

InDetail Paper by Bloor
Authors **Philip Howard**
Publish date **January 2019**

IBM Cloud Private for Data 1.2



“

We would especially like to emphasise that one of the difficulties we have seen with companies trying to implement AI is a disconnect between data scientists who develop the relevant models, and those responsible for deploying those models in production.

”

Executive summary

Artificial Intelligence (AI) and machine learning (a subset of AI) are the topics du jour within not just the IT and analytics communities but for business in general. They are constantly being referenced across the media and within boardrooms. There is general agreement that there is huge potential value in implementing the techniques and technologies associated with AI and machine learning, both directly in terms of increased efficiency and as a competitive differentiator. Not to mention the ability to introduce new business models and services. Nevertheless, while it may be easy to see the potential benefits, implementing them is not so simple. Forecasts suggest – see [Figure 1](#) – that very large numbers of companies will be investing in AI over the years to come but, today, only a relative handful have been able to do so. There are many reasons for this: cultural change issues, security concerns, talent acquisition and others, which we are not in a position to discuss. However, there is one issue – limited or no technological capability with respect to data and analytics – that we will consider in this paper and, specifically, we will discuss IBM Cloud Private for Data. This was initially released in May 2018, but this paper reflects the most recent release of the product as of December 2018. The product is currently available in two versions: a Cloud Native Edition (with basic support) and an Enterprise Edition (premium support). There is also IBM Cloud Private Experiences, which allows you to test drive the environment without installation or cost.

IBM Cloud Private for Data is an integrated data science, data engineering and app building platform built on top of IBM Cloud Private (ICP). The latter is intended to **a)** provide all the benefits of cloud computing but inside your firewall and **b)** provide a stepping-stone, should you want one, to broader (public) cloud deployments. Further, ICP has a micro-services architecture, which has additional benefits, which we will discuss. Going beyond this, ICP for Data itself is intended to provide an environment that will make it easier to implement data-driven processes and operations and, more particularly, to support both the development of AI and machine learning capabilities, and their deployment. This last point is important because there can easily be a disconnect

between data scientists (who often work for business departments) and the people (usually IT) who need to operationalise the work of those data scientists.

During the course of this paper we will first consider ICP and then ICP for Data. However, it is worth discussing, briefly, some general reasons why you might want to deploy either of these. In the case of ICP there are a couple of reasons that go beyond the flexibility and scale provided by cloud computing in general. The first is the ability to modernise enterprise applications by refactoring these, using the microservices offered by ICP. And, secondly, there is the ability to build cloud-native applications which may either be new or which leverage existing applications and data, and perhaps including public cloud services, while keeping your data securely behind your firewall. Many of the use cases for ICP for Data are really an extension of these capabilities, allowing operational applications to have decisioning or machine learning built into them: trained within ICP for Data but deployed either directly or via a cloud-native application, as you prefer. And there are other use cases: for example, you could easily deploy a data lake (or lakes) using ICP for Data with a data catalogue and data virtualisation (see below) spanning all of these, so that they do not become siloed.

“
There is huge potential value in implementing the techniques and technologies associated with AI and machine learning, both directly in terms of increased efficiency and as a competitive differentiator.
”

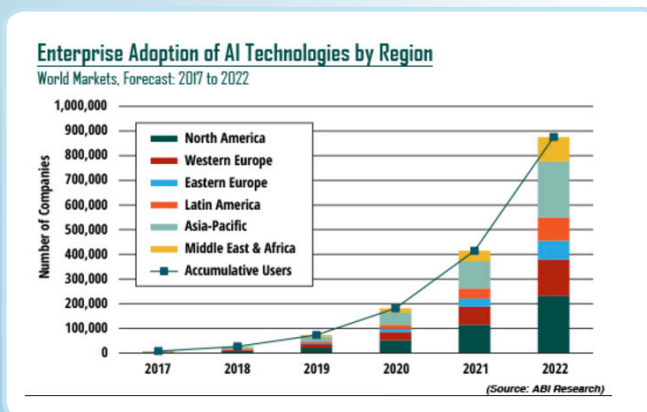


Figure 1:
Forecast
adoption of AI

However, it is not our intention to write reams about the features of these products nor to go into details about use cases. Rather, we intend to reference the business benefits these offerings are intended to address, as well as to discuss the issues that can arise within this environment. We conclude with a section on IBM's short-term roadmap for new releases of ICP and ICP for Data.

