Why an Ontology Engine Drives the PointCross Orchestra Platform

table of contents

by Suresh Madhavan, Ph.D.
CEO, PointCross Inc.

Why an Ontology Engine?

This white paper is a response to frequent requests recently from customers who are curious about our choice of dynamic ontology to drive the entire data representation within the unique software architecture of our Orchestra platform instead of the traditional monolithic data model based software solutions in the industry.

Orchestra is a horizontal platform for building, presenting, and encouraging the re-use of institutional memory and knowledge for business processes, collaborative discovery and decision making, as well as putting all structured data and unstructured content into structures of business contexts. All basic applications, or integration to desktop applications (Outlook, Office) for emailing; setting up meetings and capturing notes; creating, version controlling, and publishing documents; and accessing data are intrinsic to Orchestra. A context based authorization model drives search, information security, and compliance within the platform. A set of tools (FDK – Funnel Development Kit) and over 1,200 APIs and web services allow complex solutions and business processes to be built on top of Orchestra. These solutions orchestrate disparate processes across organizational boundaries including tacit interactions with external stakeholders.

Orchestra has a very different architecture from the typical enterprise software platform or solution. It is rooted more in concepts from neuroscience than computer science. Orchestra creates the closest representation of institutional memory in today's market. The analytics that are being continually added and the constructive growth of the ontology with use at each enterprise, Orchestra and solutions built on it will enable it to begin representing the glimmers of institutional intelligence. This is not an experiment or yet another attempt at artificial intelligence or expert system; this is about building a fundamentally robust contextual representation of knowledge; automating processes to use it; and then creating the means to learn from it.

To understand why an ontology engine is a rational choice for such a platform as Orchestra, it is important to first understand the limits and restrictions that the traditional data base model rooted architecture pose. This paper takes a constructive approach by building on the basics and attempting, perhaps not well enough, to develop the picture from first principles.

MONOLITHIC DATA MODEL BASED ENTERPRISE SOLUTIONS

For a generation of software engineering, nearing three decades now, the fundamental approach to enterprise solutions architecture and development has remained about the same. In fact enterprise software services has become a heavily formalized industry, incented by a world of T&M projects and process change management practices, has failed in delivering value for money in less time. Compared to three decades ago, computer, storage, and network hardware cost a fraction today while delivering orders of magnitude higher performance. Business software, on the other hand, costs much more for the same functionality today. Why is it so much more difficult and why are there so many reports of poorly executed software projects and dissatisfied business users? Why are more than 50% of all software development projects deemed complete failures and abandoned, while close to 100% are seriously over budget and late? Why do business users, who are able to find almost anything on the internet at home, spend 75% of their time searching for knowledge at work? Why do companies have to maintain ACL (Access Control Lists) for each of their enterprise applications? Why are software updates so expensive and time consuming? In many ways the traditional n-Tier software architecture and its monolithic data model that is managed in a relational data base engine is the cause of the problem.

Enterprise business solutions and business process management solutions are built on an “N” tier – usually a 3 tier – architecture consisting of a data tier, an application layer where all the business logic and rules are programmed, and the presentation layer which, depending on if it is a web application or a client server application, works to make the user interfaces and controls available to the user.

The data tier which resides within the database itself consists of a monolithic data model of tables and indexes that relate one table to another. This data model is absolutely important to the solution and its performance – it determines how efficiently data can be stored and retrieved and how related data can be found by navigating logically using a standard database query language. When building a business solution, a number of solution architects and process consultants work with the key stakeholders of the process and translate their real world business need into a process map. Techniques such as UML, ARIES, ORM and many others exist to help in formally defining the business process that must be modeled. These solution architects then work with a number of DBAs (Data Base Architects) who then design a data model consisting of a system of data tables. Each table carries one type of data (whether it is text, integers or floating point numbers and such) and key fields are indexed to other fields and tables. The relationship among these tables is defined and indexed. The space allocation for these tables is defined. A considerable amount of time and effort is expended in developing this model which can be seen diagrammatically through an ER (Entity Relationship) diagram.

The data model defines the business process that will be delivered. As business changes the way it operates; changes its governance; establishes new business rules (new attributes, or variables, must be defined and used); and the data model must be re-designed. These data models are far removed from the original business problem or even the language of the business stakeholders. There is nothing wrong with this time tested approach. This approach is still ideal for business processes and solutions that are relatively static because they are not likely to change very much. For example, accounting applications like the general ledger must follow certain GAAP standards; budgeting and financial packages are expected to remain constant year after year; point of sale and transactional systems are likewise fairly consistent. These are all relatively stable processes and the challenge is more in how the data is stored in databases so that application are responsive or stored in warehouses so that it can easily be retrieved for such things as business analytics. The application layers are not very complex – their logic is well codified, and the logic does not evolve or have to adapt to the organization or its stakeholders’ decisions.

Some industries have most of the applications they need for their core line of business using this traditional architecture – retail, manufacturing, trading, service industries, and utilities are all example of such transaction-heavy industries. In database terms these businesses are modeled by data tables that are row heavy (Wal-Mart, the giant among retailers, for example, handles as many as 20 million transactions per second – each one forming a row in one of many tables)!

Most enterprise applications beyond ERP are also based on an underlying monolithic data model. Examples include document, records, content management systems, collaboration solutions, plant design applications, business process management solutions, engineering applications, warehouse applications, and the analytical engines that operate on these applications.

ORGANIZING INFORMATION

Going beyond the data model, let us consider how information is managed in enterprises. Business and enterprise applications typically manage the data and information content that their specific application needs or creates. Products of collaboration and tacit interaction which include documents, meeting notes, and, most importantly, emails are usually managed by their designated application which in turn organizes them into folders that must be either centrally managed as in the case of a DMS or CMS, or managed locally by each user in their machine. For years the story has been that individuals need to manage their own data store and so in a typical organization all the content that is moving about ends up in users' Explorer folders or in their email client's folders. Books are kept on shelves, organized by theme, title or author. Mail is put in boxes. And so the software industry kept that "paradigm" and tried to replicate it with cabinets and folders.

Why it is that applications need to be coupled with their data? Why should information be organized into folders? Why is it that people still need to look for content based on matching text string rather than topics that relate to their business? The software industry and enterprise architects rarely have this conversation and it has never been clear why companies still rush out and buy not one, but as many as half a dozen content and document management systems – each time arguing that they will solve one or the other business problem.

It does beg the question, "how do I find what you put away," but no one seems to be asking that question except for those who are struggling with the question, "What don't I know that my company knows?" Even people who set up their folders do not necessarily remember the path to the content they are looking for – why expect someone else to find it? Lynn Chou, ITC's general manager of Global Technology and Strategy, wrote in 2007 that, "11 different document and content management systems are used within Chevron to manage an estimated 283 million online documents. The average number of systems used per employee is 3.5. And that doesn't even include email or those documents saved on the C drive of individual computers. "Approximately 45 million emails are sent and/or received each month, approximately 850 emails per person per month. With that volume of content, just how difficult is it to find what we're looking for? According to the American Records Management Association, the average user spends the equivalent of 1.5 to 3 workdays a month searching for documents. With the size of Chevron's workforce, that's equivalent to losing the productivity of 7,000 employees per year."

It has been said that madness is doing the same thing over and over again and expecting different results. But this is, apparently, what corporations are doing.

The real need for an organization, or even a team, is to be able to instantly re-align all their available content into taxonomies of their choice – a taxonomy that is relevant to their specific needs at the moment and through which they can intuitively navigate. For a learning organization it is also important that these taxonomies reflect not just their local and global, accepted or controlled, definitions; they should also include the current state of all their business activities including projects, processes, and collaborative activities.

