DNV GL STRATEGIC RESEARCH & INNOVATION
POSITION PAPER 4-2014

BIG DATA
- the new data reality and industry impact
“Big data” promises business benefits in terms of timely insights from data, real-time monitoring and forecasting of events, more fact-based decisions, and improved management of performance and risk. We see the rise of a new industry, largely Internet-based, the main asset of which is data, and which offers completely new value propositions based on data and insights from data.

The main enablers of these changes are: increased connectivity, the proliferation of small, cheap sensors, and new technologies for capturing, storing, accessing, processing, and analysing vast amounts of heterogeneous data. These enablers are, in sum, the essence of the “big data” trend.

Big data has already had a considerable impact and changed the competitive landscape in many industries. We believe that it will also have significant impact within the industries in which DNV GL operates.

One of the main trends that we see in all our industries is the increased use of sensor data; these data are used for performance monitoring, condition monitoring and optimisation of systems and components. This trend has paved the way for a new type of service provider, helping their customers by implementing “intelligent operation” of systems or across systems. Key aspects of the services offered are: performance monitoring and optimisation, condition-based and predictive maintenance, and optimisation across systems and system overview of operations.

The purpose of this paper is to establish a framework for identifying opportunities and threats in our industries and discuss possible changes to business models and operation modes. Through the framework and examples from various domains, we hope to trigger discussions within our industries on how big data and the new data reality can contribute to make our industries safer and more sustainable.
“There is a disruption on the horizon, in fact it is upon us already...”

CONTENT

EXECUTIVE SUMMARY ................................................................. 2
INTRODUCTION ................................................................................... 4
THE NEW DATA REALITY ................................................................. 7
TRANSFORMING INDUSTRIES AND CHANGING PLAYING FIELDS ...... 10
    Maritime .................................................................................. 14
    Oil and Gas ............................................................................. 20
    Energy ..................................................................................... 22
    Health ..................................................................................... 26
IMPACT ON BUSINESS MODELS ................................................. 30
IMPACT ON MODE OF OPERATION ............................................. 34
REQUIRED CAPABILITIES .............................................................. 38
CONCLUDING REMARKS ............................................................... 40
ACKNOWLEDGEMENTS ................................................................. 41
BIBLIOGRAPHY .............................................................................. 42
INTRODUCTION

There is a strong trend across all industries and sectors to move towards fact-based decision-making, and to use data analytics for managing performance and risk. At the same time, we have seen the rise of a new industry, largely Internet-based, whose main asset is data, and which offers completely new value propositions based on data and insights from data.

Increased connectivity, new capabilities for capturing, storing, processing, presenting, and visualising data, and, in particular, transmission of large volumes of varied data at high velocity, have developed significantly in the past few years. Sensors have become cheaper and more commonplace, and deployment of sensors in all possible settings has increased dramatically and will continue to do so at an even higher rate in the years to come.

“There is a trend towards automation of problem solving and decision making” – Espen Andersen, Norwegian Business School (BI)

This new reality has changed the playing field of many traditional industries and has been the platform for many new businesses. It will also have a significant impact within the industries in which DNV GL operates, requiring many players to shift to new business models and operation models. Sensor data is increasingly used for performance monitoring and condition monitoring of components and systems. This trend has paved the way for a new type of service provider, helping their customers by implementing “intelligent operation” of systems or across systems.

Key aspects of these services are: performance monitoring and optimisation; condition-based and predictive maintenance; optimisation across systems and system overview of operations.

We have chosen an example from the aviation industry, Taleris, as an illustration of such a new service provider. Taleris is a joint venture by General Electric (GE) and Accenture that provides intelligent operations services across systems in the aviation industry. The service uses its customers’ own data streams and facilitates optimisation of asset utilisation and predictive maintenance of critical systems (see fact box for more details).

This initiative by GE is only one of a new suite of technologies and services that it says will help airline, railroad, healthcare, manufacturing and energy companies increase productivity and reduce costs; it claims it will save $150 billion in waste. GE believes that machine-to-machine (“M2M”) advances will fundamentally alter the way in which businesses operate. GE calls the initiative the "Industrial Internet," and believes it to be the next revolution in technology.
“The Internet has changed the way we consume information and talk with each other, but now it can do more,” GE CEO, Jeff Immelt, said. “By connecting intelligent machines to each other and ultimately to people, and by combining software and big data analytics, we can push the boundaries of physical and material sciences to change the way the world works.”

This is an example of a trend that can challenge the current business models of knowledge-based service providers. Such services could possibly be partly replaced by automated monitoring, insights, advice and decisions based on continuous sensor data streams.

The “Industrial Internet”, or “The Internet of Things” not only enables data from components and systems to be available across traditional barriers, but also allows systems and components to exchange and act upon information from each other. This opens up a whole new level of automation of systems and decisions.

In the shorter-term, knowledge-based service providers may be expected to be able to work efficiently with large data sets and apply proprietary models and expertise to any data sets that are brought into projects. The ability to execute such projects efficiently and to optimise the service delivery can be a deciding factor for staying relevant and preferred in the new data reality.

Taleris, a joint venture company of General Electric and Accenture, is operating with progressive technology assets and capability from both Accenture and GE. Taleris is dedicated to providing airlines and cargo carriers around the world with intelligent operations services focused on improving efficiency by leveraging aircraft performance data, prognostics and recovery. The business case for Taleris is that they believe that they can save airlines 10% of their unscheduled cancellations, which, for the big operators, could save them in the order of hundreds of millions of dollars.
In the long-term, one way to sum up the impact of the new data reality on our future is that for any situation or process, where the outcome may be influenced by the availability of more relevant information, somebody will offer to make this information available. The new data reality will continuously offer new capabilities for making this happen at an even lower cost. When you get in the car to go to work, you would like to know which route to take, based upon the route that everybody else is taking that morning (Google offers this today). If you have merchandise to sell, you would like to know if it is attractive to the person who is passing your shop window right this minute (push messages based on location is already being offered) etc.

Any actor that is currently involved in delivering data, information, and knowledge, will be affected by this, and it will change the competition field.

The changes described above will be challenging, but, at the same time, this new data reality will provide huge opportunities for existing knowledge-based service providers and new players.

This paper seeks to give a short introduction to the topic of big data and the new data reality and give a brief overview of how it is already impacting the particular industries of maritime, energy, offshore, and health. It will also have a significant impact within the industries in which DNV GL operates, requiring many players to shift to new business models and operation modes.

In order to illustrate our points and tell a more interesting story, a number of relevant examples are included. These examples are taken from various industries, including those of maritime, offshore, energy and health. We hope that our readers will find these interesting and that they inspire ideas, discussions, and innovations.

It must also be noted that while this paper is intended to be comprehensive, it is not meant to be exhaustive, nor can we present the entire picture of either big data in general or any of the industries discussed. The ideas and examples presented are mainly the result of reading and reviewing publicly available information and documents, in addition to workshops and interviews with key experts and practitioners with experience in creating value from big data (see acknowledgements).
Since the beginning of time, human activity has generated data, and the notion that the amount of data is growing “too fast” is not new. Furthermore, using data and statistics to gain knowledge has a history stretching back almost 400 years. Nevertheless, over the last 10 years or so, two major factors have changed radically: availability and usability. We are in the middle of an explosion of available data, and the ability to store, combine, and analyse these data is now available to anyone. In combination, these two factors provide endless possibilities for creating value by producing actionable knowledge from data. This is the essence of the current “big data hype”.

There are several drivers for the increase in data volume. The price of sensors is falling rapidly and therefore the quantity of devices and equipment that can sense and log environmental factors has increased dramatically. One of the main drivers for this development is the inclusion of digital cameras, GPS receivers, motion sensors, etc. into the mobile phone consumer market; this has made sensors a commodity. In addition, both people and devices are increasingly networked and connected, such that it is possible to gather datasets for systems in time and space.

The development of Web 2.0 technologies and the resulting explosion in social media has generated data about human interactions, interests, and sentiments at tremendous rates. Another important trend is that public institutions, governments, and private companies make datasets, application programming interfaces (API), and modelling results available for free on the Internet. By making these “public data” sets freely available, anyone can innovate by presenting them in new ways or by combining them with other data. Prime examples of such “public data” are the map data from Google and weather forecasts from the Norwegian Metrological Institute, yr.no.

But just “having” data is not enough; the ability to store and process the data is as important as the data itself. And these are the final two pieces in the big data puzzle. The falling prices of processing power and storage space have resulted in the necessary computing power becoming affordable to many more. In addition, the major relevant innovations in software, algorithms and statistical methods have been developed or implemented as free software. These two factors have significantly lowered the material cost of analysing data, and have effectively made the necessary tools available to anyone with the skills and will to deploy them.
Managing and using data have always been challenging; but when is data no longer just data, but “big data”? The most common definition is the 3 (+ 1) Vs of big data: Volume, Velocity, and Variety. These ‘3 Vs’ were originally coined by Dough Laney, and IBM has introduced the important fourth dimension: Veracity.