IBM Cloud Private



This significantly speeds up release cycles (continuous delivery) and enables incremental improvements to applications.



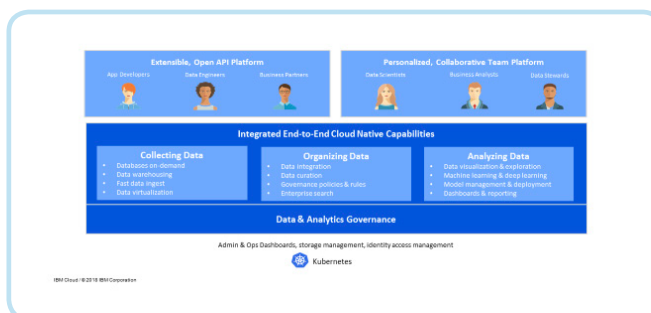
ICP is a private cloud. Note that this is distinct from a virtual private cloud where a public cloud vendor acts as the service provider, with ICP either your IT department acts as the service provider or you can get a third-party to manage it for you, but in either case business units act as tenants. The basic principle is that you can get the benefits of a cloud deployment but with your data remaining protected behind your firewall, though it can be hosted externally, if you prefer. ICP interoperates with traditional on-premises deployments as well as both public and virtual public clouds, so it supports hybrid environments and is seen by IBM as a potential stepping stone to public clouds in the future. It is implemented on Linux-based x86 platforms and runs on IBM SoftLayer as well as the Red Hat (soon to be a separate business entity within IBM) OpenShift Container Platform, for which both ICP and ICP for Data are Red Hat Certified

discussing the benefits that this brings, as the use of microservices, based on Docker and Kubernetes, is a relatively new concept.

According to www.microservices.io “microservices – also known as the microservice architecture – is an architectural style that structures an application as a collection of loosely coupled services, which implement business capabilities. The microservice architecture enables the continuous delivery/deployment of large, complex applications.” Microsoft goes further when it says (amongst other things) that “in some ways, microservices are the natural evolution of service-oriented architectures (SOA) but there are differences. Some of the defining characteristics of a microservice include the fact that services are small, independent and loosely coupled; that each service is a separate codebase, which can be managed by a small development team; that services can be deployed independently and that an existing service can be updated without rebuilding and redeploying the entire application; and that services communicate with each other using well-defined APIs with implementation details hidden from other services.”

What IBM has done with ICP and ICP for Data is to take its existing capabilities and recreate them – where that makes sense – as services, as shown in [Figure 2](#). Some of the advantages of this approach are described in the quotes above, but there are others. For example, we would should note that this encourages interoperability. It makes it much easier to take a database service and combine it with a machine learning service, for example, because you can pick and choose what services you want to use. The other point that we should emphasise, is that this significantly speeds up release cycles (continuous delivery) and enables incremental improvements to applications. This is important, not just because you get new features faster, but also because traditional release cycles of a year or eighteen months, are usually disruptive and result in delays to new releases being implemented. Moreover, bearing in mind that this is a cloud-based environment, this means that IBM can also incrementally introduce new capabilities in the same fashion.

Figure 2:
The components of
IBM Cloud Private
for Data



Containers. The products are also available on AWS and IBM has announced that ICP and ICP for Data will also be ported to the Azure platform. The company also plans to make ICP for Data available as an appliance and to make the products available on the Power and Z mainframe platforms. ICP for Data leverages all the underlying capabilities of ICP itself, for administration, security, logging, monitoring and so forth.

