# designing future data and information interactions

Current computer science operates with models of information, networking, and databases that were conceived in the mainframe and client server eras and cannot serve the needs of a truly connected world. Data today is locked in silos, on both a system-by-system basis and within applications. Software development is duplicated and applications remain unable to share data and interact fluidly. We are using outmoded data and information architectures for the next generation of Smart Systems and the Internet of Things.

smart systems design | Harbor Research

fathym

The convergence of networked computing and large scale data management with real time machine intelligence is driving the integration of the physical and virtual worlds. The intersection of these trends - the Internet of Things, Data and People - should create unimagined new values. But will it?

The term “convergence” implies unification, but you wouldn’t know it from today’s rapidly evolving “Internet of Things” market—a fragmented landscape full of incomplete platforms, narrow point-solutions, and software incompatibility seemingly all based on the premise of some outsized dependency on big data and the cloud.

What’s required is a distributed, peer-to-peer information management and application development space designed for everyone and every conceivable kind of data – a platform to overcome the Web’s “information islands” and offer the scalability, interoperability and information-freedom needed for the pervasive-computing and Internet of Things era.

Fathym’s platform is designed for a genuinely connected world in which there are no artificial barriers between data types and information. A world that facilitates free flowing data relationships, visualization for data fusion and discovery, and collaborative application development.

Fathym’s approach is not about leveraging aging IT and software technology into a new context; it’s about looking forward to a single, unified platform for interactions to which any PERSON or any THING can contribute, and which liberates data and information interactions by abandoning traditional relational database architectures and the client-server computing model.

# TABLE OF CONTENTS



## ABOUT HARBOR RESEARCH

An internationally recognized research, technology, and business development consulting firm, Harbor Research has predicted, tracked, and driven the development of the Internet of Things since our inception in 1984. While our history is long, our strategy is simple: capture and create value by combining accurate data discovery and analysis with creative systems-thinking. It is this mindset that has given us the privilege of working with some of the greatest companies in the world. Today, we continue to work with C-level executives and top management of some of the world's most consistently successful companies and innovative startups. In the same way that the market has flexed and grown over the years, our services and experience have grown to make us the premier service organization you see today. We work with clients in a variety of ways including consulting, advisory, research and content development, thought leadership and workshop facilitation.

## THE ADVENT OF SMART SYSTEMS AND SERVICES

For quite a few years now, Harbor Research has focused most of its research and consulting on what we call "Smart Systems and Services"—the convergence of pervasive or embedded computing with the packet-switching "network of networks" called the Internet.

Smart Systems—also commonly called "pervasive" or "ubiquitous" computing or "machine-to-machine" or "the internet of things" —usually refers to digital microprocessors and sensors embedded in everyday objects. But even this makes too many assumptions about what the Smart Systems phenomenon will be. Encoded information in physical objects is also smart computing—even without intrinsic computing ability, or, for that matter, without being electronic at all. Seen this way, a printed bar code or DVD disc, a house key, or even the pages of a book can have the status of an "information device" on a network.

But very few people are thinking about smart connected systems on that level. Current IT and telecom technologists are operating with outdated models of data, networking and information management that were conceived in the mainframe and client-server eras and cannot serve the needs of a truly connected world. "Smart Systems" should automatically be understood as "real-time networked information, computation and information interactions," but it isn't. The Internet's most profound potential lies in the integration of smart machines, information systems and people—its ability to connect billions upon billions of smart sensors, devices, and ordinary products into a "digital nervous system" that will smoothly interact with individuals and systems. The nature and behavior of a truly distributed global information system are concerns that have yet to really take center stage—not only in business communities, but in most technology communities, too.

current IT and telecom technologists are operating with outdated models of data, networking and information management

## ENTER FATHYM

In the following pages, we examine a new software business and unique platform developed by a group of people who are thinking on the scale that the Internet of Things and the next wave of Smart Systems needs—Fathym.

The Fathym team understands that the tools we are working with today to make products "smart" on networks were not designed to handle the scope of data types and the scale of diverse interactions. Fathym's innovation, at its core, is aimed at facilitating data fusion which allows not only patterns, but a whole higher order of intelligence to emerge from large collections of ordinary data.