**Volume**
The size of data sets is ever-growing (for many enterprises it is already in the petabyte range); this poses a technical challenge in terms of storing, finding, and analysing the data.

**Velocity**
For some problems, making the right decision in time is the most important thing. Who wants to cross the street using data that is 2 minutes old? And to stop credit card fraud, the decision has to be made before the transaction goes through. Tackling such problems in real-time adds additional requirements to the system.

**Variety**
Data can be of any type. It can be structured or unstructured. It can be text, time series from sensors, click streams and log files from web sites, multimedia, position data from GPS loggers, etc. The challenge is to combine the various data and to set up the system to integrate new data as it becomes available.

**Veracity**
Veracity is about trusting the data. It has been argued that having access to massive amounts of data will, in itself, guarantee the right answer. But this is grossly over-optimistic; having the wrong data or biased data will probably give you the wrong answer. Trusting the data is therefore an important aspect of any data problem and as the size and complexity of the data grows, the problem of trusting the source and quality of the data also gets harder. This makes data governance even more important.

**Too big, too complex, and too fast**
Another similar and common definition of “big data” focuses on the technical limitations of currently used computer systems and software. Following this line of thought, you have a big data problem when the current (standard) toolset limits what you can do with the data. Use of new hardware, software, and/or algorithms is required to solve the problem.

These kinds of definitions are relative, and what is currently considered as big data should be expected to vary and change with time. Nevertheless, these definitions are quite useful in terms of demystifying “big data” and making the technical aspects more concrete.
The new data reality has created novel possibilities for business and value creation. This is the core of the big data shift; the ability to use data to obtain actionable knowledge, insights, and predictions and, eventually, to automate this process.

One example of such predictions could be that some major US insurance companies have started estimating their exposure to claims during hurricane season in real-time, by combining real-time data about hurricanes and their projected trajectories with the locations and contract details of their claims. Insurance companies are also increasingly using a wider set of data for assessing risk and setting premiums.

Other typical applications of big data analytics include:

- Targeted marketing and sales
- Risk management
- Health diagnostics and targeted treatment
- Process and value chain optimisation
- Component and system condition monitoring and predictive maintenance

Companies in established industries, such as telecom, retail, media, healthcare, insurance, finance and transport/logistics, are being transformed by utilising large data sets about their customers and internal work processes. Insights from these data streams are used for better assessment of utilization, to improve operational efficiency, and to increase sales and the efficacy of marketing.

**VALUE CHAIN OPTIMISATION**

Big data capabilities enable companies such as FedEx (see example later), Wallenius Wilhelmsen (car carrier shipowner, see example later), and Maersk “Daily Maersk conveyor belt” (container shipowner, guaranteeing daily shipments between major Asian and European ports), to streamline their supply chain and improve efficiency. The optimisation may include predictions using weather data, tracking data, traffic data, and other data that may influence the movement of cargo and parcels.

**RETAILERS**

Retail companies such as Walmart combine point-of-sales data (what’s in our shopping basket) with other customer and operational data to increase
advertising efficiency and to optimise stocks and logistic chains. Walmart has taken this further by giving suppliers access to their systems and, in effect, has transformed Walmart into a platform of trade in which ownership is not transferred until the customer pays for the goods.

**SPORTS**
Another story, on how the use of data literally changed the game, is told in the book and movie “Moneyball”. This is the story about how available data and analytics were applied in a new way that challenged, and entirely altered, the traditional way of valuing baseball players. This has completely transformed the way in which players in the US National baseball league are valued, and similar methods are also being applied in other sports.

**NEW INDUSTRIES**
Armed with vast amounts of data and affordable capabilities for storage and processing, existing and new companies have been able to create new business and revenue streams from selling data or delivering services that are based on data analytics. There are many different roles that can be adopted in this ecosystem, ranging from simple “raw data providers” to full implementers of analytical solutions for the customer. Various aspects of the different skills/knowledge that are need, and the corresponding strengths of the value propositions for each of these roles are discussed in the chapter ‘Impact on mode of operation’.

One example of raw data providers is a number of US banks and credit card issuers that are now selling raw data about their customers’ shopping habits and spending behaviours through intermediaries like Cardlytics. Retailers use these data to offer targeted discounts to customers, and pay for the service by a commission (10-15 %) for each sale.

The most prominent examples of new use of data occur in the new Internet-based industries. Here, companies exist that have created big business based entirely on access to data and their ability to use data. These companies have moved the technology forward and have created a whole new marketplace for advertising and innovation on the data generated through people’s activities on the Internet.

One of the largest online retailers, and cloud and digital content providers is Amazon. Amazon was originally founded as an online bookstore in 1994, but has evolved into a complete retail platform/store. Amazon has also been able to leverage its unique position as “the world’s largest bookstore” to become one of the world’s leading digital content providers, basically creating the whole e-book market (the Kindle ecosystem). One of Amazon’s key success factors was the early use/adaption of machine learning and big data for targeted marketing, with as much as 30 % of the sales driven by automatic customer recommendations. Amazon has also been able to bundle and sell its internal development efforts and is now the world’s largest cloud computing and storage provider. With the recent acquisition of The Washington Post, Amazon’s co-founder and CEO, Jeff Bezos, is now also venturing into the publishing business.
Google’s main business model is based on its ability to use search phrases to connect advertisers effectively with potential customers. However, they are also continuously innovating and finding other uses for their data.

Twitter mines the content of Tweets to increase the value of advertising for their customers by enabling them to target relevant customer groups and measure the impact of their marketing campaigns. Twitter also sells raw data to insight aggregators and academia that can use the data to create new services and to research anything from epidemics to human behaviour.

PRIVACY: CHALLENGES AND OPPORTUNITIES

The new data reality and our new capabilities of handling the data represent both opportunities and challenges to society as a whole. The data and technology themselves are value-neutral. However, the use of these is not, and using these opens up new possibilities for surveillance, abuse, and big brother dystopia. We all leave behind a digital trace (often called “digital exhaust”), which could be used to give a near real-time view of where we are, with whom we communicate, what we like, and which opinion(s) we hold. If these new powers are abused, the result will be loss of privacy, and, ultimately, free speech and the foundations of democracy will be threatened.

Another aspect of this is that in many cases the potentially sensitive data are in the hands of private companies. For some of these companies (like Google, Facebook, Instagram, etc.), the data about their users is one of their main assets, and already a part of the product that they sell (targeted marketing). But even if this is not the company’s main business, the ability to protect the privacy and data of its customers and users will represent a major risk. Potential privacy issues, data protection, and privacy-related risks should therefore be addressed early in any big data project.

Despite these challenges, the utilisation of big data could also have huge benefits for society at large. By using demographic and mobility data we can obtain important insights into human behaviour, including: traffic patterns, crime trends, disaster responses, early indicators of poverty and social unrest, etc. These, in turn, can be used by business and policy makers to create better, safer, and more efficient societies.
FedEx stays competitive by using data-based value chain optimisation for all its processes. At the core of the system lies the package tracking system that tracks and monitors every phase of the delivery cycle. Data from this system enables the company to maximise utilization of assets and personnel, while delivering on time. FedEx also gives customers direct access to information about on-going deliveries, adding value for the customer by allowing better control of their own value chain. The FedEx system is integrated into a central system that collects and coordinates data from: airlines, connection hubs, positions of individual vehicles, weather forecasts, and real-time traffic information. This allows for real-time routing information to be pushed out to individual drivers and optimisation of pickup/delivery and asset utilization. FedEx has also been able to leverage on this infrastructure to expand and enhance the product. One example of this is the sensor-based SenseAware Service; this is a parcel logger that tracks temperature, relative humidity, light exposure, and can detect if a parcel has been opened. This provides the customer with a more complete picture of the whole shipment process and ensures the quality of the delivered product (adds value for the customers’ customers).

DATA DEPENDENCY - A CATCH
Data dependency arises when a service delivery relies upon critical data from a sub-contractor, or on publically available data over which the service has no governance. There is already a clear trend for near data monopolies to establish within each industry (e.g. providers of ship tracking data, ship parameters, financial indexes, insurance risk factors, customer behaviour etc.).

Building high value services with a strong dependency on data from a ‘near monopoly provider’ is obviously not all straight forward and will require considerable efforts directed towards reaching mutually acceptable agreements.
Some of the current changes in the maritime industries seem to be propelled, by the new data reality and the increased connectedness of ships and players in the maritime industry (more satellites and reduced prices enable 24/7 ship-to-shore connection).