The virtues of cloud-based computing, particularly the advantages to be derived from elastic scaling, rapid deployment, scalability and so forth, are well-known and we do not intend to rehearse those issues here. However, ICP (and ICP for Data) has a microservices architecture, and it is worth

IBM Cloud Private for Data

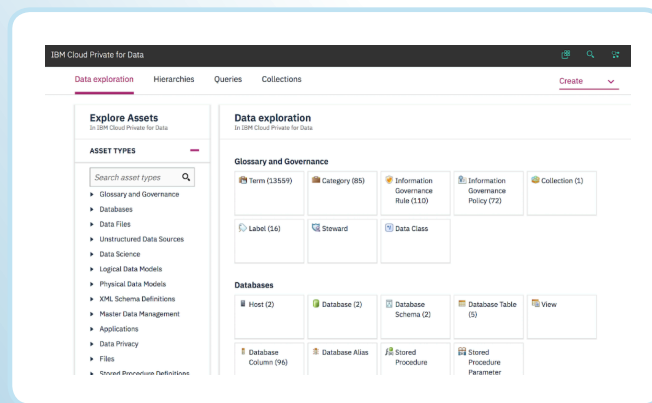
Supporting AI initiatives requires that you obtain all the data you need, govern it to ensure it is trustworthy, analyse and build the machine learning and other algorithms necessary for the project in hand and, finally, to be able to put the results of this exercise into production. This is not a trivial task. The individuals and groups – shown in [Figure 2](#), which illustrates the components of ICP for Data – responsible for these activities are often disparate and disconnected and it requires a collaborative approach to make this work efficiently. Moreover, it requires a set of capabilities that are beyond most software providers and, to enable the sort of collaboration required, you would really prefer a consistent user interface across all the underlying software capabilities that are necessary. While we will discuss ICP for Data in more detail shortly, one of the most notable features of this product is precisely that it provides a common user experience across its software stack. And it is worth commenting that it is the micro services architecture that underlies ICP for Data that has enabled this, by decoupling the user interface from the individual software components, and then creating a new consistent interface, a number of examples of which are included in this paper.

There are a number of elements of the ICP for Data architecture that require discussion. Fundamental is the Enterprise Data Catalog, which forms a part of the Data and Analytics Governance layer. Essentially, this is like a library catalogue in that it holds details of all the data assets that are available to the organisation. However, unlike a library which simply gives you a reference for each named book, the data catalogue allows you to search by category so that you can, for example, find all assets relating to sales, customers or products. In other words, it allows you to find assets of value to your role that you might not otherwise be aware of. This is illustrated in [Figure 3](#).

Where we take issue with IBM on [Figure 2](#) is the division between the top two boxes. We understand that this is a marketecture diagram, but this does not accurately reflect IBM's intentions or capabilities. It suggests that App

Developers and Data Engineers do not share the “*personalised, collaborative team platform*” – which includes crowd sourced recommendations as well as more prosaic capabilities such as workflow - available to other users. Our understanding is that this is inaccurate. As it should be.

Above the Data and Analytics Governance layer in [Figure 2](#), are three boxes. IBM refers to these as “*collect, organize and analyze*”. If it were down to us we would probably put data integration into the “*collect*” box and rename “*organize*” as “*governance*” but that is merely a semantic quibble. We will discuss each of these separately.



“There are implications with respect to AI and machine learning that may not be immediately obvious.”

Figure 3: Exploring data assets via the Enterprise Data Catalog

Collecting data

The “*collect*” box in principle supports the ability to leverage all sorts of data sources, which is obviously necessary: if you don’t have the data you can’t analyse it. There are two principal aspects to this: what is listed as “*databases on demand*” and “*data virtualization*”, which need separate discussions. In the latest release the former covers the entirety of the Db2 family (Db2 itself, Db2 Warehouse, Db2 Event Store and so on) as well as Hadoop, Big SQL (IBM’s SQL on Hadoop engine) and MongoDB (a new addition in this release). Attention should be paid to the fact that Db2 Event Store is supported, because this will be especially important in Internet of Things (IoT) based environments where you need to store data originating in sensors, actuators and other edge devices.