Their platform development represents a true shift in thinking about how devices, people and physical systems will be integrated and how they will interact. Their approach is not about leveraging aging IT technology into a new application or usage context; it's about looking forward to a single, unified data and information architecture to which any PERSON or any DEVICE or any DATA can contribute, and which liberates information interactions by abandoning traditional databases and the client-server computing model.

## THE WEB IS NOT THE INTERNET

Before delving into the new thinking that makes all this possible, let's talk about why it's necessary at all. Don't we already have a vast public information space called the World Wide Web? Didn't the Web completely revolutionize human communication and interactions? And isn't the Web working and scaling quite handsomely? Well, this may not possibly be the case.

IT professionals rarely talk these days about the need for a new information architecture or pervasive information services based on data relationships that can be made available anywhere, anytime, for any kind of information. Instead, they talk about "web services," and now "cloud services." Even high-tech business people use the terms "the Web" and "the Internet" interchangeably without giving it a thought.

But the Web is not the Internet. The Internet itself is a simple, elegant, extensible, scalable, technology-neutral networking system that will do exactly what it was designed to do for the indefinite future. The same cannot be said of the Web, which is essentially an application running on top of the Internet. It is hardly the only possible Internet application, nor is it the most profound one conceivable.

Fathym's platform is a well thought-out distributed information architecture coupled with collaborative application development tools

After more than a decade of rampant, unruly Web proliferation, we see that the Internet's inherent scalability has been both a blessing and a curse. The Internet was able to "give the Web all the rope it wanted," and today the Web finds itself trying to be something that it was never designed to be: the fundamental platform for the future of humanity's information.

This is the dilemma that Fathym's new software tools and architecture proposes to resolve with an extremely well thought-out distributed information architecture and collaborative application development tools. Fathym's architecture and tools are focused on "returning to first principles" by creating a data and information architecture that makes the fewest possible assumptions about devices and data and how they will be used in the future. Fathym's innovation underscores the need to build simple, extensible foundations for intense complexity. They represent a true paradigm-shift in IT thinking.

The Web's Achilles heel does not originate in its browsing software, or markup languages, or the other superficial aspects that most users touch directly. Rather, the Web's weakness lies with its basic enabling technologies—in particular, relational databasing and the client-server model—and the restrictions they place upon structuring, storing, and retrieving data.

The Web stores information in one of two basic ways: utterly unstructured, or far too rigidly structured. The unstructured way gives us typical static Web pages, blog postings, etc., in which the basic unit of information is large, free-form, and lacking any fundamental identity.

The overly structured way involves the use of relational database tables that impose rigid, pre-ordained schemas on stored information. These schemas, designed by database administrators in advance, are not at all agile or easily extensible. Making even trivial changes to these schemas is a cumbersome, expensive process that affects all the data inside them. Just as importantly, they make deep, inflexible assumptions about the meaning and context of the data they store. Both of these approaches to data-structure enforce severe limitations on the two things you want most in a global, pervasive-era information system: scalability and interoperability.

The client-server model underlying the Web greatly compounds the problem. Regardless of data-structure, when you put information on the Web you put it in a specific physical location on a specific server. Thus, a Web URL does not point to information per se; it points to the information's physical location, which is a very different thing. This means that the Web is hardware/machine-centric, not information-centric. Information on the Web is not free to move, and because its life is tied to the life of a physical machine, information on the Web can become extinct. If a particular server goes offline, temporarily or forever, all its URLs are dead and the information is unavailable. You can "mirror" your data so that copies reside on multiple servers, but this does not change the fact that Web-based information is always tethered to specific locations on specific machines.

today's software and approaches to data structure enforce severe limitations on scalability and interoperability

All of this adds up to a huge collection of information-islands called the World Wide Web. Assuming the islands remain in existence reliably, they are still fundamentally incapable of truly interoperating with other information-islands. We can create bridges between them, but islands they remain. That's what they were designed to be.

Multiple succeeding generations of "bridges" once called Web services and today mostly refered to as cloud services are still prey to the client-server rigidity and vulnerability we've described. In order to concatenate information from diverse sources, these services depend upon data-tagging conventions that require yet more layers of prior agreement, schemas, markup, and manual human administration.

With each additional layer of such engineering and administration, the Web comes closer and closer to resembling a fantastically jury-rigged Rube Goldberg contraption. The reason is simple. The Web, with or without "services," was not designed for a world driven by pervasive information flow.