Looking more closely, traces of the new data reality can be found within the following main areas of the shipping industry:

- Technical operation and maintenance
- Energy efficiency (cost and environment)
- Safety performance
- Management and monitoring of accident and environmental risks from shipping traffic
- Commercial operation (as part of a logistics chain)
- Automation of ship operations (long-term)

**TECHNICAL OPERATION AND MAINTENANCE**

The emerging connectedness, increased use of sensors in components and systems onboard ships, and improved big data analytics capabilities enable component and system vendors to monitor conditions remotely and advise ship management on system operation and predictive maintenance. These approaches may eventually result in equipment and systems being leased or the possibility of subscription to specific functions performed by the system and components as is occurring in aviation, land-based power generation and the automotive industries. Current examples include: Electronic Power Design Inc., which provides diesel electric propulsion for some of the most sophisticated DP3 vessels in the world; the Rolls Royce Hemos System, a system that draws on their system in aviation, for transmitting sensor data from various components to land-based service centres where system specialists create health asset reports for the customers; and Wärtsilä’s Propulsion Condition Monitoring Service, which enables detection of maintenance requirements some 2-6 months in advance.

As with the example of Taleris, the independent intelligent operations service provider for the aviation industry, there are already examples in the maritime sector. One example is US-based Engineering Software Reliability Group (ESRG) that has built a commercial cross-silo analytic platform for ships, the OstiaEdge® Monitoring Suite for real-time analytics. This is derived from a major engagement with the US Navy, capturing and analysing more than 5000 data points for more than 120 Naval ships.
The OstiaEdge® Monitoring Suite
This system enables vessel owners and operators to use data and analytics to identify issues prior to failure, reduce maintenance costs, and prioritize maintenance spend across a fleet of assets. This technology has been developed over the past decade and has been tested in a variety of environments, including the US Navy, and can monitor a wide variety of equipment onboard commercial and defence assets, ranging from gas turbines to diesel engines, to a variety of auxiliary equipment.

OstiaEdge® enables improved use of onboard data to help implement a Condition-Based Maintenance (CBM strategy). This creates value for the operators, managers, and owners in a variety of ways:

- Identifies issues prior to failure with automated predictive diagnostic analytics
- Troubleshoots issues more rapidly with real-time analytics onboard and onshore
- Plans maintenance periods more effectively with real-time transparency into equipment health and performance
- Prioritizes maintenance spend across a fleet of assets to maximise return on investment.

The FRAM Project
The FRAM Project is an industry initiative inviting shipping companies to voluntarily measure, report and verify CO₂ emissions from their vessels, and use the resulting data and insights to become more energy efficient and environmentally conscious.

The partners of the FRAM Project consist of a coalition of ship-owner companies, an environmental NGO, a maritime professional service company and a ship industry association.

ENERGY EFFICIENCY AND ENVIRONMENTAL PERFORMANCE
For a long time early movers among shipowners have worked actively to improve their fuel efficiency for cost reasons. Maersk is a class act to follow in this area with proprietary solutions for collecting, analysing, and presenting data from all the vessels in their fleet to drive continuous energy efficiency improvement. In the general market, there are many smaller niche players offering solutions to support shipowners in this quest. Marorka and SRG group are examples of niche players with tailored solutions and services in this area.

Ships will be impacted by new requirements for emission control by 2018. Both IMO and EU are establishing guidelines for MRV reporting (Monitoring, Reporting & Verification) and has initiated projects to develop systems to manage such reporting.

Interest organizations such as the Norwegian Shipowners’ Association, are taking the lead in encouraging and enabling their members to document emissions and prepare for the new legislation (the FRAM project). Furthermore, consumer concerns have prompted retailers like IKEA to require that shipping companies document their emissions for their product footprint documentation.

Complying with the new requirements requires extensive and continuous reporting from the ships. Ship managements will seek to meet the requirements by taking the right measures in their operations and will need assistance. Thus there is likely to be a boost in demand for advisory services in these areas.
The increased availability of data will be advantageous to those that can deliver advice from an analytic platform that is continuously fed data captured onboard the vessels from existing systems or from purposely installed data loggers.

To date, this market has been dominated by niche vendors like Marorca and Maersk (offering a version of their own internal system in the market), but class societies like ABS, NK and DNV GL have been well established in this field for a long time with traditional advisory services founded in their in depth knowledge and competence and class customer base. The class societies are now actively seeking to move this to a more analytically based platform. Through the acquisition of a PMS vendor ABS has attained a position on board the vessels for data capture and analysis of performance (ABS Nautical Systems). Societies like DNV GL and NK have developed onboard reporting solutions from scratch or together with partners like NAPA.

A more recent trend is that actors that already offer data capture onboard ships are extending their offerings to providing energy efficiency advice. Examples here include component-manufacturers/system integrators like Rolls Royce and Wärtsilä and other independent intelligent operations service providers such as ESRG. In addition, vendors of other onboard systems such as routing and navigational safety decision support, like Amarcon with their Octopus suite, are expanding into this space.

**SAFETY PERFORMANCE**

Safety performance, verification and assessment have been key domains of the class societies through their role as standard setters, delegated body and advisory offerings. In addition, there are a variety of complementary assessment schemes like vetting inspection schemes for the tanker fleet and various condition assessment programmes. Traditionally, these offerings are based on empirical and analytical rules and ad hoc processing of disparate datasets, but more data-based offerings for assessing and monitoring ship safety are gradually emerging. For a while actors like RightShip have offered safety rating services that are based on available data outside shipowners’ organizations (vetting, port state control, class, ship registration data, etc.). Others, like Lloyd’s List Intelligence, offer advisory services founded in similar datasets. Others, such as OceanIntelligence, are more concerned with financial risk aspects and use similar datasets as input to, for example, debt and credit analyses. Shipping KPI is another major initiative that provides independent safety assessments and performance management schemes for the shipping industry based on a whole range of data reported manually from ships into a standardized reporting system.

With the new data reality, data from a multitude of sources onboard ships may become available for safety assessment activities on a continuous basis.

The drivers for such availability of data will certainly be component and system vendors/integrators as previously described. However, vendors of onboard operational support systems such as weather routing,
hull monitoring etc. that already collect and process vast amounts of data onboard the ships will also play a part, as the data will be attractive for secondary use in safety assessments and predictions (e.g. fatigue life of a hull girder).

In addition, regulatory bodies like the EU will also push for increased transparency of safety of shipping, by requiring disclosure of more data and information directly from the ships.

The MRV for environmental performance has already been mentioned, but EU also has an initiative for developing methodologies for improving existing risk management procedures and processes for inspections, incident detection, and recording, compliance monitoring, contingency plans and emergency responses (ref. CP - Call: FP7-TRANSPORT-2013-MOVE-1). This includes dynamic collection, processing and use of real-time information from ships, also considering vessel health status information through real-time information from ‘intelligent’ sensors. We believe that there is potential for creating new business opportunities around the use of these data.

Lloyd’s List Intelligence

“Lloyd’s List Intelligence is founded on the heritage of Lloyd’s List plus a global network of specialist sources of business-critical maritime data. It brings together the expertise of a global staff of maritime analysts and journalists with an extensive system of shore based and satellite intelligence gathering to create a complete information support service for the maritime industry.

One of their products is a Hull Risk Profiler. Hull Risk allows you to immediately assess the risk associated with an owner, manager, fleet or an individual vessel.”

MANAGEMENT AND MONITORING OF ACCIDENTS AND ENVIRONMENTAL RISKS FROM SHIPPING TRAFFIC

Most busy shipping lanes and ports currently have traffic control centres to manage traffic in real-time to avoid collisions and accidents. We expect to see more of this type of activity, and they will expand to include more complex risk models, enriched with data from multiple sources, including data from the ships themselves. The risk models will be able to combine real-time information about course, speed, weather etc. with facts concerning the safety condition of the ship, cargo, and other elements that influences the likelihood or consequences of an accidents, and determine the shore areas most exposed to the risk.
The services will also expand beyond the busiest shipping lanes to, for example, position emergency responses such as tug and oil recovery services. These will reflect dynamically the volume, safety conditions, and weather conditions through monitoring and predictions.

Ship tracking data, in general, is emerging as a platform for many new services. Service providers range from mobile ship finder apps, intended for the general public, to more professional services aimed towards the maritime industry. Examples of such applications could include continuous monitoring of emissions from shipping, and real-time monitoring of accident and emergency response risks (oil recovery, tug preparedness, pilot schemes etc.). Operational and navigational risk monitoring is one of the problems addressed in the European Maritime Safety Agency (EMSA) EU initiative. We believe that there are clear business opportunities for actors that can combine in depth knowledge of continuous advanced analysis of ship traffic with other data such as weather, ship parameters, ship equipment, cargo, safety condition etc.

Through the efforts of both governments and commercial companies like exactEarth, the global network of satellites will soon cover all the oceans and the poles. The availability of ship tracking data can change the way in which many services are delivered, including monitoring the safety of ships based partly on knowledge of weather conditions to which the ship has been exposed at any time.