TRANSACTION HEAVY BUSINESS PROCESSES

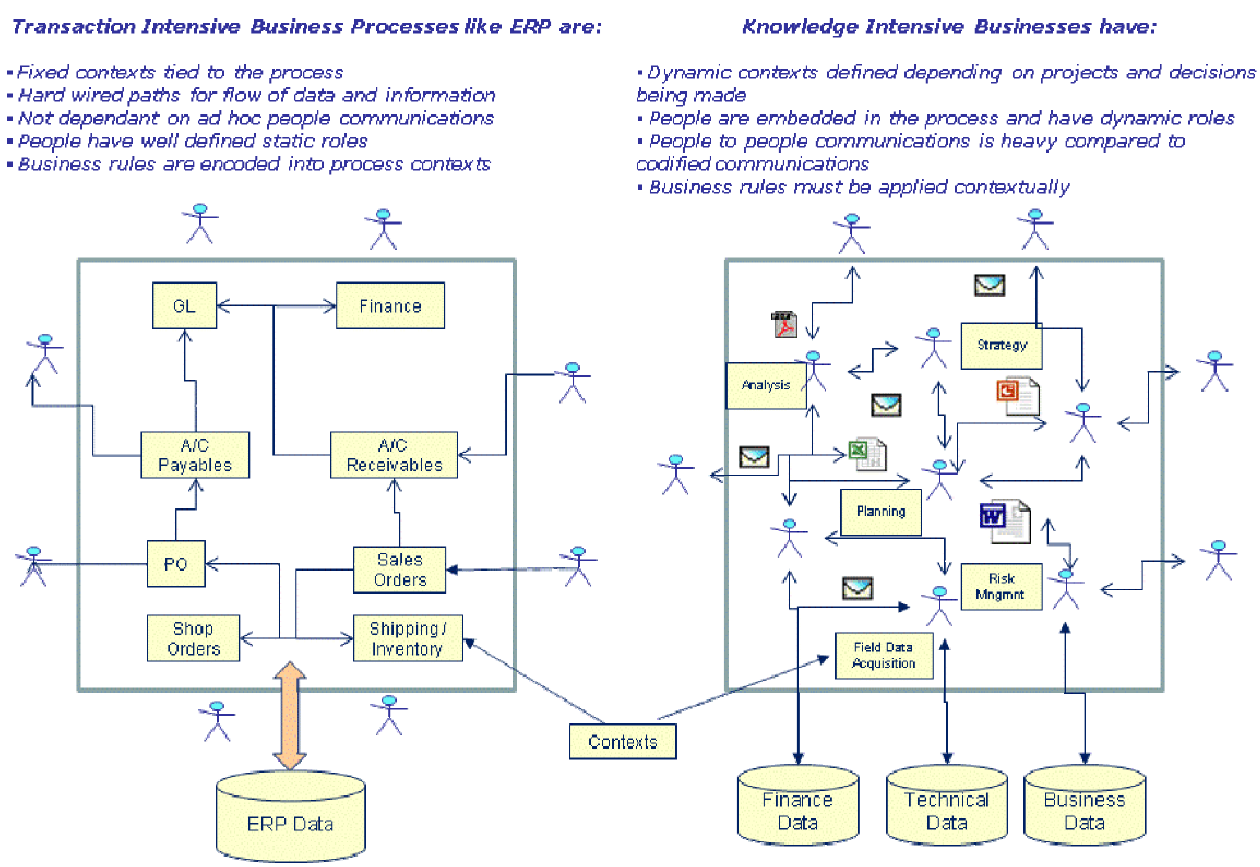
Now let us consider the nature of business decisions and communications in organizations. The nature of decisions in these transaction-oriented industries is highly codified and based on data. For example, a purchasing agent is typically presented with the current inventory, on-order quantities, and the short term and medium term demand for those items under that agent's control. All the possible vendors, their past history on delivery performance, their standard pricing and discount schedules and other performance parameters are clearly presented on a dashboard to this agent. The agent has sufficient leeway in making a decision on how much to buy, from whom to place the order, and to negotiate a suitable price with the selected vendors. If the current inventory is not known the agent would not be able to decide how much to buy. There is very little "judgment" needed in making this purchasing decision. These are data centric decision processes. If you have the data you can make the decision. If you don't, you wait until the data is there.

Consider Wal-Mart. Their complex business intelligence analytics software might make assessments by the second and their shipping staff and inventory staff might find their screens giving them data (in fact they are being instructed) about how many umbrellas will suddenly be in demand in Seattle in the next 48 to 72 hours because of predicted rain showers while New York will need a lot of beachwear because of the upcoming long holidays and the anticipated hot summer weather over the same period. So the New York store shipping agent might get an instruction from the ERP system to ship his umbrellas to Seattle while Seattle ships out most of their beachwear to the New York stores.

These are data-centric industries: their stakeholders do not need to send emails or make phone calls to discuss the data they possess or the decisions they contemplate making. The left side of the diagram below shows life in a data-centric, transaction-oriented business. Stakeholders in each business function are connected by the application layer based on the business logic and the roles of the people. Ad hoc communication is neither necessary nor encouraged because the ERP or other system drives the displays for each stakeholder with all the data they need at the right time; and elicits their decisions so that the appropriate effect is shown to the other stakeholders (e.g., a purchase decision of a widget causes the accounting system to be updated, and the inventory is updated to show the new widgets on order).

Now, this is not how strategic decisions are made in these industries by their CEO, CFO, or chief marketing people. These executives decide the direction of the company: should they increase their inventory to hedge against a rising Chinese Renminbi against USD? Should they invest in building stores in Europe? Should they call back pet food of a certain manufacturer on the basis of a news report that some batches were found to be tainted with melamine? Should they cut back on their cap-ex spending on new store facilities based on the latest consumer confidence index? Clearly the data they have, whether external or derived from business intelligence analytics from their transactions, is insufficient to make these decisions, but decide they must. That's why they get paid the big bucks – to make good judgment calls.

There are entire industries where the bulk of the processes and decision making has elements of strategic thinking. A geologist and geophysicist who interpret a wide range of data put



themselves through considerable dilemma before they decide whether to drill; where to drill; and how to drill. Their decision commits millions of dollars and the possibility of a dry well at the end with no new knowledge gained. Or, they might find oil and reap big rewards.

NATURE OF STRATEGIC DECISION MAKING

What is special about the kind of decisions these people make? For one, they make decisions based on a collection of situations, or scenarios, that they presume; which in turn they create based on their collective judgment; which in turn is affected by the data they have; and the prediction of what that data and their trends suggest to them.

Secondly, the kind of decisions they make tend to be "strategic," meaning that they define "what" must be done and the direction of the follow-on planning and operational decisions that must be made to effect these strategic decisions. For this layer of management, ERP is an information environment that is necessary but not sufficient for planning or decision making.

Stacked on top of each other, the layers of decision makers in a business are its operations, guided by planning, and directed by strategy. Operational decisions are largely based on data. Operations also collect transactional data as part of their activity. Planning uses information - that is structured data, as well as unstructured data in the form of text within emails, documents, meeting notes, and analytical output from BI tools. The strategic layer of management typically relies on knowledge to make decisions.

Knowledge is information put in context.

Contexts are the virtual spaces where people with shared purpose come, work collaboratively, and communicate; while related information is continually anchored forming a layer of rich metadata and links to all kinds of other sources of data and information. These contexts have relevance to the line of business and its strategic direction: a facility that must be built; a product that must be withdrawn from the shelves; a new market to be entered.

Knowledge is the confluence of intelligent, skilled, experienced people working with disparate information that has been contextualized. The information environment for the strategic layer must support tacit interactions as opposed to purely codified transaction notifications as in ERP. Since the risks of decisions are high as are the rewards, these decisions and the information they rely on must also be audit traceable and be compliant to various policies or regulations.

Compared with industries such as retail, manufacturing, trading, refining, or utilities, there are a number of other industries whose core line of business is dependent on knowledge and decisions based on the judgment of people who are knowledgeable, skilled, and experienced. These knowledge industries include upstream exploration and production of oil and gas, Pharmaceuticals R&D, investment banking and insurance, innovative R&D of high technology or defense, and many others.

These industries are characteristically:

- ☑ Knowledge-intensive;
- ☑ Decision centric - in the sense that their core line of business proposes strategic direction and decisions;
- ☑ High risk, but also high rewards;
- ☑ Decisions are made with imperfect data and information.

Another way to look at it is that the total revenue in the traditional transaction heavy industries comes mainly from its operations, efficient planning and execution to deliver their goods and services. On the other hand knowledge intensive industries return huge rewards on large investments into ventures such as finding and developing a new drug; finding and exploiting oil or gas reserves; or inventing a new technology; or crafting a new deal or security. In the first case, knowledge serves its strategy; while in the latter their strategy is to apply their knowledge to make money.

NEEDS OF KNOWLEDGE INTENSIVE INDUSTRIES

Knowledge is disparate but related data and information put into contexts for the purpose of supporting decisions or learning.

In 1959, Peter Drucker coined the phrase "Knowledge Worker," which Wikipedia: defines as: "an individual that is valued for their ability to interpret information within a specific subject area. They will often advance the overall understanding of that subject through focused analysis, design and/or development. They use research skills to define problems and to identify alternatives. Fueled by their expertise and insight, they work to solve those problems, in an effort to influence company decisions, priorities and strategies." McKinsey has described the nature of the modern knowledge based industry as based on "tacit interactions."

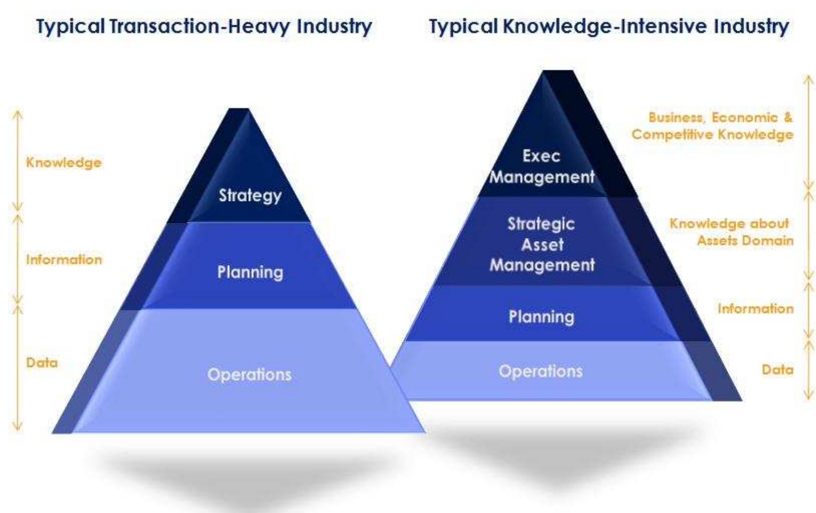
For a knowledge worker, data is an essential commodity that should be accessible, but it is not sufficient to make decisions that affect their assets. Information, or structured data, is essential and useful. But knowledge, which is contextual information, contains the insights and nuggets that drive them to new heights of innovation, invention, and concepts that they share and collaborate with their colleagues. Scientists in Pharma R&D, or explorationists, geologists, geophysicists, reservoir engineers in the upstream oil and gas industry seek, use, and re-use their personal, team, departmental or enterprise level knowledge.