## INFORMATION IS NOT FREE

Information on the Web is not free (and that's free as in "freedom," not free as in "free of charge"). In fact, it's closer to being imprisoned than it is to being genuinely liberated. Thanks to the client-server model, Web information is not free to move anywhere at will. And thanks to our present information architectures, it's not free to merge with other information.

What would truly liberated information be like? It might help to think of the atoms and molecules of the physical world. They have distinct identities, of course, but they are also capable of bonding with other atoms and molecules to create entirely different kinds of matter. Often this bonding requires special circumstances, such as extreme heat or pressure, but not always.

In the world of information, such bonding is called "data fusion." The special circumstances that provoke fusion might be nothing more than a user asking a certain multi-dimensional question. And the newly created matter would be an easily perceivable, manipulatable, or mappable "model" of the answer to that question.

facilitating free flowing data relationships and data fusion is one of the fundamental purposes of Fathym's technology

Facilitating free flowing data relationships and data fusion is one of the fundamental purposes of Fathym's technology. Most people have had little, if any, experience with it, and so its power and importance are often difficult to grasp. The ability to perform true data fusion is the holy grail of information systems because it allows not only patterns but a whole higher order of intelligence to emerge from large collections of ordinary data. The implications for research in almost any area of human inquiry are obviously immense.

Radical new thinking about information technology must begin at the most basic levels, with new conceptions of information and data types and software componentry and of the fundamental units of information itself.

## FATHYM'S BUILDING BLOCKS

Fathym is a fractal-like data "ecosystem" built from a catalog of granular building blocks. Each element acts autonomously but also knows how to co-exist with other data elements and display itself. These building blocks can be combined to form complex parts or even applications. Fathym is future-proofing their innovations in these areas by making the fewest possible assumptions about the nature of networked information objects and the data they carry or process.

Fathym takes a much broader, all-encompassing view of information objects than is typically found in conventional computer science. The information objects are the basic

building blocks of Fathym's data and information architecture. Fathym's schema-free data management approach avoids the confinements and limitations of relational databases. It allows data to maintain their fundamental identity while bonding freely with other data. Furthermore, the attributes of any data object can be changed or extended without changing any other. Fathym imagines an Internet-connected world beyond the Web, and it does so with architectures that embrace every conceivable data and information type, however large or small.

All information objects have inputs and outputs, but little else about them is taken for granted. The essence of an information object is completely abstracted from its real-world embodiment so they are mutually interchangeable. Information objects also transmit information in a significantly different way from most distributed computing systems popular in industry today and, as they scale, encourages and rewards parallel computation, which generally means about the same thing as "pervasive computing."

we need new tools for a genuinely connected world that do not create any artificial barriers between units of information, data or information objects

In the past, the art of parallel computing has involved solving problems fast by getting lots of machines to work on what is essentially the same computation. With Fathym's approach, the art is to get machines to work together productively when their users have several different, independent, but overlapping information needs.

Information needs to be truly portable in both physical and information space, and among any conceivable information devices. As we have seen, chunks of information that lack structure or have too rigid a structure tend to live in isolation, unable to discover and bond with others to create new information. What's required is a data-structure that provides unmistakable identity for every chunk of information but also facilitates flow and fusion.

To this end, Fathym's new information architecture is built upon the simple notion of "nodes" and "links." This way, all data (nodes store heavy data and create structure with attributes) and data-relationships (links that create relationships between nodes and store light attribute data) are free and not trapped in relational database tables. This recognizes that computing systems need a simple, extensible, and universal way to represent data.