We also see the emergence of other data aggregators that compile data about the ocean in general. There are millions of sensors throughout the coastal areas and even across the oceans deployed by academic and research institutions, national bodies and by commercial organisations in order to monitor things such as water temperatures, currents, waves, chemical compositions, sounds, fish movements etc... Some of these may also become influential in the maritime and oil and gas domains. One example is Marinexplore who offers the combination of a public and commercial big data platform for everybody with data from the oceans to contribute, combine and analyse. Another example is Google who is expanding it’s Google Earth platform to include oceans data (Google Ocean).
COMMERCIAL OPERATION OF SHIPS AND FLEETS
Shipping companies have begun using analytics as part of the means to optimise their place in the value chains, and to optimise their own operations. Combining and analysing data about availability of cargoes, available space to take cargoes, port slots, weather, ship performance, fuel prices, etc. does in some cases materialize as a huge business case (see fact box for Wallenius Wilhelmsen Logistics).

AUTOMATION OF SHIP OPERATIONS
Increased availability of sensor data and the ability to process data across systems will drive ships to become more and more automated, and maybe one day we will even see autonomous ships controlled from big onshore operation centers. This trend will challenge regulatory aspects, as well as safety performance assessment and verification processes and roles. This topic is further covered in another strategic position paper, “The Digital Ship” to support the DNV GL 150-year anniversary portfolio of showcase projects.

“Real-time analysis helps us make contingency plans for different situations – for example, a sudden rise in oil prices, or a closure of a major port – enabling us to adapt our shipping schedules accordingly.” - Donald Duggan, Head of Trade Strategy and Performance Management, WWL.

The solution enables planning for best- and worst-case scenarios as well as the standard forecast, enabling the company to respond more effectively to normal demand fluctuations as well as emerging situations, such as the 2011 Japanese tsunami. A six-month forecast helps to plan routes for individual vessels, and the 12-month and five-year forecasts help with capacity planning (for example, whether to charter or build new vessels).

Wallenius Wilhelmsen Logistics are saving in the order of 100 million dollars annually from route planning and general reductions in speed.
In the oil and gas industries, companies have grappled with big data challenges for quite some time, for example to process seismic data. The new trend is that data and analytics are playing larger roles in other parts of the operations such as integration of offshore operations across offshore assets and system barriers, improving efficiency of drilling operations and optimizing production processes etc. These uses of data pave the way for further improving reliability and managing risk across the whole enterprise, including the risk of environmental impact as exploration and production move into more sensitive areas.

**Integrated operations**

Integrated operations (IO) is arguably one of the main drivers for big data in the oil and gas industry. The main drivers behind IO are the perceived cost reductions and increased efficiency from bringing together data and expertise across assets and disciplines in onshore operational centres. IO are already very data intensive and big data technologies will play an important role in the further development of these operations. At the same time, manufacturers and system vendors are moving towards remote monitoring and diagnostics. This fits well with the IO trend, and, in combination, will allow operators and manufacturers to move away from periodic service regimes, towards systematic, risk-based maintenance strategies.

We believe that this trend will pave the way for new actors in this domain. Both data aggregators and intelligent operations service providers (like Taleris in aviation) offer advanced analytics and optimisation across systems and supply chains.

**Efficient drilling and production**

The cost of drilling is a major concern for oil companies, and increased automation has the potential to increase the efficiency and reduce the cost dramatically.

Sekal AS (see fact box) is an example of a company that offers model-based real-time monitoring and control of the drilling process using data from the well itself and other systems involved in the drilling process. Again, analysing data from across system barriers can provide insights that are not visible in the individual systems alone.

There are probably many more examples of data and analytics being used in ways that already impact on the ability of operators to work more effectively. For drilling this trend will strengthen as cost focus increases in the years ahead.
Managing risks

Conoco Philips use of big data analytics to manage the risk from ice in real-time (see fact box) is a good example of how real-time analytics has enhanced the ability to manage risks in such a way that it has a direct and positive impact on production yield.

Another trend is to automate environmental risk management through real-time monitoring of environments around offshore drilling and production installations. DNV GL, together with Statoil, Kongsberg, and IBM, is participating in one of the first projects to establish the technology and methodology for such solutions in the “Integrated Environmental Monitoring” (IEM) project. The goal of IEM is to develop and deploy a system that will access data both from the operations and from the surrounding environment via a combination of data imports and sensors. Through implementing environmental models on the data streams, operators and stakeholders are able to evaluate continuously the risks for stationary and transient (pelagic species) organisms. In the future, such systems may well become a pre-requisite for compliant operations in sensitive areas like the Arctic.

PIPELINES

Pipeline systems represent a significant financial investment and constitute a key element of the industrial and public value chain. The extensive number of new pipelines and the vast number of ageing pipelines means that there is increasing focus on asset integrity management and life extension of pipelines. Part of the basis for this is an increasing amount of pipeline inspection data and sensor data. These activities in themselves require processing of huge amounts of data from in-line inspections.

There are various software solutions that already address pipeline integrity management such as Associated Technology pipeline Ltd and DNV GL software with their Synergy pipe solution.

There is reason to believe that future solutions will become more real-time and based on monitoring and predictions on more and more disparate datasets (including weather, shipping traffic, seismic activity, political situations etc.)

Statoil has recently joined SåkorninVest as an investor in the newly established company Sekal AS to commercialise the products DrillTronics and DrillScene. These products are aimed at reducing risks and costs in drilling processes - during both exploratory drilling and production drilling - and employ advanced physical models that give a much fuller and more refined real-time picture of the state of the well than has been possible previously. They are used to manage drilling control systems in real-time and to catch drilling problems developing in the well several hours in advance, such that undesirable occurrences related to control of the well and the drilling can be counteracted and prevented much earlier. The result is a safer, more efficient drilling process. After successful testing and verification offshore during the last two years, DrillTronics and DrillScene are ready for commercial use internationally. Both these products have been developed by the International Research Institute of Stavanger AS (IRIS), with financial support from Statoil, BP, and the Research Council of Norway (Demo 2000).

In order to help improve productivity at its Arctic facilities, ConocoPhillips wanted to monitor and forecast ice floe movement in or near real-time. Big data technologies helped the company to collect and manage thousands of data points per second from multiple sources. Through predictive analytics, ConocoPhillips can visualize in real-time the position of ice floe near its Arctic facilities. The result is increased productivity and yield as the company has been able to extend the drilling season by weeks.
GRID MANAGEMENT
The energy mix is changing, with increasing inclusion of volatile renewables. This poses considerable challenges for grid management, load balancing, and energy security.

In addition, implementation of “smart meters” and “smart grids” is spreading. The huge investments in these technologies create enormous amounts of data, which, in turn, create business opportunities for companies that are able to analyse and act upon this knowledge. All the large grid players, such as GE, Siemens, ABB, Alstom, Schneider Electric, and Eaton/Cooper Power, are investing to establish businesses and services for “smart grids”. This can be exemplified by GE’s cloud-based “Grid IQ Solutions as a Service”. In this, for a structured fee, they offer to utilities hosting of their data on dedicated servers in its GE Digital Energy data centre in Atlanta, with the opportunity for the utilities to choose from a list of products and functions that they would like to deploy.

In this emerging market, the seven highest valued data analytics applications are:

- Grid analytics
  - Outage management
  - Voltage optimisation
  - Asset management

- Consumer analytics
  - Revenue protection
  - Load forecasting
  - Detailed customer segmentation
  - Energy efficiency, advice and assessment

In addition to the change in energy mix, there is also a growing deregulation of markets; the emergence of consumers as producers and of micro-power plants will make the picture even more complicated.

Utilities face some regulatory challenges and uncertainties in turning over key parts of their operations to a third party. At the same time, they are under pressure to meet a whole new array of requirements, including smart grid security and data privacy, which may well be better managed by a big central provider than individually by each small utility.

Ultimately, services, most of them probably founded in big data analytics, will be the key to unlocking the small utility smart grid market.
The smart grid market is an open playing field where startup companies like Opower already collect more than 90 billion meter readings annually from more than 75 utilities in the US. Opower offers to do four things for utilities: increase behavioural engagement, deliver better energy information to customers and utility workers, fine-tune retail and marketing offerings, and provide home control - all from a scalable Hadoop big data platform. In these service offerings they perform the roles from data aggregation right through to insight offerings (see Figure 4).

**Condition monitoring and predictive maintenance**

As with the trends in maritime, offshore, and aviation industries, gathering and processing data from across systems will enable a further shift towards predictive maintenance.

Accenture’s offering to the utility Hydro One is an example of the potential impact of even just basic analytic applications on data across the customer organization (see fact box).

Again in the energy industry, component manufacturers such as ABB, Wärtsilä, Rolls Royce etc. are already offering remote monitoring of their products through their remote control centres.

Similarly, new entries may also be expected, like Taleris (in aviation), and these will offer insights and advice based on their abilities to process data across system and component barriers.

As components in the energy segment will be easier to connect through the ‘Industrial Internet’, a higher degree of automation of complex decisions might also be envisaged.

**WIND**

There are already many untapped sources for insights and performance enhancements in the data available in the wind industry today. There are huge amounts of weather data available from field assessments and field operations, and performance and technical data are available from installations. Initiatives to develop solutions for improving operational effectiveness exist, underpinned by analytics across system barriers, and provide opportunities to increase the profitability of the assets.

