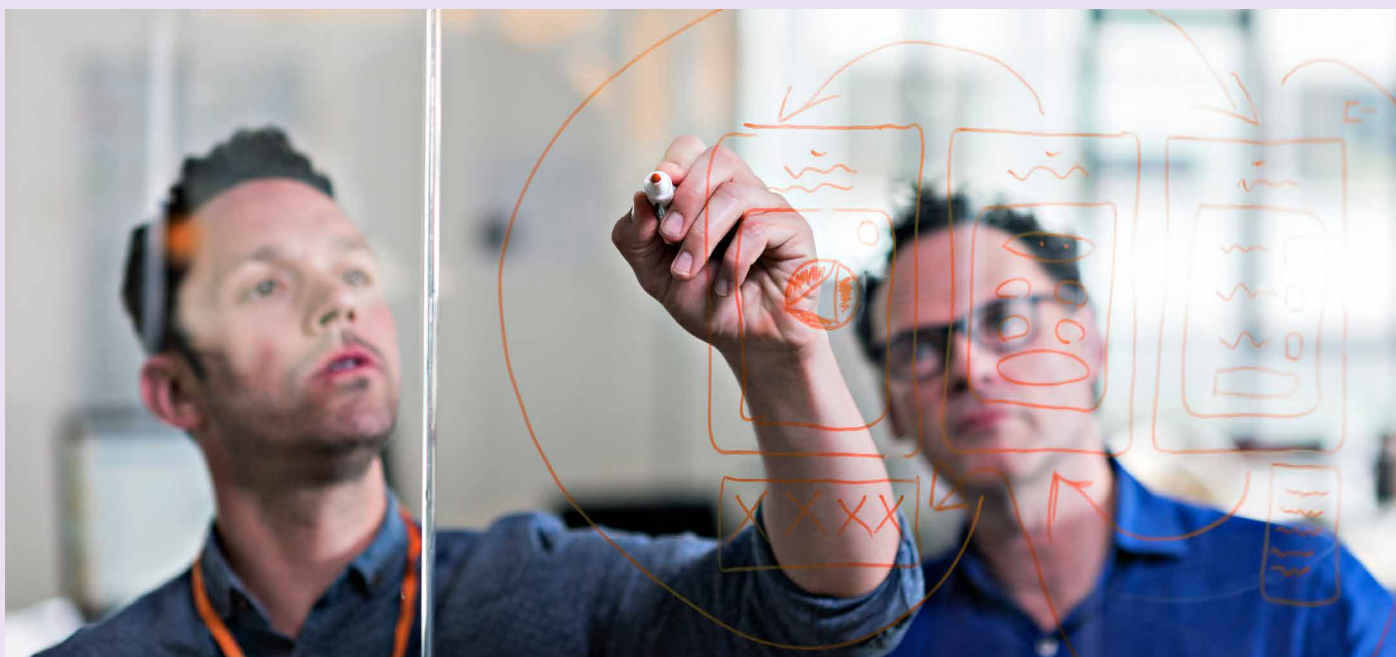
¹² *Augmented Analytics 101: A Starter's Guide*; 27 November 2019; Reilly, Pete; SelectHub - <https://www.selecthub.com/business-analytics/augmented-analytics-guide>.

¹³ *Market Guide for Augmented Analytics*; 30 October, 2019; Krantz, Austin; Gartner - <https://www.gartner.com/en/documents/3970874/market-guide-for-augmented-analytics-tools>.

¹⁴ *Gartner Glossary: Augmented Intelligence*; 29 March 2020; Gartner - <https://www.gartner.com/en/information-technology/glossary/augmented-intelligence>.

¹⁵ *Look For Banking Applications With Embedded Artificial Intelligence: Off-The-Shelf Capabilities Offer A Fast Path Toward AI*, Forrester.com <https://www.forrester.com/report/Look+For+Banking+Applications+With+Embedded+Artificial+Intelligence/-/E-RES141762#>.

5. About the IET



We are the IET - a charitable engineering institution with over **169,000 members in 150 countries** – working to engineer a better world.

Our mission is to inspire, inform and influence the global engineering community to advance technology and innovation for the benefit of society.

As a diverse home across engineering and technology, we share knowledge that helps make better sense of the world in order to solve the challenges that matter. It's why we are uniquely placed to champion engineering.

We bring together engineers, technicians and practitioners from industry and business, from academia and research, and from government and the third sector. We are member-led, independent and impartial.

Passionate about the digital sector, we bring together expert practitioners from the digital industry, academia, the public and third sector.

We champion engineers and technicians working in the sector by offering networking, volunteering and thought leadership opportunities.

Together, we campaign on issues of the day and provide policy input to Government. Your specialist digital knowledge can inspire others and make a difference.

To find out more contact sep@theiet.org

Our offices

London, UK

T +44 (0)20 7344 8460

E faradaycentre@ietvenues.co.uk

Stevenage, UK

T +44 (0)1438 313311

E postmaster@theiet.org

Beijing, China

T +86 10 6566 4687

E china@theiet.org

W theiet.org.cn

Hong Kong

T +852 2521 2140

E adminap@theiet.org

Bangalore, India

T +91 80 4089 2222

E india@theiet.in

W theiet.in

New Jersey, USA

T +1 (732) 321 5575

E ietusa@theiet.org

@TheIET      

theiet.org

The Institution of Engineering and Technology (IET) is working to engineer a better world. We inspire, inform and influence the global engineering community, supporting technology innovation to meet the needs of society. The Institution of Engineering and Technology is registered as a Charity in England and Wales (No. 211014) and Scotland (No. SC038698). Michael Faraday House, Six Hills Way, Stevenage, Hertfordshire, SG1 2AY, United Kingdom.