Disparate facets of what knowledge (or knowledge management – which suggests that we already understand what knowledge is and the only thing left is to "manage" it) is or should be in an enterprise occasionally glints seductively like diamonds catching the light just so. There is something there, but it is elusive. For example, taxonomies are important. Search is important. Social networks are important. Semantic web, which in effect marries both, is important. Web 2.0 offers some important possibilities. Collaboration and shared sites play a role in – well, sharing. Document and content management were big earlier this decade, but mostly because, before search, people thought of content as they do physical books and documents that need to be shelved properly so they can be found; a habit of thought shackled to the past. SharePoint, the fastest growing collaboration application in enterprises and potentially the biggest threat to Web 2.0 players, is mostly used for file sharing – which the reason that document management systems exist. The problem as we see it is that the industry and customer organizations are stuck at thinking of knowledge as information, information as content, and memory as finding specific content. The success of search in the consumer world seems to have caused us to stop in our tracks and look for "stuff" rather than discover past business experiences.

ENTERPRISE SEARCH

In the diagram shown below, typical consumer searches fall in the 3rd quadrant, as would an enterprise user who goes outside the company firewall to look for information. Searches within one's desktop fall in the 1st quadrant, while searches within the firewall of an enterprise would fall in the 2nd quadrant. This last scenario continues to be a major challenge because unlike a user who searches within their machine, users cannot be shown everything within an enterprise, such as private email written by colleagues or sensitive documents to which the user does not have authorization.

Implementations of enterprise search technologies that fulfill the needs of the 2^{nd} quadrant may show some success in the pilot phase when there are a limited number of documents and content. When realistic quantities of content are injected, however, the Access Control Lists (ACLs) apply huge burdens on



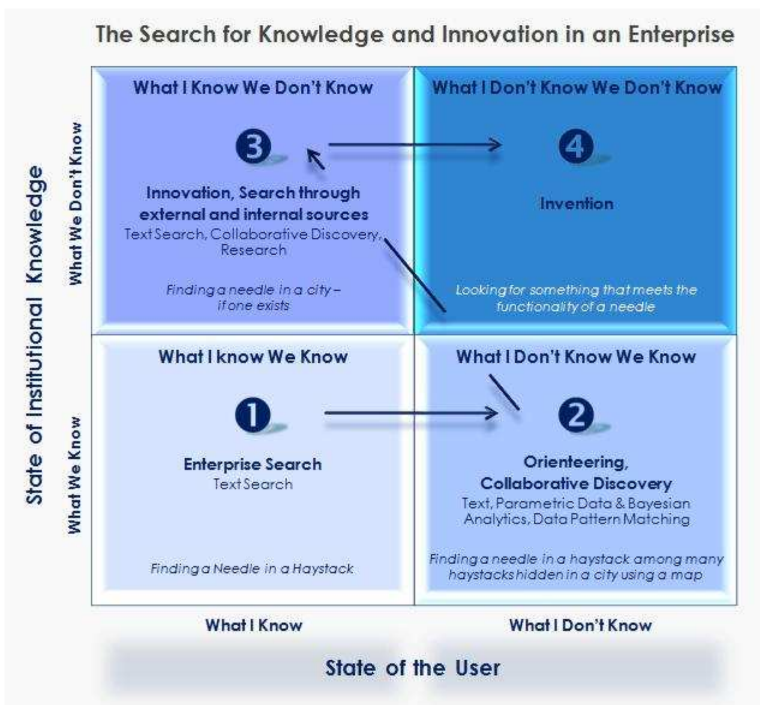
the post-processors or even the indexing and search phases of the search lifecycle because each of the tens of millions of content objects must have an inclusion or exclusion list for authorization purposes. The problem is not the search engines, which have marvelous technologies behind them. The problem is that organizational changes cause dramatic changes in the list of authorized people and each time this happens the ACLs have to be updated and the indexing has to be re-done.

Quadrant 3 represents the search of things that a user doesn't know that the institution does know. If an enterprise wants to put its collective institutional knowledge to real use it must make it discoverable – not just the content but the institutional experience itself. Neuroscience holds many of the answers to the question of how we as individuals or teams create memory, contextualize information, and use pattern recognition and rapid simulated projections to recollect contexts, moods, scenes and higher levels of consciousness.

Consider how we constantly see, touch, hear and smell our environment. These individual snapshots and their sequence are captured and stored in the brain all the time. Without attempting to get into the science of neuroscience, the essential point is that the brain builds contexts from these snapshots – contexts that represent the time, setting, smells, emotions, etc., of the memory. Images that are not contextualized are removed periodically, as far as neuroscientists understand, such as during sleep. Contexts are linked to other contexts; and higher-level contexts are created. As new events happen these contexts get linked or re-wired continuously. When we recollect, we hardly ever recollect just an image, or just a smell. We recollect the experience holistically as a collection of stored sensations around a context; perhaps time was the context, or perhaps it was a thought that came up at that moment.

What is also interesting is that these collections of disparate sensory images can be, and usually are, related to multiple contexts. For example, imagine being invited to a friend's new house on the beach for dinner. Hanging out in the kitchen with the host and a few good their other friends, one face in particular catches your eye. The scene is beautiful with the setting sun and the waves crashing on the beach; the smell of cooking wafting in the air; and this special face who catches your eye. Months and years later someone somewhere is cooking with similar ingredients and smells, and it triggers a memory of that experience, making you wonder whatever happened to that person who caught your eye. Or, you enter a subway and as the doors close you catch a glimpse of someone who just got off the train and something about the face, or a mannerism, catches your eye and – snap – you are back in that friend's house on that evening with the beach and the cooking – and you get hungry for that food. Any one of the disparate sensory images has the potential, and only a possibility, of triggering a memory of that event.

As a thought experiment, consider what happens if, as a voracious reader of fictional stories, you experience a whole range of situations and mental images – possibly very vivid if they are good stories. Later, your real life experiences could trigger these fictional memories just as well as real memories. This transference is an example of the potential of institutional memory: events experienced by others that may create disparate content (similar to the images) can cause shared experiences. As real-time, hi-fidelity simulations get better with technology, it is becoming common to see people having a hard time separating reality from virtual reality. We have seen fighter pilot trainees drenched in sweat and experiencing all the symptoms they would have in a real flight or mission emergency – except they were in a simulator. The long term objective is to provide knowledge workers with a realistic support environment in which they can sense their business



reality, project forward, and make better decisions – like the pilot who gets trained on simulators. Like the simulator, a knowledge system should become a virtual institutional information and decision-making environment.

Humans don't just remember discrete sensory images. We also sense and experience or, in frequent modern parlance, we "have a moment." These moments are collections of images married to emotional sensations that create a local context. In the mind these are not necessarily transitory contexts; rather they are a series of contexts in time, or even conditional on other contexts happening. When triggering events evoke one of these time sliced experiences or contexts, and if by chance the present time real time (or virtual if it is a movie or a book that one is reading) imagery transitions into one that has a good match with the next one stored in the memory – guess what – "its déjà vu all over again" [Yogi Berra].

Creating such contexts is the secret to creating institutional memory. Ontologies are an essential part of this effort.

Another example: in an effort to push tactical decisions "downward" to the field, military tacticians use these techniques to create what they call CROP (a Common Related Object Perspective) so that people have true situational awareness or immersion in the situation. The fusion of human sensory inputs with a combination of real-time, real-world and synthetic imagery is surreptitiously entering our reality. This could work for enterprises, too.

Contexts are the building blocks of institutional memory. An ontology, which consists of contexts, their relationships, and disparate content, is the institutional memory. Real-time analytics and processes laid on top of this ontology create an active, dynamic ontology that has the potential to serve organizations in real time.

Perhaps Godel's Incompleteness Theorem limits us from ever completely mimicking the miracle of the human brain. But that shouldn't stop us from emulating it as far as we can. We can at least attempt to create a working information environment that is at first passive, and later active, serving teams and the institution at large with the equivalent of true shared memory of business experiences and basic cognitive functions needed by the institution. A solid information environment will therefore include storing content into folders and searching for text buried in content. Shared spaces where people exchange text content have been associated with knowledge activities and yet they appear naïve, pedestrian, and a pejorative attempt at simulating the human kind of knowledge. Knowledge as applied by people in real life is an unscripted orchestra of ideas, concepts, cognition, reasoning, perception, pattern recognition building and testing hypotheses, synthesis followed by analysis, and more observations and communications – all for the purpose of making decisions. So how exactly does a man-made system claim to be a true knowledge management system, let alone an agent that can support knowledge- related human activities? We are a far way from the human brain, but let's consider a few baby steps that will attempt to mimic nature.