When ICP for Data was initially released the “*collect*” box included data federation but this has been replaced with

data virtualisation. This represents a significant enhancement to the product and it is worth describing how it works because IBM's capabilities in this area are both innovative and ahead of other offerings in the market. To understand how this technology (which was previously codenamed QueryPlex) works, it is worth comparing how it differs from conventional approaches to data federation, shown in [Figure 4](#), with IBM's use of a computational mesh, illustrated in [Figure 5](#).

Netezza, BigSQL, Informix, Derby, Oracle, SQL Server, MySQL, PostgreSQL, Hive and Impala. Further support is in the pipeline and is discussed in the Roadmap section below. Note that Data Virtualization integrates with the Enterprise Data Catalog.

While [Figure 5](#) gives a conceptual view of how IBM Data Virtualization works, what the user sees is illustrated in [Figure 6](#). In other words, as is described in [Figure 5](#), the edge devices organise themselves and you don't need to know how that is done. Note that although this refers to "edge devices", which would be entirely appropriate for IoT environments, it also applies to databases and other sources that need not necessarily be on the edge.

Organising and analysing data

As far as "organize" is concerned, ICP for Data leverages familiar technologies from the InfoSphere brand, with support for data cleansing, data masking, governance and so forth. In this context, it is worth pointing out that there are implications with respect to AI and machine learning that may not be immediately obvious. For example, consider automated decisioning. That is, where some sort of computerised algorithm makes decisions automatically, as a part of a relevant process, based on information received. That algorithm has to have absolutely trustworthy and unbiased data to work with, otherwise its decisions may be flawed, which could have very adverse consequences for your business. In other words, data quality – often thought of as just about names and address deduplication and cleansing – is fundamentally important within this context, and applies not just to structured data but also to semi-structured and unstructured data such as sensor readings and text. Note too, that in IoT environments there can be issues such as out of sequence, missing and duplicated readings, each of which must be catered for. In addition, you can get situations such as "sensor drift" where maximum and minimum sensor readings gradually increase or decrease (typically, because of extreme environmental conditions) and anomalous spikes (was this an event of interest or a loose connection?).

We would especially like to emphasise that one of the difficulties we have seen with companies trying to implement AI is a

Figure 4:
How traditional data federation works versus how IBM Data Virtualization works

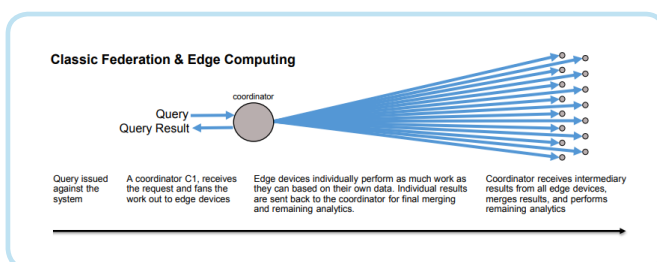
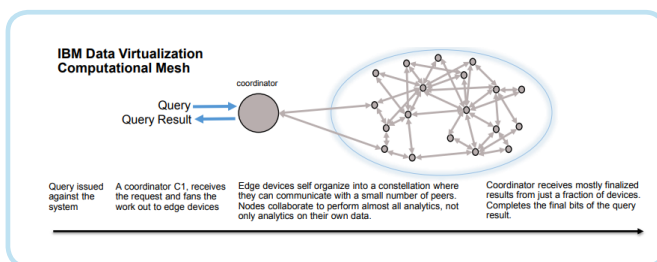
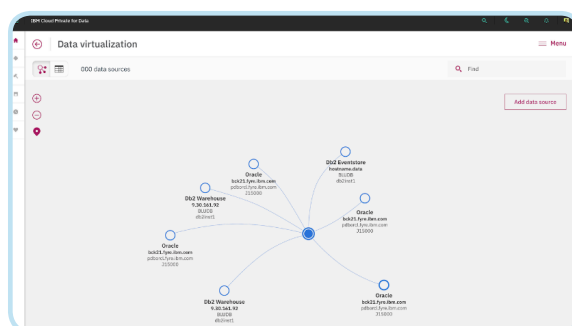