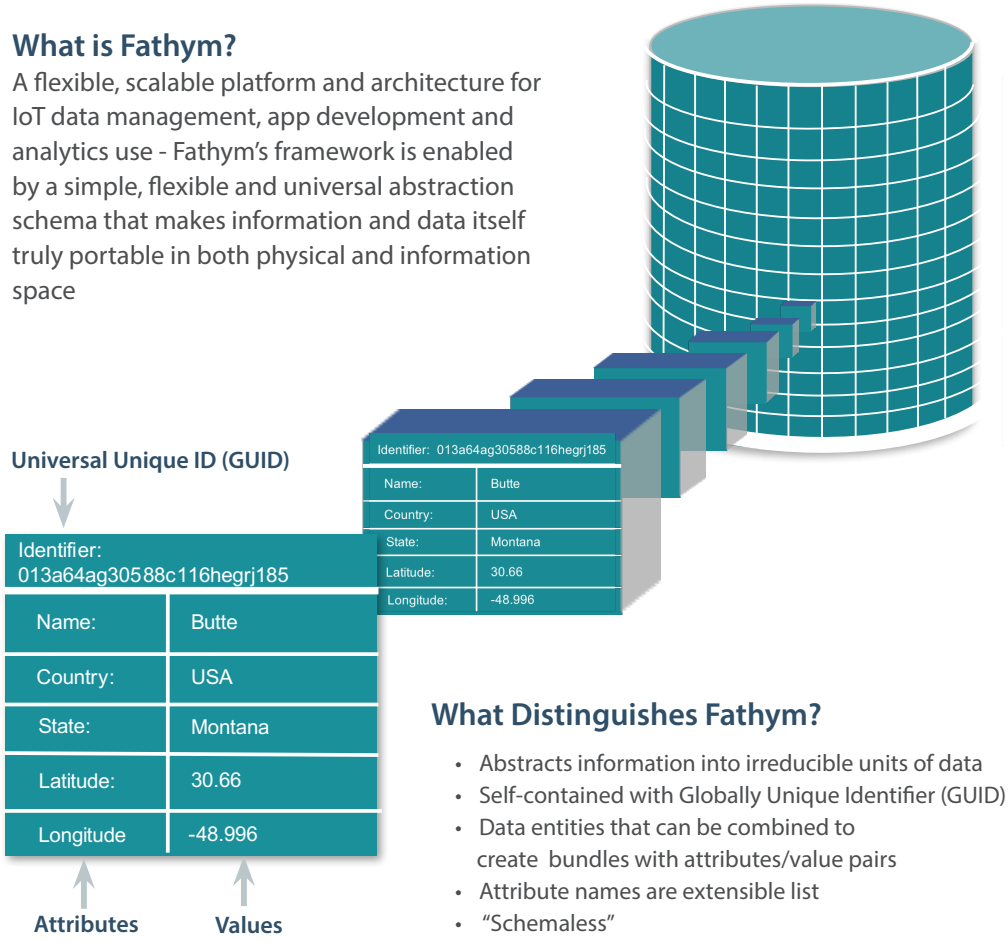
To this simple structure Fathym has made one very important addition: a universal or globally unique identifier (GUID). The GUID travels with the node's data wherever in the world it may go, and in whatever device it might find itself.

Why does the node (a new more universal data container) need a GUID? We all recognize the need for unique identifiers in many aspects of our lives and work. ISBN numbers for books, Social Security numbers for U.S. citizens, license plate numbers for automobiles, and so forth, allow us to refer unambiguously to a single object in some large population of objects.

The aforementioned identifiers, however, are unique only in their own little worlds, not in the world at large. They are not universally unique. They operate within islands of reference, and they do not interoperate with other "islands." They allow us to point unambiguously at an object in a given "namespace," but nowhere else.

## What is Fathym?

A flexible, scalable platform and architecture for IoT data management, app development and analytics use - Fathym's framework is enabled by a simple, flexible and universal abstraction schema that makes information and data itself truly portable in both physical and information space




## What Distinguishes Fathym?

- Abstracts information into irreducible units of data
- Self-contained with Globally Unique Identifier (GUID)
- Data entities that can be combined to create bundles with attributes/value pairs
- Attribute names are extensible list
- "Schemaless"

In an unconnected or minimally connected world (a world of information islands), this works well enough most of the time. But Fathym has designed a system for a genuinely connected world in which there are no artificial barriers between pieces of information (data or information objects). In such a world, there is only one identity space—namely, cyberspace—and thus the entities (information objects) that inhabit that space must be universally unique, not simply locally unique. Thus the absolute need for every data object to have a unique identifier.

Establishing unique identifiers for every piece of information in the world, however tiny or granular, will sound impossible to most people. In fact, technically speaking, it's comparatively easy to generate identifiers that virtually eliminate the possibility of "identity collision" for all time.