Transforming Data and Information into Knowledge

The practical uses of knowledge in business are twofold: to decide or to train. Decisions include research decisions, judgments made in support of decisions, search, orienteering, planning, discovery, exploration, commitments, and inventions – the creative use of knowledge. Traininginvolves and overlaps with education, learning, discovery, search, orienteering, research, and exploration – the transfer and imparting of knowledge. Since we are thinking about organizational knowledge and not just individual knowledge, an information environment that supports knowledge work must also include the tacit interactions among people and consider the needs for sharing information contextually.

The business context

By context, we mean the virtual space that represents a commonly useful concept, an idea, a shared purpose around which people who have a need to know or contribute can work to achieve that purpose while using relevant data and information. Contexts may represent business topics; matters where intellectual work and contemplated decisions are made; subjects that require collaboration among people who have a shared business purpose; and disparate types of information that must be securely accessed in order to arrive at decisions. These topics can represent any business artifact that persists during a process or whose information must be recorded for compliance. Examples of contexts include:

- ☑ Business events for which people work;

- ☑ Project or project elements;
- ☑ Deals such as acquisition, divestiture or swaps; in/out licensing;
- ☑ Assets and their components;
- ☑ Plants and their subsystems;
- ☑ Functional organizations and their departments;
- ☑ Roles – of people and their responsibilities;
- ☑ People – and their expertise.

Contexts are organized into structures, and contexts in one structure may be related to contexts in other structures. Contexts are meta-tagged by links to controlled and loosely-controlled taxonomies carrying definitions.

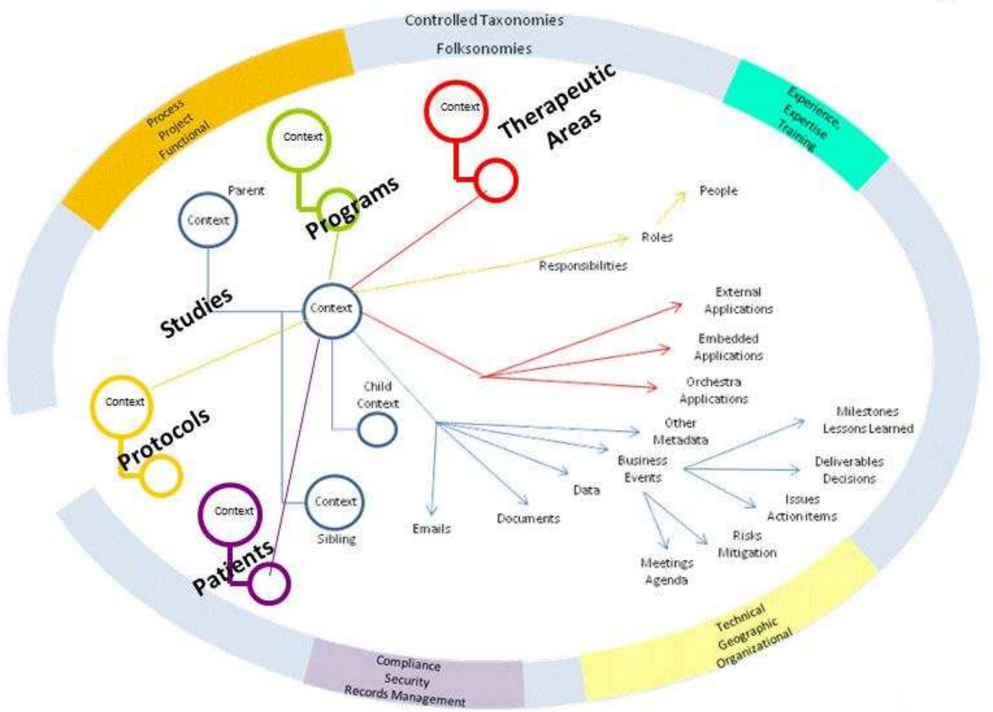
ROLE-BASED ACCESS TO CONTEXT

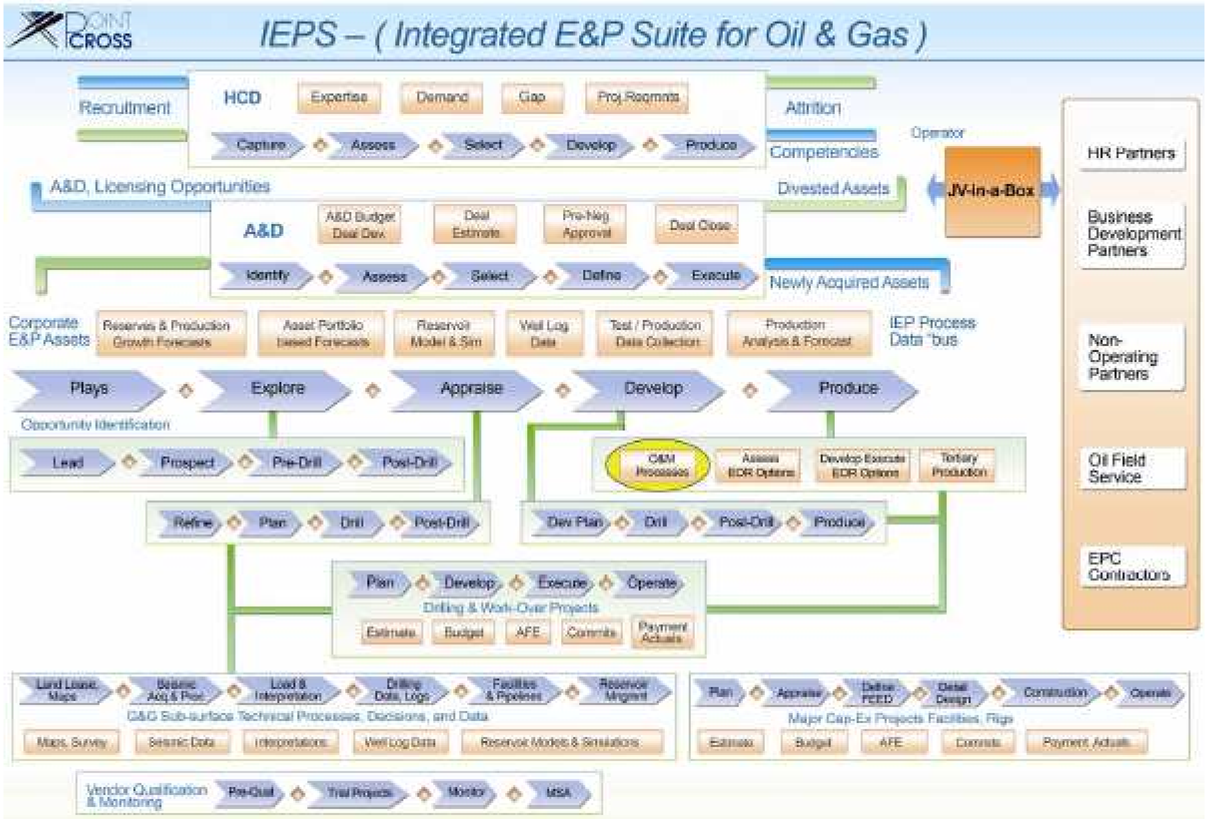
Unlike an individual and her brain, institutional memory involves multiple parties whose roles must be determined by context in order to best offer – and limit – access to specific information. In fact, as our experience has shown, "roles" are of paramount importance. This is because knowledge industries tend to be project-oriented, and they rely on internal social networks more so nowadays as departments and teams are dispersed globally.

The formal roles assigned to people in a knowledge enterprise are the least important when it comes to authorization, access, compliance and security (a separate paper on this topic is available). We look at the project organization as central to the conduct of business and where people are constantly moving in and out of roles that are critical to the specific project but only loosely related to organizational formal role. Social networks, while important, do not play a role in the project other than being an important influence.

Authorized access to contexts is based primarily on project role assignments. Contexts and roles have a many-to-many relationship. Roles and people also have a many-to-many relationship. Each role has a specific level of responsibilities as defined by a RACI (Responsible, Accountable, Consulted, Informed; also referred to by companies as LACTI, RASCI and others) chart that specifies what kind of expectations are levied against a role.