Figure 5:
How traditional data federation works versus how IBM Data Virtualization works



As can be seen, the key difference is that the computational mesh not only performs analytics locally but also within a local constellation. Given that moving the data across the network is the biggest issue with traditional data federation techniques this should significantly improve performance. Moreover, this should have a significant impact on costs because you need less infrastructure to get the same (or better) performance. Data sources supported by the computational mesh include the Db2 family (as above),

Figure 6:
The user view of IBM Data Virtualization



disconnect between data scientists who develop the relevant models, and those responsible for deploying those models in production. And, if data engineers are defined as those people who prepare data ready for analytics, while the data scientists actually do the science, then this applies to them too. So, what is required is “*AnalyticOps*” – analogous to DevOps – in the sense of bridging the gap between the analytics and their operational deployment, which is precisely what the collaborative environment in ICP for Data is intended to provide. And, we should add, being able to provide collaboration across this entire spectrum of roles and capabilities is only possible because of the breadth of IBM’s capabilities. To further enable this, IBM AI OpenScale has been built on top of ICP for Data and is available as a separate offering.

AI OpenScale has various capabilities that enable business to operate and automate AI across its lifecycle, with the aim of bringing enterprise-scale AI more easily within reach. First, the product allows you to monitor how the deployment of AI is affecting business outcomes, whether positively or negatively, no matter where models were developed or deployed.

Figure 7 shows a dashboard that is used to automatically detect unfair or inaccurate results based on user-defined conditions. As can be seen in this example, “*driver performance*” has two issues: it is both inaccurate (that is, performing badly) and is producing unfair outcomes. In total there are three accuracy alerts and six bias alerts.

A second function of AI OpenScale is to detect, and automatically mitigate, bias: see Figure 8. Various companies have announced such capabilities for training purposes but, as far as we know, AI OpenScale is the only such product that is also available at run-time, when it compares the performance of the deployed models with recommended de-biased models. No retraining is required. However, it will provide a way for a user to download and inspect the de-biased model before deploying it into production.

Thirdly, in addition to audit capabilities that allow you to trace the

recommendation behind any decision, AI OpenScale provides an explainability function that allows you to explore the factors involved in any decision, how much those factors contributed to the decision and how different factors might have changed the outcome. Confidence levels are provided for the transaction in question, and there are facilities to support the business in explaining a decision to a customer, regulatory body or other party.