Accenture developed an Asset Analytics solution for Canadian utility Hydro One, based on Hydro One’s existing ERP, GIS, data historian, and other legacy asset systems. It helps them with improving CAPEX/OPEX investment decisions, provides a common understanding of asset risk, supports regulatory compliance, and facilitates communications. “If it generates even one per cent in savings through better decision-making, it pays for itself in a year.” Rick Stevens, Hydro One Vice President of Asset Management.

**GE’s “Grid IQ Solutions as a Service”**

General Electric is taking the smart grid platform to the next level. GE’s smart grid-as-a-service pitch can be illustrated as follows: the city utility of Norcross, Georgia, USA, wants a smart grid, but doesn’t have the means to buy one. GE’s “Grid IQ Solutions as a Service” will build the grid for a monthly fee, and then run it from a giant cloud-computing platform in Atlanta.

GE claims that most utilities don’t want to have to buy the servers that are needed to run today’s smart grid, or to pay the salaries of the people that are need to run it. But GE Energy is headquartered in Atlanta, and has a lot of “computing horsepower” that can be shifted around the region.
Integrating the wind forecast with the maintenance schedule improves the capability of being on the grid when the wind is blowing at its best. This, in turn, results in better wind utilization.

The availability of technical data allows improved deployment and calibration of the turbine, leading to greater efficiency.

Higher reliability through proactive maintenance reduces the extent of unplanned unavailability. There are also examples, like Vestas, where manufacturers offer analytics on some of these data as value adding services to the customers (see fact box).

**CARBON CAPTURE AND STORAGE**
Safe handling of large quantities of carbon dioxide (CO₂) will be a requirement of all Carbon Capture and Storage (CCS) projects.

Development of standards and models will rely heavily on the availability of experimental and operational data. Sensor technology is being developed to detect and monitor CO₂ leakages from the seabed above storage sites. This is especially challenging as there is considerable natural variation in CO₂ levels in the sea. Once the sensor technology is in place, continuous monitoring and risk management of storage sites during injection and through the lifetime of the field (until the CO₂ is absorbed in sediments) will become a prerequisite. Again, such a service will require that advanced sensors are combined with capabilities in big data analytics and domain competence and models.
Turning big data into solutions that add value for the customers is a priority for Vestas.

The key is Vestas’ advanced utilisation of big data. Based on the company’s “Firestorm” super-computer, Vestas constantly gathers data about wind, weather, and real-time performance from almost 25,000 turbines all over the globe. Through the digital platform VestasOnline, these massive amounts of complex information are made available to Vestas’ customers in the form of user-friendly online-applications such as VestasOnline PowerForecast (that provides customers with an overview of the power generation forecast for their wind power plant) and VestasOnline Maintenance (that enables planning of maintenance of the wind turbine at times when power production would be least affected).

“With VestasOnline, we have commercialised our use of big data and converted this part of our business into important knowledge and useful services for our customers. VestasOnline makes it possible for our customers to optimise the revenue generated through better predictability of the power output,” says Christian Frederiksen, Vice President for Global MarCom & Corporate Relations and responsible for Vestas’ eBusiness.

To Morten Albæk, Group Senior Vice President for Global MarCom & Corporate Relations, big data is also vital for Vestas in another way. “In a market characterised by increased competition, digital services can be the decisive factor. Intelligent use of big data enables us to offer our customers some unique services that will improve their investments in wind,” Morten Albæk explains.
Huge quantities of data are collected in healthcare institutions and are increasingly available in digital format. These data are flowing: inside hospitals; between institutions, primary care, and pharmacies; to insurance companies and health authorities. There are huge potential benefits and large business cases to be gained from applying data analytics in healthcare. Important drivers for this development are the need to meet the challenges of increasing costs from the globally aging society and the rising burden of chronic diseases, such as diabetes, heart failure, and lung disease.

Data in healthcare is a form of big data, not only for its sheer volume, but also for its complexity, diversity and timeliness, probably more so than in any of the other industries discussed so far.

The main data pools currently available for processing and analytics typically include: clinical data (patient diagnostics, laboratory data, imaging data, etc.); administrative data (when people come and go, and who is in contact with whom etc.); patient behaviour and sentiment data; activity (claims) and cost data; and pharmaceutical R&D data.

Among the main hurdles in analysing these data are that devices and systems that collect the data and information exist in multiple siloes and do not communicate, that data continues to be regarded as a waste product in many processes, and that standardized data formats and data exchange protocols are lacking. These issues are being addressed, but there is still a long way to go. To further complicate this matter, there are also many legal problems related to data ownership, privacy, and sensitivity of patient data. For the latter, there is considerable variation in the relevant legislation.

Nevertheless, all the big players in IT, in addition to many new ones, are moving into or positioning themselves to deliver analytics for the healthcare industry.

The main areas for analytics seem to be:

- Diagnostics using smart agents and data mining techniques
- Patient monitoring and management (providing the right treatment at the right time)
- Management of patient care processes across system barriers (e.g. from first encounter at GP, through all specialists, lab investigations, hospitals or medicament treatments up until sign off)
- Epidemiology
In 2011, the Watson supercomputer, developed by IBM and named after its legendary CEO Thomas Watson, beat human Jeopardy champions in a TV show. In addition to providing state-of-the-art processing capacity, the revolutionary step is that the machine is able to ‘understand’ both questions and context.

Health has been the first area of commercial application of this technology for IBM.

Watson’s ability to understand questions and context, and to rifle through 200 million pages of data and provide precise responses in just seconds, can help a physician treating a patient to consider all related texts, reference materials, prior cases, and the latest knowledge in journals and medical literature. This can assist physicians in rapidly determining the best options for diagnosis and treatment. Such an approach would also enable more use of evidence-based medicine, which is designed to standardise patient treatments by identifying proven best practices. A simple example of evidence-based medicine in action would be when a health provider automatically places someone who has suffered a heart attack on an aspirin regimen upon leaving the hospital.

**MANAGEMENT OF PATIENT CARE PROCESSES**

Different actors are already entering this arena with offerings based on big data. One example of an entry as a data aggregator is the Premier Healthcare Alliance in the US teaming up with IBM.

In a pilot project with local hospital authorities in Spain, Accenture targets the 4% chronically ill patients that tie up more than 60% of the hospital resources. Through a service that monitors the chronically ill patients that can live at home, anomalies in their condition can be detected at an early stage. When anomalies occur, nurses employed by Accenture visit the patients and, in collaboration with the patient’s doctors, intervene early to adjust medication, make further tests etc.

Avoiding readmission to hospital is a huge business case for the healthcare system and also for the patients in the form of improved life quality.

Monitoring patients’ conditions is made on the basis of medical and lab records, as well as self-reporting of key parameters from instruments available to the patients.
to integrate healthcare data to gain insight, measure performance, and improve population health. This is claimed to be an industry-first data-sharing model among hospitals, doctors and outpatient clinics, supporting accountable care and performance improvements.

On the analytics side there are many actors, including OptumInsight that has made the following statement regarding their capabilities:

“We gather information and apply analytics to make it useful. We create secure, interoperable networks that enable the exchange of information among communities. We understand how those in the health system actually do their work, so information can be seamlessly introduced into the flow of their everyday activities.”

**PATIENT MONITORING AND MANAGEMENT**

In the area of patient monitoring and management, many hospitals have started to apply big data analytics in the care of chronically ill patients. This group represents the largest cost per patient, and large gains can be made by identifying correct and timely treatments and adjustments to avoid unnecessary admissions to hospitals. IBM, through their ‘Smarter Cities’ initiatives, has worked closely with hospitals worldwide to use big data analytics to, for example, support home care of chronically ill patients (e.g. Stavanger Hospital in Norway).

The example of Accenture as a service provider, leveraging its vast analytic capabilities to support Spanish hospitals in a novel way, may serve as an illustrative example of how new service providers may find a place in a novel industry by leveraging their analytic capabilities (see fact box).

**EPIDEMIOLOGY**

In the area of epidemiology, an unexpected actor has created a major impact. Through creative research into their search statistics, Google has found that they can determine how influenza spreads geographically as based on search terms used by individual people (see example box).

A more mainstream example within epidemiology is the Snow Agent developed by the hospital in Tromsø, Norway and the Norwegian Centre for Integrated Care and Telemedicine. The Snow Agent system enables epidemiologists to view spread of diseases in time and spatial dimensions. The Snow Agent system automatically collects anonymous data about incidences of infectious diseases etc., from different health and laboratory systems. A symptom-based surveillance system such as this can help healthcare workers by providing information about the occurrence of a symptom pattern in a patient population at an earlier stage of a disease outbreak.

Also under consideration is providing an app to the general public that can be used to monitor the current occurrence of contagious diseases in the local community.

In terms of patient behaviour and sentiment, it seems that analytic applications can provide information about treatments that are commonly chosen by patients, in a similar way to applications currently available to consumers. Such applications may potentially be as transformative in healthcare as they have been in retail.