Orchestra Meta-Data Server Relationships



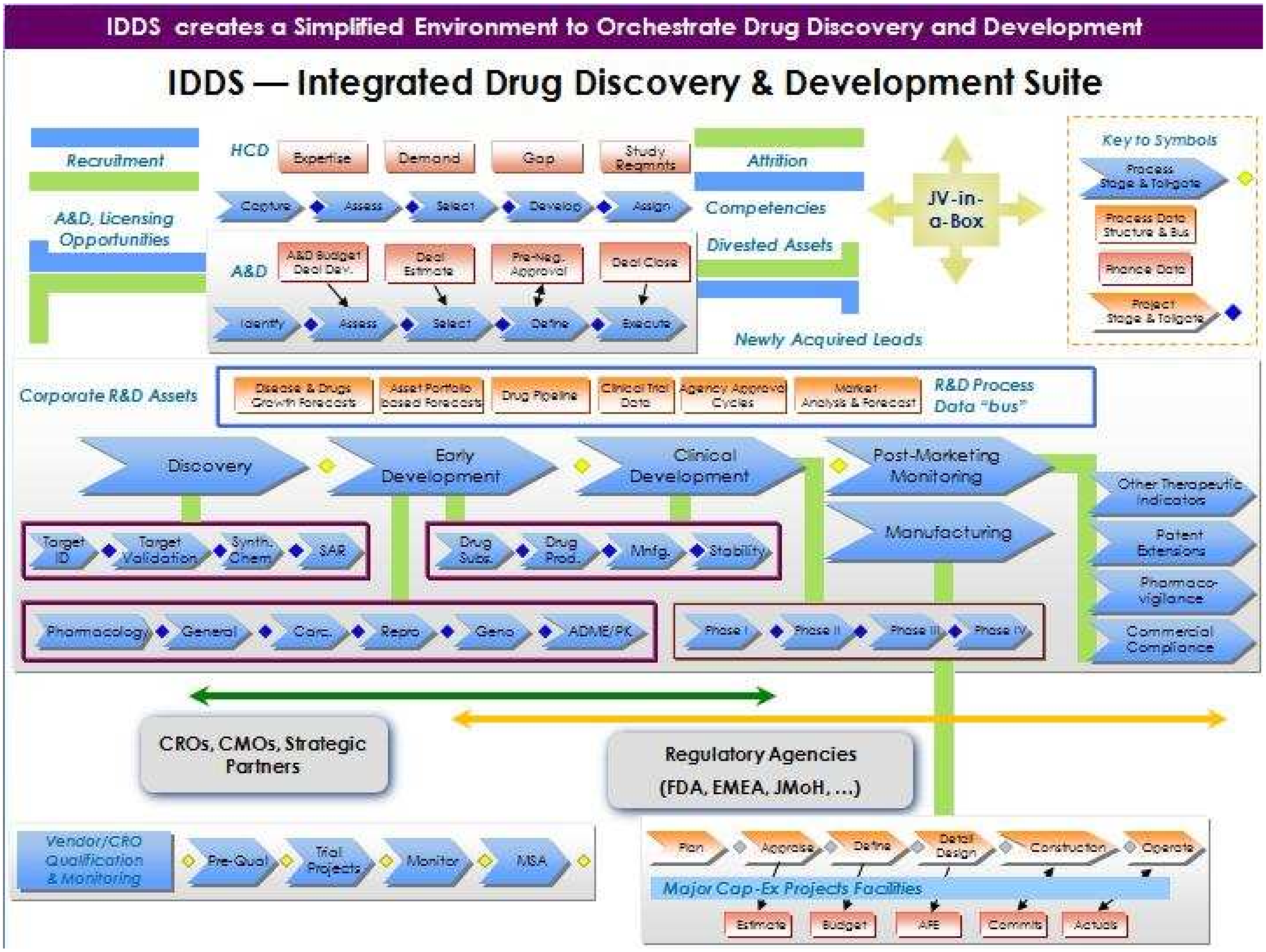


There is a whole discussion around security that is germane – security classification of contexts, people, and their matching and assignment of "need-to-know" – that we discuss in another paper.

Processes: The source and sinks of data and decisions

People don't come to work and just start making decisions. Nor do they randomly pick up something or the other and start looking for data. Their minds are turned on by the context, and the context is established by events such as an email asking them for something; or a deadline that is approaching on a project; or the workflow that beckons their attention towards a technical or business process.

Consider the process map for an E&P company, or a Pharma R&D company. These are orchestrated, interoperating process streams – each stage within them containing their own set of workflows. These processes levy demand for intellectual labor from people based on their roles in the process. Therefore these processes become the source of data for people, which they access through contexts. Contexts are portholes into the process and they provide the view point for the people working on the process. It may be a process step, a project work task, a system component or a HSE audit. The context is connected, as shown earlier, to the data behind the project or process. As employees carry out their daily work, the products of their work are automatically captured into the context, which in turn organizes it into the thematic data representations for the project, system, or process.



Processes are essential sources and sinks for all data and information in a company. These processes need not be very formal – even ad hoc processes have the same effect.

Role of the metadata server

Data and information – large quantities of them – are the way of life in E&P, Pharma R&D, and other knowledge-intensive industries. Raw data may include data collected from animal studies, human trials or assays in Pharma R&D; or seismic data and SEGY files, drilling data, simulation outputs in E&P. But email content, documents, and meeting notes are all raw data as well. If knowledge workers could only find and work with raw data life would be unbearably tough. Instead, to some level or the other, we index, and extract metadata – or data that describes the raw data – and store it in databases. This is because the description of the data allows us to find the same extracted conclusions again and again by different people at different times with consistency.

That sounds good – consistency, simplicity, ease of search are all good things to have in a productive environment. However, there are a number of things we must consider related to metadata. There are some metadata that are objective, factual, and simply existential. Headers of documents and emails come to mind. But then there are additional metadata that attach themselves to the same content over time as people touch it, look at it or change it. These are essential as we continue to create a dynamic ontology or institutional memory of the company. As processes and projects are executed, the contexts in which people work cause them to interact with the data and metadata. Capturing this is an essential requirement to provide the correct perspective for those who look at the data in the future. It is one of the ways in which the learning organization builds its institutional memory.

Consistency is essential for delivering quality out of a product, but it is also dangerous for people who are engaged in discovery, innovation, and invention. Geologists, geophysicists, and reservoir engineers in E&P; or scientists and toxicologists in a Pharma company are not just trying to find what is already known; they are trying to gain new insights and discover hidden trends in the data, or they are trying to test their hypothesis against the data. Therefore metadata cannot be a snapshot-once-use-many exercise. It is important to continually build on the past and capture the annotations and other activities such as search patterns and observations, and to capture them not just against the data and metadata, but in context.

Keeping in mind the earlier discussion about authorization, and particularly role-based-contextual authorization, metadata should be served up as needed to the right people at the times. Analytics plays an important role in this because it is analytics that allows smart extraction of metadata from a mass of data, as well as the ability to find the nuggets buried in the data. Therefore the knowledge system becomes valuable as a resource to knowledge workers when it can serve up the metadata and search the metadata with rapid drill down capabilities into the underlying data that the metadata represents.

Metadata Extraction

The extraction of metadata is not trivial. It is an essential part of creating an enterprise-wide semantic web. It is essential in creating data normalization that allows disparate types of data of disparate standards and their versions within. It plays an essential part in unit normalization when dealing with disparate data. One of the holy grails for all industries is standardization. The madness is not in the search for standards but in expecting that companies across the globe will instantly adopt a single standard and freeze a version of that standard so that everyone can move forward in lock-step. It doesn't happen too often. What will happen is that certain standards that define the behavior of commodity components will stand – like the standard size of a car tire. They define the outer diameter, width, inner diameter, but they provide freedom to the manufacturer in the selection of threads, material compositions and such. Knowledge industries require some of these standards as well, such as the valves, controls, and sensors in the plants, fields, and factories.

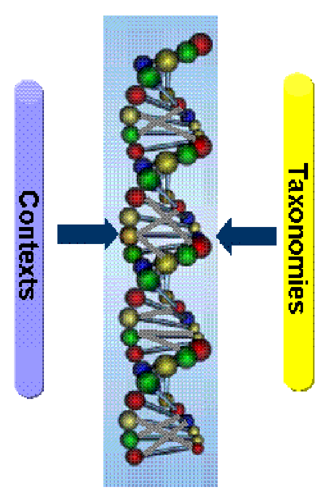
Meta-Analysis

Meta-analysis is one of the most important exercises that knowledge workers should be able to do. It is what allows them to discover what they didn't know but what the institution knows. Remember quadrant 3 in the Search for Knowledge diagram? It is natural that research and project staff intent on a specific assignment will focus, often, exclusively, on that project with less attention to their peripheral projects and activities that others work on. Yet, during discovery, safety analysis, solving a tough technical problem, or other research activities, these same scientists and engineers will need to conduct meta-analysis across disparate projects, slicing, dicing, comparing, looking for trends and anomalous behavior, correlations where none was suspected, outliers that point to a physics or chemistry trait that wasn't understood. When we hear about toxic side effects of an approved drug it is quite often the result of meta-analysis and diligent research of data that already existed that led to such life-saving discoveries.

Ontology

People work in teams to meet the needs of processes, occasionally breaking off into collaborative interactions – tacit, formal or ad hoc – and create new concepts and apply analytics to learn from their work. They use data and information, measure new data, and create or change data and information

Foundations of a Business Ontology



that needs to be organized.

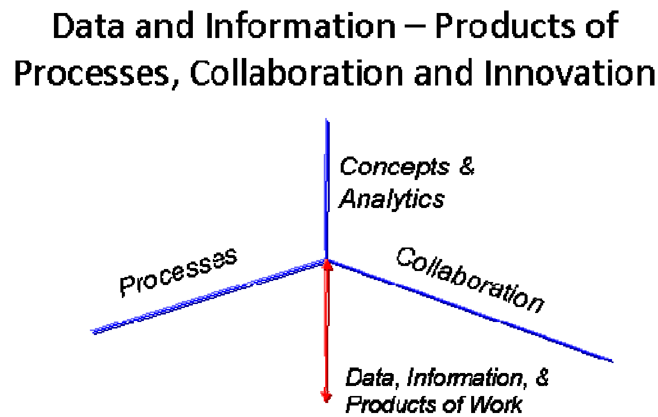
Contexts are the topics, purpose, tasks, or matter that people work on and they form structures that represent the process, equipment systems, or task structures (WBS – work breakdown structures) of the projects. To give meaning to these contexts, the elements of the contexts are meta-tagged by linking them across to the elements of definitional taxonomy – either folksonomy or controlled taxonomies. These are represented as a paired set of structures: one that represents the business as it evolves, and the other that represents the semantic meaning as it evolves over time.