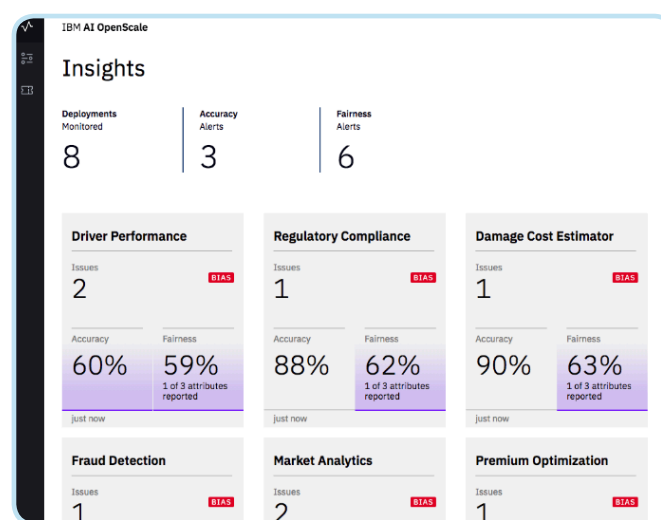


Figure 7: Accuracy and Fairness in AI OpenScale

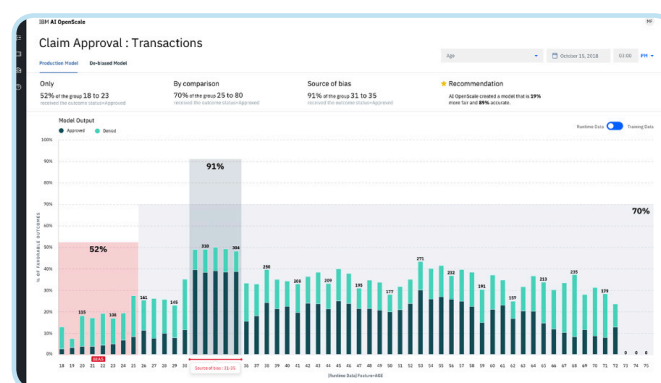


Figure 8: Bias detection in AI OpenScale

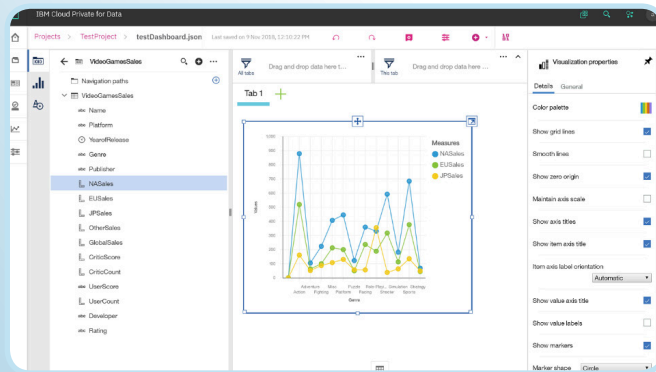
Under the covers the product also provides what is known as payload logging. This means that all the data being scored is logged and persisted to a database and it is against this data that the various dashboards and other capabilities of AI OpenScale are based. You can query the database using IBM Cognos or using third party tools, in order to get performance metrics. In addition, these metrics can also be exported and combined with application metrics to measure business outcomes.

Finally, AI OpenScale automates other aspects of the AI lifecycle through NeuNets (Neural Network Synthesis), which is currently available in betas. This is a function that will automatically create and recommend customised neural networks for AI deployments in any runtime environment, creating significant productivity gains for the data science teams tasked with building and retraining AI.

With respect to additional analysis capabilities ICP for Data leverages existing IBM technologies, including SPSS Modeler, Watson Explorer and Watson Studio,, which has specific capabilities for training, persisting and scoring machine learning models. However, there is a gap between trustworthy data and machine learning models. In this release, ICP for Data also includes support for IBM ILOG CPLEX Optimization Studio as an add-on. This is a prescriptive analytics solution that enables the development and deployment of decision optimisation models using mathematical and constraint processing.

Finally, we should also add that model management is supported by ICP for Data. This is important because what is the best model today may not be the best model next year. This is for two reasons: firstly, next year you will have more real data to work with and, secondly, because conditions and trends change over time. This means that you may need to replace one algorithm or model with another periodically. Ideally, this should be on a “hot swap” basis with no downtime. In any case, the performance of models needs to be monitored and changed when appropriate, hence the need for model management.

Figure 9:
Drag and drop
interface for data
visualisation in ICP
for Data



Roadmap

There are three notable plans that are worth mentioning. The first is the Unified Console, which is used to manage the “*Databases on Demand*” feature of ICP for Data. Significant enhancements to this are planned. Secondly, Cognos Analytics will be supported as part of the product’s “*analysing data*” capabilities. And, thirdly, Data Virtualization will include support for Excel, CSV and text files, MongoDB, SAP HANA, SAS, MariaDB, CouchDB, Cloudant, various Amazon and Azure databases, various streaming products, multiple mainframe environments (IBM and others), generic JDBC access and several third-party data warehousing databases.

Conclusion

At the Data Works summit in Berlin in April 2018 a straw poll was taken of the audience, asking how many of the attendees' companies planned to put data and analytics into the cloud. A, perhaps surprising, 34% of the 400+ people who voted, said their companies had no such plans. The truth is that no matter how impressive the hype is, there are many organisations that are reluctant to take that step, for a variety of reasons. That does not mean that they do not recognise the benefits of cloud-based computing but, currently, it is perceived to be a step too far. What ICP for Data offers is an in-between position: the benefits of cloud computing without the risk of moving data outside of your firewall.