There are also more and more examples of innovative solutions, such as Asthmapolis (now Propeller Health) that offers an app where users log their inhaler use and position (using the smartphone GPS). These data make it possible to identify individual, group and population-based trends which, when merged with information about known asthma catalysts (e.g. pollen), can help physicians to develop personalised treatment plans and identify prevention opportunities.

Google Flu Trends is an example of an innovative way in which valuable insights are derived from their own data. Google discovered that an increase in searches concerning flu-like symptoms could be an early indicator of flu outbreaks. The company has been able to change the game regarding disease prevention.
Data use is ultimately a vehicle by which the business model of companies can be changed and improved. In order to comprehend how the use of data can transform business, it is useful to understand and categorize the different business models, then discuss how data can alter the mode of operation in each category. In order to do this, the work of Charles B. Stabell and Øystein Fjelstad of the Norwegian Business School (BI) is very useful. These authors categorize business models into: value chains, value shops, and value networks.

**VALUE CHAIN**
The value chain model fits well when the end product is a physical product; the value offered to customers is the physical item that is for sale and the value chain is the set of processes and assets that are necessary or that support the process of transforming raw material into the final product. For value chains, the main use of data will be to facilitate improvement in the quality of the end product and to make the overall process more cost-effective. Typically this results in an increasingly automated production process. Although this has, of course, been recognised for industrialized processes for a long time, new technology and capabilities have made it possible to take this to new levels of sophistication.

FedEx and Walmart are good examples of this kind of optimisation.

Key aspects in DNV GL’s industries include condition monitoring of production equipment for lower maintenance cost and less downtime, and monitoring of end product for improvement.

**VALUE SHOP**
In contrast, a value shop is a company or business that helps customers in solving problems. The key assets of a value shop are the knowledge and capabilities of its workers; they are competing in the knowledge economy. The value of the services lies not in the effort needed to obtain the solution, but in the value of the right answer to the customer.

The main changes, challenges and opportunities posed by the new data reality to value shops will be related to: the mode of operation, the capabilities expected by the customers, how services are delivered, and changes in the competitive landscape. New capabilities, tools, and infrastructure have the potential to increase the efficiency of project delivery, which, in turn, could return better margins on each project, and from scaling of the services delivered.
Furthermore, customers increasingly expect that value shops are able to handle, assess, and analyse large amounts of data. For both these aspects, the ability to build knowledge into models is key. A shift towards automated data-driven (subscription) services also enables value shops to scale in a completely novel way, but can, at the same time, completely change the competitive landscape when new players enter the market.

Taleris, Sekal and GE (see fact boxes earlier in this document) are good examples of value shops using big data capabilities. Both companies offer services in which large amounts of data are combined with modelling to increase the efficiency and safety of their customers’ operations.

**VALUE NETWORK**

The value propositions from value network companies do not only aim at solving specific problems, but rather provide a network where the customer can conduct their own business, or a network for mediating interactions between customers. The main value lies in the capabilities and size of the network itself. Traditional examples of such business areas include telecom, banks, insurance, and transportation. An easy way to understand that network size matters is to imagine how hard it was to sell the first phone before there was even a single other user in the network. Thus, a very important aspect of this business model is that every customer or node in the network enhances the overall value of the network.

IT technology, the Internet, and the possibility of collecting and analysing large amounts of data have enabled the creation of many such companies. For some of these companies, like Facebook, Twitter, and Google, the product is the use of the users, and there is enormous value in the size of the network. Controlling the data gives power, and being in the spider’s position in the web, with all the data in the network passing through, gives unique opportunities for adding analyses such that the value of the services is enhanced. For example, benchmarks and global views might be provided, and the value of the benchmarks increases as the size of the customer base grows.

Becoming a value network can be highly profitable. One important point is that this can sometimes be achieved by transforming a value chain or value shop into a value network by allowing others to trade, thereby letting data flow on and into the network.
American Airline’s Sabre (yield management system) is a good and early example of this. A more recent example is Apple’s success with iTunes and the AppStore, where Apple has leveraged on its position as a hardware and software vendor to create a platform for trading of digital content (music, films, books, podcasts, etc.) and mobile applications (apps). Apple have successfully bundled these digital outlets to its hardware products (especially the iPod and iPhone) and created a very lucrative eco-system. These are examples where companies have used their existing position to transform themselves into a network provider. Companies like GE and Philips are trying to do similar things in industries where they are in a strong position as hardware vendors (aviation, grid, and healthcare); their aim is to move away from being mere producers and become service providers inside their own networks.

The other option is to invest in building a network that is above critical size. In the latter case it is important to invest early, before the market even exists, and basically to invest in order to create the market. Being a first mover requires having and building trust. Networks can exploit and build upon the fact that customers may not have the analytic capabilities in-house. Carfax (see fact box), eBay, and Marinexplore are good examples of the latter.
Founded in 1995, eBay Inc. is an American multinational consumer-to-consumer Internet corporation with presence in over thirty countries. eBay is now a multi-billion dollar business, with a business model of simply providing a network that connects buyers and sellers of goods and where anyone can trade.

One of the eBay’s main assets, which has also been used to expand the network, is the user and user interaction data that are collected. eBay seller ratings are the foundation of the trust between the buyer and the seller, but can also be used for other purposes such as credit scoring.

Carfax is a US company and system that allows car owners to collect and store the complete history of a car in a central database. Whenever the car visits a workshop, the work done and the condition of the car is uploaded to the system. For the car owner, being part of the network increases the secondhand value of the car. The system has simplified buying used cars and, as the service is available for all brands, circumvention of the brand/vendor barrier adds extra value.

Table 1. Summary of business models with examples.

<table>
<thead>
<tr>
<th>BUSINESS MODEL</th>
<th>IMPACT FROM BIG DATA</th>
<th>EXAMPLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value Chain</td>
<td>■ Big data will enable further optimisation of operating value chains.</td>
<td>FedEx, Walmart</td>
</tr>
<tr>
<td></td>
<td>■ Big data is also used to move towards becoming a value network.</td>
<td></td>
</tr>
<tr>
<td>Value Shop</td>
<td>■ Mode of operation (e.g. automation of knowledge services)</td>
<td>McKinsey, Accenture, DNV GL, Taleris, Sekal, ESRG</td>
</tr>
<tr>
<td></td>
<td>■ Types of problems that customers can be assisted in solving</td>
<td></td>
</tr>
<tr>
<td></td>
<td>■ Optimisation of efficiency and scalability of services through efficient re-use of data and automated services</td>
<td></td>
</tr>
<tr>
<td>Value Network</td>
<td>■ Increase the network value of the network</td>
<td>Google, eBay, FedEx, Carfax, Twitter, Vestas Online, Maersk line, Shipping KPI, Apple, Amazon, Marinexplore</td>
</tr>
<tr>
<td></td>
<td>■ Capitalize on the data generated by the network</td>
<td></td>
</tr>
</tbody>
</table>

33
Operational modes are about how projects and services are executed, packaged, and sold to the market. In the context of big data, a knowledge service provider, operating through a value shop business model, typically helps customers to gain insights from their data through applying expertise and models to data. These insights will, in turn, be used by the customers to improve performance, to manage risks, etc.

But there may be additional services in the portfolio that also generate data about customers from which insights can be generated. These data and insights can be anonymised and then utilized to increase the value for customers by combining them with the customer’s own data, and possibly other external data, in analyses and modelling.

Thus, three possible scenarios or ways of using data to enhance and transform existing services, and to create new data-centric services, can be envisaged.

Data, data feeds, and analytics can be used to support existing services. This implies that customers may not be aware of these internal services and capabilities, but the service provider will use them internally to improve their traditional deliveries. Better capabilities in this domain will also enable scaling through data sharing and reuse of models across projects (see Figure 2).

Figure 2. Three scenarios for combining traditional knowledge offerings/services with data-fed services and business intelligence.
Data feeds and analytics dashboards may augment and become integrated into existing customer-facing services. The customer will be aware of and use some of the data-centric services, but always in the context of an existing service. The data-centric service adds value to the service that has been purchased by the customer. Typical examples include the automated customer recommendations of Amazon, or parcel tracking offered by most delivery firms.

Data and data analytics can also be sold as standalone services. In such cases the customer may purchase just the data or the insights from data independently of any other service. Sales of this kind of service may also result in higher demand for other more traditional services (cross sales). Examples of standalone data-driven services are Vestas’ online power forecast for the wind-power industry and DNV GL’s calculations of emissions to air from shipping traffic, using AIS data in combination with ship data and proprietary emission models. The latter have traditionally been delivered on a project basis, but are now also delivered as a continuous online service.

In this context, a stronger and more complete value proposition to the market may be viewed, as shown in Figure 3, as a composition of three main service areas: expertise, insights, and data.

Figure 3. Combination of different data-driven services; the total package make a much stronger value proposition to customers.
In considering commercialization of projects and services and how they can be packaged and sold to the market, it can be useful to identify a clear position and role from which to act.