Connecting the processes, collaboration, and other ad hoc innovation spaces or contexts to the ontology foundation, the ontology starts to look like the diagram. titled “Foundations of an Ontology”. The business activities at the top cause the data and information that is churned there to percolate down into the ontology, thus contextualizing the metadata. The physical raw data itself moves down into the data stores while the metadata, including pointers to the raw data, are retained within the contexts.

People work on processes; they collaborate and generate new ideas. But for the most part these are permanent employees who form a pool of expertise. While they work on these projects and processes, they are also the key part of the company’s assets and knowledge. The metadata related to people include the roles that they occupy in various projects, processes, and other activities; it also includes the metadata related to their expertise, as well as their ownership of data and information.

The expertise of people, and the processes and projects to which they are assigned are also defined in the taxonomies (domain trees of the ontology). Therefore the connection of roles into the ontology effectively connects people into a many-to-many relationship between people and taxonomies for disciplines, organizations, geography (spatial) and concepts.

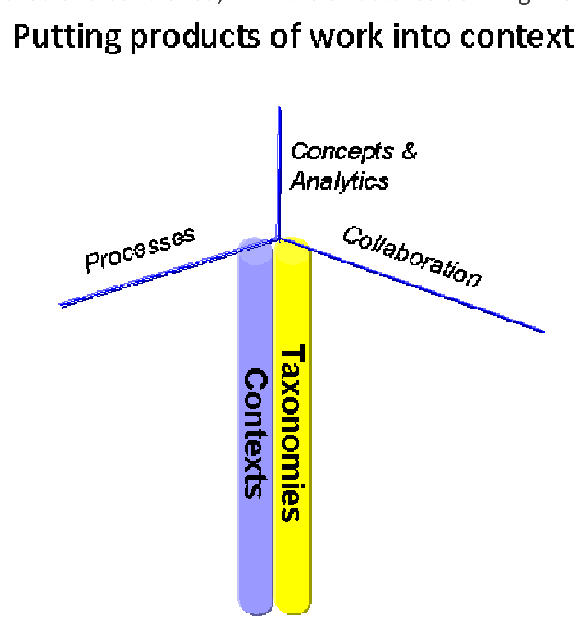
We discussed the vital importance of metadata and how it is extracted and contextualized. We’ll describe the types of extracted metadata a little later, but they include unstructured information, structured data, as well as other external data stores directly or through their parent applications.



DYNAMIC ONTOLOGIES

The ontology of a business is the interlinked structures of contexts that represent various ongoing and past business activities, processes, and functions along with their related data and information. Their links to definitional taxonomies give them meaning and the possibility of relating them semantically. Together the ontology and everything it contains represents the addressable knowledge of the organization. (See Tom Gruber’s formal definition for what we think of as static ontologies in the world of semantic webs on the internet: http://tomgruber.org/writing/ontology-definition-2007.htm).

In a closed, organic enterprise where there is a tight closed loop between new creative work and semantic search and orienteering, ontologies need to dynamically adapt to growth in knowledge as users gain insights from existing information or make decisions. Contrasting this, static ontologies are structured and controlled, with the structures defining the



classes and their organization in an ex-ante way designed to help make sense of an otherwise senseless mass of information by relating them to these ontologies.

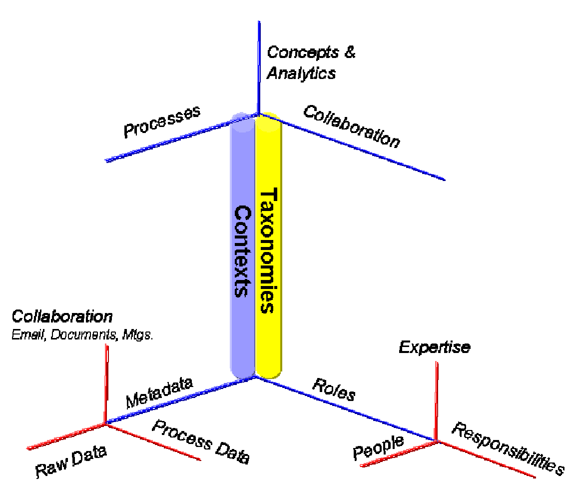
Parsing the activities that knowledge workers engage in, and what they need to support those activities, we arrive at the following:

- ☑ Research: ontologies enable researchers to find what they didn't know, that is known either within or outside the walls of the organization. This means showing exceptions among groups of results that have similarities and unearthing results that have underlying relationships and trends that were not initially obvious.

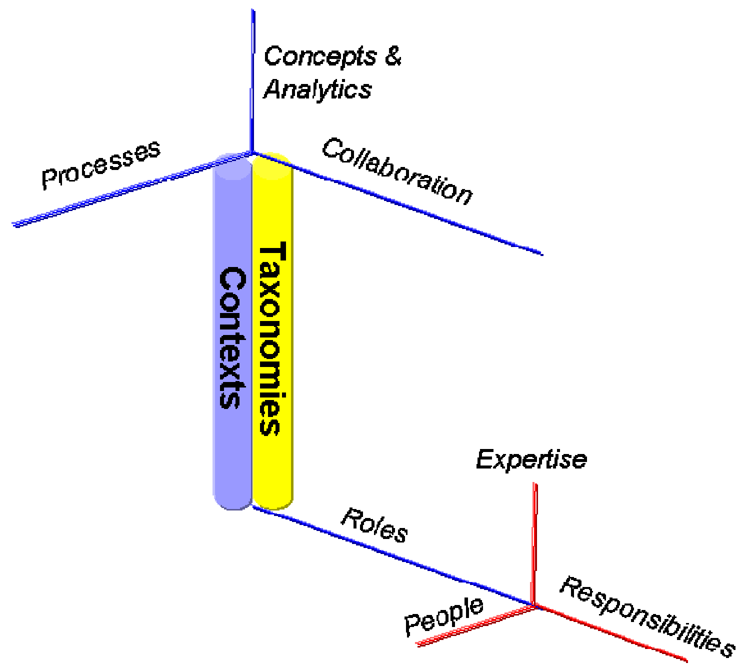
- ☑ Judgments made in support of decisions: show relevant information that is current and with clear indication of the level of its maturity in a form that allows people to make decisions; log the process of decision making; provide the means to collaborate with checks and balances in the absence of an accepted common process

- ☑ Organizing information for Search and Orienteering – the organization of disparate information in a contextually relevant way and where the contexts themselves form elements of a dynamic business ontology – or a structured environment which have

Ontology for Institutional Knowledge and Intelligence



Representation of people form an essential component of the business Ontology



rich relationships that link contexts to meanings and definitions; people with the skills or associations to those definitions; and to process steps that mature the knowledge to a decision point.

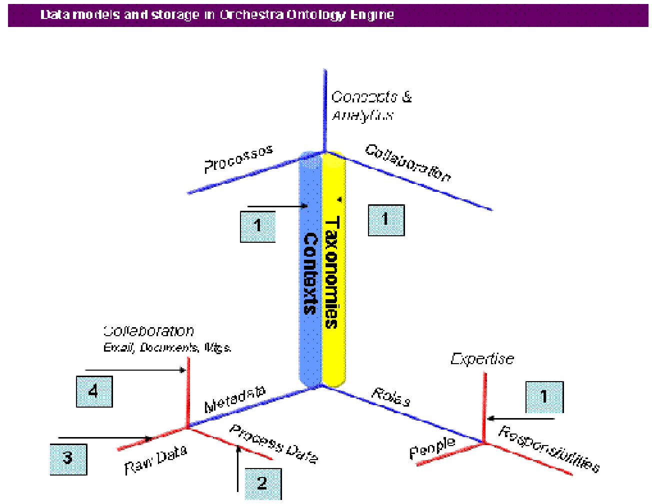
- ☑ Search – Providing the ability to index and search for text within the metadata or the actual content within the documents, meeting notes, or emails that are associated within a context. Search supports the quest of what the user knows exists within the team's, department's, or enterprise's content store. Search may be – ideally – extended to include search through the data and its attributes. Content, data, and their metadata should be indexed so that text string or phrase level search can be conducted from within a context; from a context; from a collection of contexts that represent a concept; or the entire scope of the information organization to find the kind of results that are relevant to the scientist.

- ☑ Orienteering leading to Discovery – Orienteering supports self-navigation by a knowledge worker looking for elusive insights, nuggets, and concepts within a huge body of studies and other products of works and communications using a combination search technologies. Orienteering requires that users be able to set up searches using simple criteria to complex structured criteria that uses a combination of text, semantic, and patterns within data sets - all combined logically (as a Boolean expression or that can be extracted as a Boolean expression from a syntactically expressed phrase) into a criteria that can be captured and recalled in the future. The search results can be seen clustered around the elements in a familiar taxonomy that is created based on the attributes and data sets that the results point to; or clustered around elements of a made for purpose taxonomy. As the name suggests, “orienteering” is not a single search attempt – it is intended to support sequential searches where each subsequent search might take the previous results; or the criteria that delivered a previous set of results and build on it or use a modified version of it. Search criteria might include quantitative criteria applied on the parametric data associated with attributes that describe the data sets. Since orienteering is intended to help those who do not know what the organization might know; the paths of inquiry conducted by the user and past users will leave trails within the information and data sets making it easier to navigate through it and find paths out of the information maze – much like cow paths in a forest that ultimately become trails that can be easily navigates by those who come by later. Orienteering will help users build paths of inquiry through the data.