However, this isn't all that ICP for Data offers: if you want to deploy machine learning – and almost everybody does – then you need an environment that facilitates that. IBM refers to this by saying that you can't have artificial intelligence without an information architecture (*"AI requires IA"*). And the problem with building an information architecture is that it involves many moving parts, many software requirements and many personas. To make this work requires that companies adopt AnalyticOps as a principle, and this requires not just a broad range of base functionality but collaborative support across all of the personas involved. Even though ICP for Data is still developing you can see that this is the direction in which the product is headed. It would be infinitely harder to achieve with a set of disparate products from multiple vendors.



Even though ICP for Data is in its first incarnation you can see that this is the direction in which the product is headed. It would be infinitely harder to achieve with a set of disparate products from multiple vendors.



FURTHER INFORMATION

Further information about this subject is available from www.BloorResearch.com/update/xxxx



About the author

PHILIP HOWARD

Research Director / Information Management

Philip started in the computer industry way back in 1973 and has variously worked as a systems analyst, programmer and salesperson, as well as in marketing and product management, for a variety of companies including GEC Marconi, GPT, Philips Data Systems, Raytheon and NCR.

After a quarter of a century of not being his own boss Philip set up his own company in 1992 and his first client was Bloor Research (then ButlerBloor), with Philip working for the company as an associate analyst. His relationship with Bloor Research has continued since that time and he is now Research Director, focused on Information Management.

Information management includes anything that refers to the management, movement, governance and storage of data, as well as access to and analysis of that data. It involves diverse technologies that include (but are not limited to)

databases and data warehousing, data integration, data quality, master data management, data governance, data migration, metadata management, and data preparation and analytics.

In addition to the numerous reports Philip has written on behalf of Bloor Research, Philip also contributes regularly to ***IT-Director.com*** and ***IT-Analysis.com*** and was previously editor of both ***Application Development News*** and ***Operating System News*** on behalf of Cambridge Market Intelligence (CMI). He has also contributed to various magazines and written a number of reports published by companies such as CMI and The Financial Times. Philip speaks regularly at conferences and other events throughout Europe and North America.

Away from work, Philip's primary leisure activities are canal boats, skiing, playing Bridge (at which he is a Life Master), and dining out.

Bloor overview

Technology is enabling rapid business evolution. The opportunities are immense but if you do not adapt then you will not survive. So in the age of Mutable business Evolution is Essential to your success.

We'll show you the future and help you deliver it.

Bloor brings fresh technological thinking to help you navigate complex business situations, converting challenges into new opportunities for real growth, profitability and impact.

We provide actionable strategic insight through our innovative independent technology research, advisory and consulting services. We assist companies throughout their transformation journeys to stay relevant, bringing fresh thinking to complex business situations and turning challenges into new opportunities for real growth and profitability.

For over 25 years, Bloor has assisted companies to intelligently evolve: by embracing technology to adjust their strategies and achieve the best possible outcomes. At Bloor, we will help you challenge assumptions to consistently improve and succeed.

Copyright and disclaimer

This document is copyright © 2019 Bloor. No part of this publication may be reproduced by any method whatsoever without the prior consent of Bloor Research.

Due to the nature of this material, numerous hardware and software products have been mentioned by name. In the majority, if not all, of the cases, these product names are claimed as trademarks by the companies that manufacture the products. It is not Bloor Research's intent to claim these names or trademarks as our own. Likewise, company logos, graphics or screen shots have been reproduced with the consent of the owner and are subject to that owner's copyright.

Whilst every care has been taken in the preparation of this document to ensure that the information is correct, the publishers cannot accept responsibility for any errors or omissions.



Bloor Research International Ltd
20-22 Wenlock Road
LONDON N1 7GU
United Kingdom

Tel: **+44 (0)20 7043 9750**
Web: **www.Bloorresearch.com**
email: **info@Bloor.eu**