The various roles in the data-centric ecosystem are illustrated in Figure 4, and range from being a data provider, data aggregator, data refiner, through being an insight and foresight provider, advisor, to being a full system implementer for customers. Moving up this ladder, both the knowledge required and the value of the delivered services increase. In this context it is important to note that for most of these roles, access to data is essential. Thus, companies that have access to valuable datasets typically use these to climb the value ladder and to transform their business model from value shop to value network.

Hence, assessment of current position in terms of these roles is useful and can be used to sharpen the value propositions from this position and also to map out competition. In preparing a long-term strategy, an appreciation of what is required to take on a new position can be obtained and competition in that position can be mapped.

The ultimate success criterion for an offering based on data and analytics is obviously to prosper in the market, so existing actors need to revise their sales and marketing operation modes as well. New entrants have the possibility of leveraging their flexibility and adopting advantageous new go-to-market strategies.

Typically, the current go-to-market strategy for knowledge companies is selling the experience of an advisor by the hour. An automated data-based offering requires a subscription-based model, with scaling through sales of a solution rather than a project. New, data-based offerings may even cannibalize existing offerings during a transition phase, and hence require strong strategic commitment. Recent changes in the media and publishing industries, as precipitated by the digitization of content, can serve as examples of this.

For a while, newspapers had to offer much of their content free of charge on the internet, cannibalizing on their sales of paper editions. This has severe consequences for their bottom line. New revenue models have recently begun emerging as novel sustainable platforms, in which ‘light’ news is offered for free, but in depth articles is only available as a paid service.

New data-based offerings may also require different pricing models. For example, a service may only be profitable as a value network business model, but in order to achieve a network of critical mass, parts of the service may have to be offered for free. Flickr, for example is adopting this model for image sharing and management, and Spotify is using this approach for music streaming.

Another challenge is related to sales channels; data-based service offerings may require different sales skills and have different target groups. For example, the traditional contact points in a customer organization may not be able to respond to a data-based offering.
### Table 2. Summary of operational modes with examples.

<table>
<thead>
<tr>
<th>Operational mode</th>
<th>Offering</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implementer</td>
<td>Automated decisions from insights and foresights based on data or aggregated data (accountable)</td>
<td>Next maintenance scheduled (Taleris), drilling operation decisions automated (Sekal)</td>
</tr>
<tr>
<td>Advisor</td>
<td>Advice from insights and foresights based on data or aggregated data</td>
<td>Next maintenance advice (ESRG, Rolls Royce, Wärtsilä), 'ask the expert about safety risk' (Lloyd’s List Intelligence)</td>
</tr>
<tr>
<td>Insight/foresight generator</td>
<td>Insights and foresights based on data or aggregated data</td>
<td>Carfax, eBay, Marketing insights based on credit card transactions (Cardlytics)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Grid IQ Solutions as a Service (GE, Opower etc.)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ship ratings (RightShip, Shipping KPI),</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ship fuel quality benchmarking (DNV Petroleum Services)</td>
</tr>
<tr>
<td>Data aggregator</td>
<td>Sharing, selling combined datasets, or combining data for own use</td>
<td>Access to aggregated ship data (IHS, Lloyd’s List Intelligence), Access to aggregated data about the oceans (Marinexplore)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Aggregated bunker price information (Petro Media, Bunkerworld)</td>
</tr>
<tr>
<td>Data generator</td>
<td>Sharing, selling or generating data for own use</td>
<td>Credit card transactions (American Express)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Social media (Twitter)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ship parameters, incidents (IHS Fairplay)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DNV GL Survey findings database</td>
</tr>
</tbody>
</table>

### Table 3. Examples of positions built on data and analytics.

<table>
<thead>
<tr>
<th>Operational mode</th>
<th>Offering</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implementer</td>
<td>Automated decisions from insights and foresights based on data or aggregated data (accountable)</td>
<td>Next maintenance scheduled (Taleris), drilling operation decisions automated (Sekal)</td>
</tr>
<tr>
<td>Advisor</td>
<td>Advice from insights and foresights based on data or aggregated data</td>
<td>Next maintenance advice (ESRG, Rolls Royce, Wärtsilä), 'ask the expert about safety risk' (Lloyd’s List Intelligence)</td>
</tr>
<tr>
<td>Insight/foresight generator</td>
<td>Insights and foresights based on data or aggregated data</td>
<td>Carfax, eBay, Marketing insights based on credit card transactions (Cardlytics)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Grid IQ Solutions as a Service (GE, Opower etc.)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ship ratings (RightShip, Shipping KPI),</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ship fuel quality benchmarking (DNV Petroleum Services)</td>
</tr>
<tr>
<td>Data aggregator</td>
<td>Sharing, selling combined datasets, or combining data for own use</td>
<td>Access to aggregated ship data (IHS, Lloyd’s List Intelligence), Access to aggregated data about the oceans (Marinexplore)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Aggregated bunker price information (Petro Media, Bunkerworld)</td>
</tr>
<tr>
<td>Data generator</td>
<td>Sharing, selling or generating data for own use</td>
<td>Credit card transactions (American Express)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Social media (Twitter)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ship parameters, incidents (IHS Fairplay)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DNV GL Survey findings database</td>
</tr>
</tbody>
</table>
REQUIRED CAPABILITIES

In order to use big data and secure a strong position in the market, a wide set of skills and capabilities are needed. Broadly speaking, these can be divided into three categories: a culture for data and business development; the skills and knowledge to handle and analyse data and to build models using the data; and the availability of tools and infrastructure.

Successful big data solutions are driven by good business questions. Therefore a culture in which data is actively pursued, questions asked, and fact-based answers sought will drive development forward and create new business. The ability to challenge and change current services within the organization can also be a key factor when creating new business opportunities.

Another key factor when building big data applications is the ability to handle and analyse data, to build knowledge into models and to visualise and communicate the insights. In order to address this, skilled workers with the domain knowledge, visualisation and communication skills, the ability to implement (i.e. write code) and build knowledge into models are needed. The scarcity of interdisciplinary talents and the inability of organizations to nurture such talents have been identified as being among the main obstacles for realizing big data opportunities. In this context, we need to emphasise further one more aspect of this process; data cleaning, preparation and formatting. Preparing data for analysis is very labour-intensive, typically accounting for as much as 80% of the time spent in projects. But without it, results simply cannot be trusted. In this process domain knowledge and technical skills are again key, but it is challenging for organizations to recognize and appreciate the effort and complexity involved in this process as it is usually not visible when the end result is presented.

On the more technical side, we need to point out that there is really no “one size fits all” solution for big data; different problems need different tools. Another interesting aspect is that the majority of the necessary software components are available as open source software. There are potentially large benefits to be gained from the freedom and flexibility that comes with the use of free software: adaptation of the software to specific needs; use of open data exchange and storage formats; and in the form of availability of people from the large user-base able to contribute to projects. In addition, large cost savings can be made by those who look beyond the large software vendors. Partners and solutions can be sought in the open source software community and from among the many new specialized data analytics and big data companies that sell and support solutions based on open source software.

We believe that the following capabilities should be prioritized when investing in big data solutions: scalability, flexibility, availability of powerful tools for analytics and visualization, and data governance.

Building for scalability makes it possible to meet variations in service demand in a cost-efficient manner. Scalability is also one of the main selling points for cloud solutions, as it enables organic growth and the possibility of trying out new solutions and ideas without making large investments.

Another key aspect is the availability of powerful tools for analytics, data exploration, data visualization, and model building. In addition to traditional data modelling and handling, the available toolset should support streams of sensor data (time series) and heterogeneous data where the schema is not known upfront. The ability to analyse data on the system without moving it around will be...
very important for performance, and will facilitate reuse of data.

Flexibility is another big issue. The aforementioned availability of tools plays a large role in creating a flexible system, but there are also other aspects. One aspect is the possibility of experimenting and exploring tools, algorithms, and data in a safe environment (sandbox). Another feature is the ease and flexibility by which ideas can be taken from prototype, through pilot, and into production. The ability to try out ideas, and fail fast and cheap is key for testing towards customers and identifying the really lucrative products. Flexibility also has a cultural and organizational side; IT departments are typically set up to handle the day-to-day operation of a company with focus on quality and cost through standardisation, and having this focus does not always promote the level of flexibility required to succeed with big data. On the other hand, experts and data scientists have a tendency to focus on delivering projects and developing new solutions, and are often not concerned with data governance, data re-use, and the day-to-day operation of IT infrastructure. We believe that the ability to balance the required rigidity for operations with the flexibility needed to develop big data solution will make or break many big data efforts.

Last but not least is data governance. When many datasets are stored on the same infrastructure it becomes increasingly important to control ownership and access. The level of access control needed will, of course, vary from case to case. In the most extreme cases, access to data must be controlled down to the individual record (medical records are an example). This will also be a requirement for backup-systems. Good solutions for governance must therefore be a part of the requirements and built into the system from the start.