☑ Planning – support the means to people and teams to plan or organize events or projects. A considerable amount of knowledge goes into planning and tomorrow's projects benefit from the lessons learned from past projects. Project costing, planning (project structure - WBS, scheduling, risk assessment and mitigation are all essential activities that must be supported and whose data and metadata must be captured.

In an information environment knowledge can be represented best if the following concepts are followed:

☑ Unified Business Information - Information of any kind that is related to a topic is kept closely associated with



one or more contexts that represent that topic and which are relevant to business or technical decision making. Consider the kind of data and information that a user or a process will be required to jump across based on their relationships or links to a common context and ensure that these are not stored in unconnected storage silos. Unifying disparate related content around a context is essential as discussed earlier.

☑ Meta-Data - Establish an extensible wrapper that is associated with each and every type of data or information that can contain the metadata about the data or information. Most of the currently available metadata is really file header information that is of some use to IT administrators but carries little value for knowledge workers. Metadata of many types must be supported:

- Standard metadata consisting of the header information
- Metatags that are automatically assumed as attributes based on business rules including text matching, text string comparisons with taxonomies, or semantic rules
- Metatags that are manually added as attributes – these can include standard terms in a taxonomy and links to those elements in the taxonomy

- Metadata representing the use of the information – who, or what (process, search) touched it, for what business purpose (context from which it is accessed), when, changes/versions

- Metadata added by annotation of the information (Web 2.0 type facilities)

☑ Metadata Extraction – All information and the especially the relationships between contexts; the content, any data should be analyzed, indexed and made available for search and analytics applications

☑ Communication – provide the ability to communicate through emails, IM, blog, SMS directly from any information or content object.

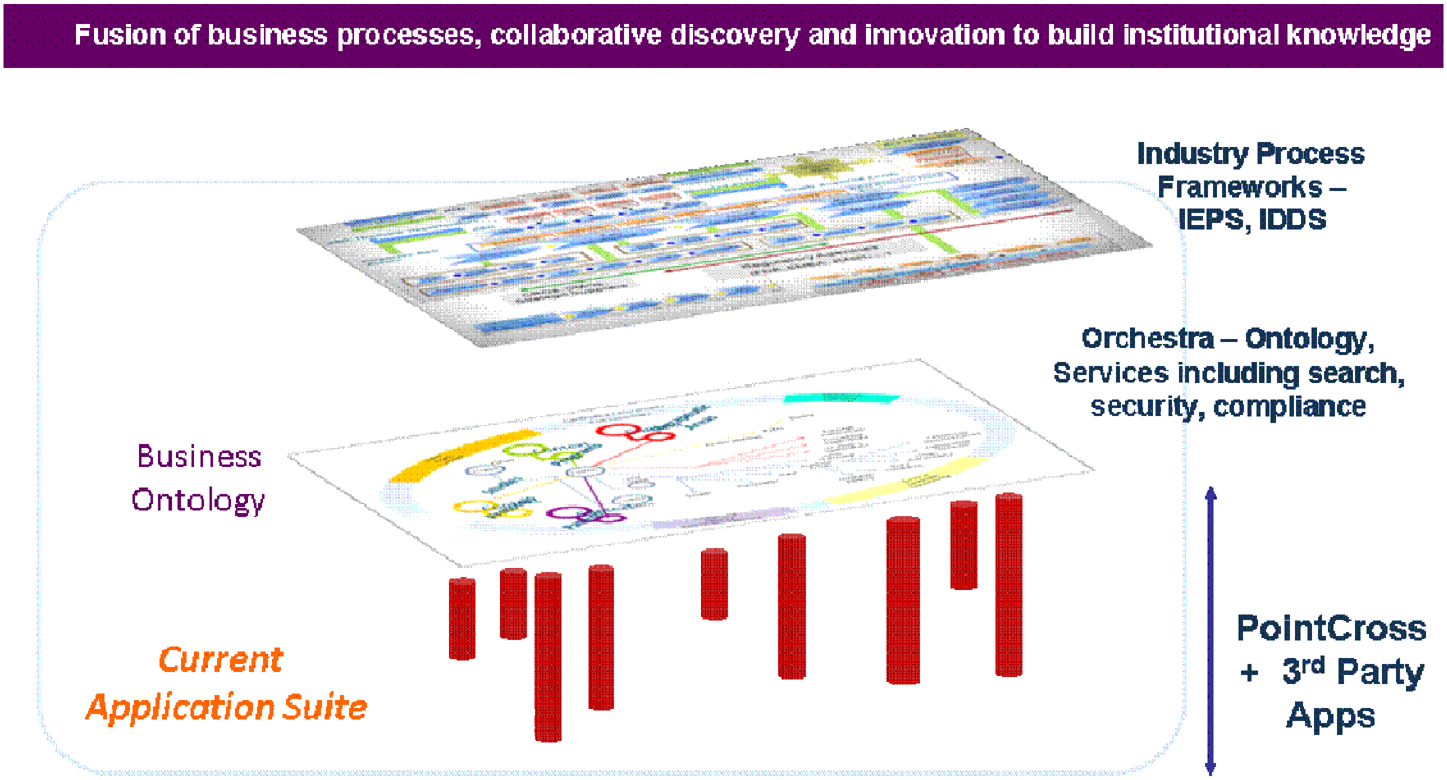
☑ Definitional Taxonomies – Taxonomies are structured definitions of terms, with child terms inheriting the meaning but offering further granular definitions. The data structure allow for synonyms and AKA (also known as) terms. These taxonomy terms are available to be linked to contexts, content (as attributes or metatags) roles, people (to describe their expertise) and other contexts that represent process stages. These links are very important because they provide the multiple paths to content by creating the semantic web that is used by search and navigation applications.

The act of organizing information in folders does not make it easy to find content in the future – even for the person who organized it. This is because not every person who is looking for the information is going to remember, or re-construct the path to the folder the information was organized. The fact that knowledge workers spend as much 70% of their time looking for information in multiple locations is proof enough. For example, it is common practice for those who are very organized among us to spend the time to create a tree of folders in Explorer, and then again in the email client such as Outlook or Lotus. When we want to find a document we go down these folders structures until we find the one we want and then we look for the files. Yet, anyone who has installed a desktop search knows they are more likely to simply search for the document by keying in something about the document or the folder it is in. That makes it infinitely more useful for someone who is not familiar with that folder organization – such as in a server side document management system. This begs the question "why not simply put all the content in a single folder and use search to find the documents" – not an irrelevant question particularly since the whole technology about semantic webs is to ex-post organize and index information against one or more taxonomies or an interconnected, but static, ontology.

The taxonomy links discusses here provides a multitude of ways to the context. Whether manually or with the help of automated search engine, it becomes much easier and it is possible to use the poser of semantic relationships to find the content within ontology particularly in an active enterprise where the structure of the ontology is dynamic and constantly growing.

This last point needs additional thought when considering how ontologies are set up and used. The way semantic web technologies are evolving, most of the work is in making sense out of the information that is "out there" on the w3 or perhaps within the intranet. The semantic parsing, analysis, and the comparisons to the taxonomies (which are pre-defined and maintained under standards) or controlled ontologies allow these ontologies to become analogous to a scaffolding that gets build around a building. The content found is connected or linked to the ontology nodes based on their semantic meaning and relevance. In this model the ontology forms the structure which is created ex-ante (before the ontology is available to the community and the semantic web) by a body of people who have the rights, skills and knowledge. The content is linked to the ontology through a process somewhat similar to a crawler but it is done ex-post (after the event of creating and publishing the content).

Contrasting this, dynamic business ontology is a hybrid of business contexts (topics) which are functional but constantly changing and being re-structured to account for the contemporaneous needs of the line of



business, its processes, the ongoing collaborative work of specialists and experts; as well as the static taxonomies that metatag the business contexts and content. In another dimension the taxonomies of roles and people are constantly changing and they attract metadata as well as metatags. Between business contexts, roles for projects and organizations, and people the ontology is dynamic with changes happening contemporaneously even as these very contexts, and people in their roles are adding, changing, and publishing content and data.

☑ The information systems must allow for multiple types of definitional taxonomies:

- Folksonomies – These are hierarchically organized (structured) definitions that are used by the working teams as part of their processes and general collaboration. These terms are relevant locally and may include colloquialisms, acronyms.

- Controlled taxonomies – These are identical to folksonomies except that they are controlled by a central authority that ensures that the terms in this data structure are formal, validated and provides the semantic relevance and when linked across folksonomy elements provides meaning equivalences.