## AN ECOSYSTEM OF INFORMATION OBJECTS

Fathym is building a profoundly different information architecture than those in common use today, which one can see manifested in the schemas of relational database tables. Those inflexible, pre-ordained schemas impose more than rigid structure upon information. They impose rigid meaning. Fathym's approach completely abstracts data not only from the devices that may transport or process data, but also from the schemas and layers of semantics that confer specific meaning upon data in specific contexts. They recognize that database designers do not have—nor should they have—any idea of how specific data might be used in the near or distant future.

Instead, Fathym deals with the meaning of data when the data is retrieved for a specific purpose, not when it is stored. In other words, Fathym does not make interpretive assumptions part of the basic unit of information itself.

Doing this allows any datum to move throughout the system as an irreducible entity, with no pre-defined role or absolute meaning imposed on it from birth. The resulting "flow and interactions" allow these units of information to bond with others to form new information relationships that could not have been anticipated by their creators. With Fathym, irreducible units of information are free to play any number of "roles."

A critical part of Fathym's platform is a software toolset for reading, visualizing, and authoring applications. It gives users multi-dimensional power to use and fuse data in many different ways.

## THE FUTURE OF DATA and INFORMATION

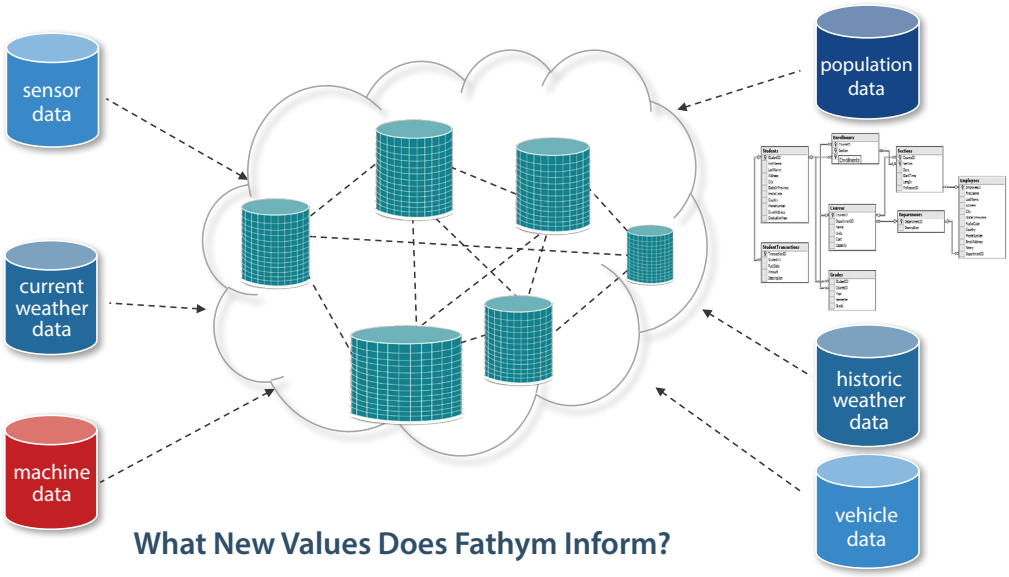
The Internet of Things really means the convergence of information, control and the cyber elements of physical systems. It will require a remarkably agile network that could comfortably scale to billions or more nodes—some of them hardware, some software, some purely data, many of them coming into and out of existence or some in obscure physical locations or changing location constantly. Obviously, such a network cannot be "designed" in any ordinary sense. Certainly, it cannot be designed "top-down."

And yet the Internet of Things must be designed in some sense. Some basic design principles must be put in place to guide the growth of a vast, distributed technological organism that must remain organized as it evolves according to a logic all its own. It demands that we design not only devices and networks but also information itself in ways not addressed by current IT.

Artificial intelligence, machine learning and the Internet of Things are all in some way trying to break from today's computing paradigms to enable real-world physical systems. These promising technologies need to be viewed as applications that will work with the next generation IoT architecture providing access to data of all types. Said another way, it will not be feasible to re-invent the underlying infrastructure for each new application. Instead, these applications need to integrate easily with an appropriate networking and data management foundation.