On a more fundamental level, companies will have to choose whether systems and competence should be built in-house, or whether parts of, or all, the necessary capabilities should be obtained from external partners and vendors.

An important big data trend is the increased use of NoSQL (Not only SQL) database technologies. For big data, these databases are important complements to traditional relational (SQL) databases. These technologies can be important components that offer better support for different data types, flexibility (schema free design), distributed data storage, fault tolerance, performance, and scalability.

Another interesting aspect is that many of the leading big data software solutions have been developed, and made available to the public, by academia and research institutions. A prime example of this is CERN and the Root and Invenio software packages originally developed to handle, organize and make available the massive datasets produced by the Large Hadron Collider.
CONCLUDING REMARKS

There is a strong trend across all industries and sectors to move towards fact-based decisions, and to use data and data-analytics actively to make informed decisions and manage performance. The main enablers behind this development are: greater connectivity, increased availability of data, and improved capability for cost-effective collection and analysis of data. In sum, this is the essence of the “big data” trend.

The new data reality has already had a considerable effect in many industries, and we believe that we have only seen the beginning of this trend, and that it will also have a significant impact within the industries where DNV GL operates. This trend will offer new opportunities and might also challenge and threaten existing companies and services by paving the way for new competitors in the markets in which we operate.

One of the main trends we see in these industries is the increased use of sensor data for performance monitoring, condition monitoring, and optimisation of systems and components.

This trend has opened the door for a new type of service provider, helping their customers by implementing “intelligent operation” of systems or across systems. Key aspects of these services are: performance monitoring and optimisation; condition-based and predictive maintenance; and optimisation across systems and system overview of operations.

In order to meet the challenges, knowledge service providers need to treat data and expert models as an important asset, build a flexible (agile) framework to execute projects efficiently, and nurture a culture for innovation on data. This requires the ability to maintain and develop competent people with the combination of domain knowledge and data analytic skills, and to attract developers who can create user-friendly solutions for customers.

Figure 5. Key components/capabilities for successful big data applications
ACKNOWLEDGEMENTS

The preparation of this position paper has been a joint effort between with resources from DNV GL Strategic Research and Innovation and Global Shared Services IT. An important part of the project mandate for the preparation of this manuscript was to reach out and discuss with external partners, vendors, and competence centres.

To achieve this, we have used our own networks and the more extensive DNV network, including the collaboration with UC Berkeley through DNV GL’s TopTech programme to meet and discuss with leading experts in big data, both in the industry and in academia.

We will especially acknowledge and thank:
Helge Stubberud, Arild Kristensen, Lars Hovind, and Marc Teerlink from IBM for very good discussions and several useful workshops and presentations, including industry relevant examples.

Naci Akkøk, Rune Syverinsen, and Inge Os, from Oracle for an interesting workshop, including a good overview of big data technology and problems related to data governance.

Frode Huse Gjendem and Geir Thomas Knutheim from Accenture for a valuable workshop and a very good presentation, including relevant use cases.

Mike Olson, Clouderas co-founder, Chief Strategy Officer and former CEO, for an interesting discussion about Hadoop, Clouderas business model and relevant use cases from many industries.

Guy Yalif from Twitter joined us for an interesting discussion about Twitter’s data, products and business model.

Espen Andersen, Norwegian Business School, for lengthy and very relevant discussions about how big data can change, create and challenge business models and modes of operation. These discussions formed the basis of our own discussions on business models, modes of operation and taking positions. Espen also provided several relevant examples of companies using big data.

Øystein Widding from the Department of Industrial Economics and Technology Management at the Norwegian University of Science and Technology (NTNU), for discussion around how big data transforms the business of knowledge companies.

Tom Thomsen and Christian Mahesh Hansen from the Information Risk Management group in DNV GL for a presentation about semantic technologies and the deployment of these technologies in projects with customers.

Emilie Udnæs, summer student, for conscientious background research.

Bjørn Haugland, Rune Torhaug, Elling Rishoff, Erik Stensrud, Arnaud LeBreton and Frank Barre Pedersen for valuable feedback and discussions throughout the project.

And last but not least Lucy Robertson for a fantastic effort in improving the language of the manuscript.
Articles, books, and whitepapers


“A different game”, The Economist, Feb 25th 2010


“Building the Foundation for Big Data”, High Performance IT Insight, Accenture

“Casting off the chains”, Øystein D. Fjeldstad and Espen Andersen, EBF issue 14, summer 2003

“Case Study: Business Intelligence Drives Competitive Differentiation at Guy Carpenter”, Rita L. Sallam, Gartner research, July 2009, ID Number: G00169209


“Data, data everywhere”, The Economist, Feb 25th 2010


“Distributed, decoupled, analysed. IT’s next frontier”, Kishore S. Swaminathan, Accenture Outlook, No 3 2011

“Every Business Is a Digital Business”, Accenture Technology Vision 2013


“Industry Analytic Services”, Kurt Schlegel, Gartner research, July 2009, ID Number: G00167258


“The data deluge”, The Economist, Feb 25th 2010


Energy efficiency and safety in shipping


European Commission DG Environment and Climate Action CLIMA B.3/ SER/2012/0014

SST 2013.4.2: Inspection capabilities for enhanced ship safety

Business examples


Amarcon, routing and performance monitoring of vessels: http://www.amarcon.com/

Carfax: http://www.carfax.com

Credit score from social media: https://www.lenddo.com/
Some background:
http://money.cnn.com/2013/08/26/technology/social/facebook-credit-score/

FedEx value chain optimization:
http://www.fedex.com/ma/about/overview/innovation.html

Electronic Power Design, propulsion as a service:
http://www.epdild.com/

ESRG, data analysis and remote monitoring for the maritime sector:
http://www.esrgtech.com/

LightStructures, hull monitoring vendor:
http://www.lightstructures.no

Lloyd's List Intelligence, advisory services in the maritime sector based on public and customer data:
http://www.lloydlistintelligence.com

Marinexplore, data hub and collaboration platform for ocean data:
http://marinexplore.com

Marorka, environmental performance monitoring in the maritime sector:
http://www.marorka.com

OceanIntelligence, commercial risk in the maritime sector:
http://www.oceanintelligence.com/

Petromedia Bunkerworld, bunker price data aggregator:
http://www.bunkerworld.com/

Technology and software

9 Open source big data technologies to watch:


Hadoop, the Apache Hadoop platform for scalable distributed computing:
http://hadoop.apache.org/


OpenStack, Open Source platform for cloud computing:
http://www.openstack.org/


RightShip, safety rating of ships: http://site.rightship.com/

Rolls Royce, analytics for better availability: http://www.rolls-royce.com/about/publications/marine/indepth/16/index.html (page 5/7)

Sabre holdings, booking system in aviation:
http://www.sabre.com/

Sekal, decision support for drilling operations:
http://sekal.com/

SenseAware, Sensors in shipments:
http://www.senseaware.com

Technology companies and vendors

Accenture, analytics services: http://www.accenture.com/us-en/consulting/analytics/Pages/analytics-index.aspx

Amazon Web Services (AWS), the world’s largest public cloud provider:
http://aws.amazon.com/

Cloudera, enterprise big data solutions using Hadoop and big data vendor:
http://www.cloudera.com

IBM, big data services and products:
http://www.ibm.com/software/data/bigdata/

Hortonworks, Hadoop and big data vendor: http://hortonworks.com/

MapR, enterprise Hadoop and NoSQL solutions: http://www.mapr.com

Ontop, semantic technologies: http://ontop.inf.unibz.it/

Oracle, big data technologies and services:

Rackspace, hosting and cloud solutions based on OpenStack:
http://www.rackspace.com/

Revolution analytics, enterprise R and big data analytics:
http://www.revolutionanalytics.com/

Tableau visualization software:
http://www.tableausoftware.com

Teradata, big data vendor: http://www.teradata.com/

Talens, “intelligent operations” for aviation:
http://www.talens.com/
DNV GL
Driven by its purpose of safeguarding life, property and the environment, DNV GL enables organisations to advance the safety and sustainability of their business. DNV GL provides classification and technical assurance along with software and independent expert advisory services to the maritime, oil & gas and energy industries.

It also provides certification services to customers across a wide range of industries. Combining leading technical and operational expertise, risk methodology and in-depth industry knowledge, DNV GL empowers its customers’ decisions and actions with trust and confidence. The company continuously invests in research and collaborative innovation to provide customers and society with operational and technological foresight. DNV GL, whose origins go back to 1864, operates globally in more than 100 countries with its 16,000 professionals dedicated to helping their customers make the world safer, smarter and greener.

DNV GL Strategic Research & Innovation
The objective of strategic research is through new knowledge and services to enable long term innovation and business growth in support of the overall strategy of DNV GL. Such research is carried out in selected areas that are believed to be of particular significance for DNV GL in the future. A Position Paper from DNV GL Strategic Research & Innovation is intended to highlight findings from our research programmes.