## What is Fathym's Impact on IoT Development?

The inexorable movement toward pervasive, distributed computing ("the network is the computer") implies the existence of an equally pervasive, distributed store of knowledge that finally does away with all theoretical barriers between the world's units of information




## What New Values Does Fathym Inform?

- Creates distributed, shared information spaces
- Allows participants to reuse data with ease
- Increases efficiency, decreases complexity of data formatting
- Increases speed of discovery of new analysis techniques
- Improves ability to collaborate across domains and business units

We have seen that we can make a computer capable of beating the reigning genius of chess, but yet we can't make a robot capable of walking across the street as well as any normal two-year-old child. The real world is not a strictly regulated, closed system like a chess game. Sensing a player's moves on a wired chessboard and responding quickly and intelligently based on "knowable" algorithms is one thing. Sensing and responding to

physical systems and states – is a fundamentally different challenge. To achieve the seamless intelligent solutions envisioned by the IoT requires an entirely new approach that can leverage common platforms and data models across diverse devices, data and domains.

The fact that a wide range of sensors, machines and equipment can transmit information about status, performance and usage, and can interact with people and other devices anywhere in real time points to the increasingly complex role of data in IoT systems. This only compounds when we consider the billions or more of networked devices that many observers are forecasting will be deployed and the scale of data they will produce.

Fathym's platform is a great example of a "catalytic" innovator that is accelerating the development of new data management, analytics and application development tools for smart systems, services and the internet of things

The tools we are working with today to make products "smart" on networks were not designed to handle the scope of new functional capabilities, the diversity of devices and the massive volume of data-points generated from device interactions.

These challenges are diluting the ability of organizations to efficiently and effectively leverage the real value of connectivity. The rigid and fragmented nature of software offerings available today make it extremely difficult and expensive to use and re-use diverse device-generated data.

Customers expect evolving software tools to be functional, ubiquitous, easy-to-use and able to be added to their smart device systems even as they come from different vendors. What would be entailed to achieve these requirements?

» An architecture that can align with and "form fit" to the physical world and its embedded and distributed computing architecture

» Software tools that are easy to use

» An architecture that addresses a broad range of real-time, and historical data analytics requirements – ideally a single unified framework to design and build solutions that can interoperate across diverse data environments and under widely differing usage scenarios

» A true distributed software architecture that can process and create value from device data locally, while enabling higher-level applications that generate value from enterprise or portfolio-wide data

» Software that can enable easy integration of any combination of inputs and data types - message, feed, and streaming data - in real-time that's independent from traditional rigid database technologies

» Software designed without the bias and dependencies of a single product or service application that will cause integration pain when attempted to be used beyond its original scope

» Truly scaleable software architecture, data models, tools and functions with no constraints on where it gets deployed – chip, device, server, cloud, or hybrid system- and provides extensibility with common features and functions at each level of the architecture

» Software and tools that allow users to easily and quickly build their own applications and automate visualizations that present critical insight and value to the user

We are reaching a critical juncture in market development where organizations will soon be crying out for a completely new approach - one where the effort invested to develop new data-centric IoT applications can be quickly and easily re-used again and again across an ever broader spectrum of devices, integrations and end-use applications.

## FATHYM'S DISTINCTION

Harbor has seen over and over that the biggest winners are those players that take firm control of their core innovations and growth strategy and shape their competitive arenas. These are the companies we call "catalytic." Just as a chemical catalyst hastens the rate of a chemical reaction, companies with catalytic strategies shape their worlds at rates that take the competition's breath away.

In chemistry, a catalyst is an agent that speeds the chemical reaction that produces a desired compound. Similarly, in business a catalytic strategy is one that hastens the arrival of a desired end result or state. Those results involve a "value compound," a unique combination of technological and business system elements which, when offered to customers, result in accelerated market penetration, value chain advantage, or advantageous market structure.

Of course, simply "rowing harder" may hasten the arrival of this state somewhat. But in business as in chemistry, catalysis is in essence the process of sneaking around a barrier that others are struggling to climb, and thereby arriving at the destination much more quickly. In other words, it is following an entirely new path. The catalyst itself may be a technology or other business system element that creates the shortcut to an advantageous structuring of one's competitive arena.

In our view, Fathym's platform is a great example of a "catalytic" innovator that is accelerating the development of new data management, analytics and application development tools for Smart Systems, Services and the Internet of Things.