☑ Pattern recognition – this involves serious analytics on the metadata as well as on the data that is associated with the contexts. A whole paper dedicated to this topic is required to pay adequate attention to this topic. It is also true that the state of the art has a way to go before some of the basic benefits such as creating a business déjà vue can happen in all facet of work. However, discussing it at a superficial level, the state of art does provide some exciting possibilities using pattern recognition and simulation. Couched as predictive analytics there are a number of solutions being built and some very interesting possibilities are emerging.

Fusion of Process Frameworks, Ontology, and Applications

The Orchestra ontology engine based platform uses a tool kit (FDK) to craft specific solutions and business process automation solutions as shown by examples such as IEPS (for the oil and gas exploration and production) or IDDS (for the drug discovery and development). The applications shown in red columns are either 3rd party legacy applications or even special purpose tools built for purpose.

The diagram shows the value of an ontology engine in a knowledge enterprise very clearly. Normally companies will acquire various tools like the ones shown as red columns and they will try to integrate these applications to improve the productivity of the people who use the applications. Enterprise Application Integration is notoriously expensive. And it is not surprising as EAI follows Metcalf's law – the power, or in this case, the cost, of a network of nodes – where a node represents an application as a data port – is n*(n-1) where n is the number of applications that need to be integrated.

With Orchestra and its ontology engine providing the contexts into which applications and their data can be provisioned means that the cost is reduced to "n". There is no better way to prove this than by experiencing the low cost and high speed of deploying solutions and processes that incorporate the legacy applications. We usually are able to demonstrate 1/10th the time or better even for the most complex business process problems when compared with database based platforms including SAP or Oracle; keeping in mind that Orchestra is not intended to replace either.

Effectively the fusing of processes, ontology and applications ends up creating a layer of reusable institutional memory. Adding new analytics in the ontology layer or as external applications end up adding layers of intelligence to this set up.

How data is managed in Orchestra

Applications grounded in a static, monolithic, data model might be appropriate for transactional business solutions but it is a disaster for applications like content and document management or to manage data like study data or assay data. It is difficult and expensive to manage change, versions, normalize , harmonize, or maintain integrity over the long haul. Of course, ISV (independent software vendors) companies keep building those because that is what they know.

We do not follow this model for business solutions or processes. We do use fixed data models to define essential constructs within the Orchestra platform. But we take a very different approach with Solution or Process data.

The Orchestra Ontology carries a multitude of data types. Orchestra handles these in different ways depending on their position in the Ontology hierarchy. The other consideration is that certain types of data are intrinsic to the Orchestra Platform; while others are specific to the Solutions and Processes built on the Orchestra platform. IEPS is such a solution framework for the E&P industry, while IDDS is for the Pharma industry; SDI, a study management tool that utilizes SDE (Semantic Data Exchanger) is an example of a tool that fits within the IDDS tool.

Orchestra's architecture breaks out the entire data handling into three sub-systems – a middle tier; a file management system; and the database itself. We use a fast, automated converter that we call DBXML to do rapid, on the fly conversion of data, or sections of it, in the database tables to XML using their corresponding XML schema. DBXML also rapidly writes data from the XML data (using its schema) into the data base tables as data enters the system or people working on processes change it. Some reasons for this architecture are discussed further below but suffice to say that we benefit from the incredible flexibility of schemas that are capable of having semantic relevance and invokable on demand by processes or user actions. This simple migration to a middle tier has allowed us to support the huge footprints of our solutions and business process with just a couple of DBAs on our staff. We do entire, complex, systems for customers with nary a DBA.

So, we use multiple ways to store, and hold, data at various points while ensuring information integrity, security, access authority down to cell level (without dependency on data base security constructs which we find to be limited and restricting. These include:

A) Data model built with tables in the database of choice (Oracle, Microsoft SQL Server, and PostgreSQL). Most of the core data stores for contextual data in the Orchestra Platform are managed in this manner.

B) As files or XML wrapped files in the file management system, XML schema files are also used by the middle tier as part of the DBXML application we use to mirror XML coded data with flat file storage of the actual data in the database (described earlier).

All the business process and solution data is stored this way.

C) Binary encoded files that we use for certain internal transactions and flags such as used for internal synchronizer caches. We also use this to handle many cached operations related to parametric data analytics and conversion into the XML data files. It is not a form in which data is stored for the long term.

We manage all metadata for “Orchestra platform level application” data-stores in the database. This is because all basic activities that people are likely to engage in are happening all the time and they are contextually relevant. People write email about everything, they share document, meeting notes, and make changes to them. They are concerned with things like content, costs, schedules, risks, issues, basic workflows. So, all of these are intrinsic properties of the key object in Orchestra - "The Context".

When it comes to business solutions, we configure the solution or process in terms of contexts and then we allow the expression of all platform level functions through the contexts. The data related to the solutions is usually thematically decomposed (a solution design decision), contextually sliced and managed in the middle tier through XML files that will contain the precise data that the application layer need in real time. This is done by our DBXML dynamic converter.

This approach allows us to manage the data base very efficiently. We use only the simplest of data base SQL queries. The queries are very simple. The speed is very high because we maintain simple flat, or quasi flat tables even though the real model may be complex; and we can handle very complex data models and be able to change them as new versions come up. For example if the 2 dozen or so assays mature and the technology causes new data types to start appearing it is a simple matter to add on new data models while managing the old ones and still be able to compare them on thematically equivalent basis. Another important benefit is that it makes it very easy to re-purpose data and merge disparate types of data and unstructured data into a common publication.

We manage business solution and process models in XML schemas – each of which is kept relatively small; and with a single theme around it. We also use this schema and the various data names within it to give it semantic relevance by indexing these model files and linking them to the folksonomies and controlled taxonomies. There is a lot of minutia related to how these are handled. By breaking out the data and its handling into these two systems what we accomplish is the ability to build ultra simple – and sometimes flat file – tables for these XML models within the database and these are done by some proprietary automation that is part of Orchestra. This allows us to depend on

For example, consider the keys in the diagram (the ones with a number in them pointing to each area):

1. All Contexts, the trees that structure them; and collections of trees are intrinsic to the Orchestra platform. Contexts, roles, people, and taxonomies are, likewise, intrinsic to the platform and therefore have a fixed data model that resides in the database. Each of these has a complex data structure (ER). As contexts, people, taxonomies are added these data models are replicated within the database like cells in a tissue – same ER structures that replicate; and which are connected to others through indexed links. That and the properties of the entity (contexts, people, roles, and taxonomies) is the only thing that differentiates their behavior.

2. Process and Solution Data: These are built on top of the platform using our FDK tool kit or using the 1,200+ APIs and web services that the Orchestra platform provides. In this case we do not begin with a DBA designing an ER model for the database. Instead, we define the data model as a XML schema.

   In the case of an assay or a study the data model is self-defined by the data source itself. Each of the studies (actually the trials within in the study) or assays will yield their own “appropriate” data model. But it is also quite often the case that the kind of data and the model itself may undergo changes over time as the technology is developed or evolves; and the same goes for the studies where the protocols themselves will go through changes. Another thing that changes is the data models used for reporting data to, say, regulatory agencies.

Because we need to deal with disparate data models (one for each assay or trial); and then again their various revisions and changes we simply maintain an XML schema – one for each. We keep the metadata about the data as well as the schema and its genesis and its semantic relationships in an XML model as well.

3. Raw Data – Includes bulk data from external ODBC sources, data stores from technical systems integrated into the platform to serve a solution such as parametric data from trials in a study, a gene expression system, a MRI machine, some equipment used for an assay. A lot of experimental or assay data falls in this category. The data itself may be imported into Orchestra – each case or assay data will be assessed before deciding if the raw data will be imported into the Orchestra data store. If the source system is stable, capable, and a system of record, it may be wise to leave it there and simply read it; if not it will be imported. Once imported, a set of analytics will be applied and they will crawl through the data to extract metadata from the raw data. Again, the kind of metadata and its structure is designed to suit each set of source raw data (each assay). The metadata extracted will be initially designed to meet the kind of search or queries that might be made from the solution or business process – therefore ti is designed to meet eventual business needs. Keeping in mind that these needs will keep changing as people learn, these analytics will evolve to meet the emerging needs and therefore the metadata structure will also change. This is another example of why the metadata structures are maintained in XML with the DBXML handling the mirroring with the database itself.

   The ESP data will therefore be kept it in its native format either in the source environment, or in Orchestra. The metadata will be extracted through a crawl by a set of analytics, and then the metadata will be stored in an XML format and managed by the DBXML. If the metadata is very row heavy and there are a huge number of records the DBXML will hold the data in the database as quasi flat files. If certain thematic groups of metadata that are extracted are best handled purely as XML because, for example they are very column heavy and they have a lot of meta tags, then the actual XML data files will be stored within the file manager; the database will hold the headers and pointers to get ready access to the files.

4. Content of Collaboration

   a. Emails - header information and pointers to the email content, summary text of the email (generated from the search indexing), author, recipients, reply-to, subject and other data are managed in the platform level with strong relationships to the corresponding context's contextual data model. The actual content is stored in the file management system with each email within its XML wrapper.

   b. Documents - the header information and pointers to the document content, metatags consisting of extracted words, word stems, and manually add or validated tags and attributes, records management attributes, (generated from the search indexing), author, and other data are managed in the platform level with strong relationships to the corresponding context's contextual data model within the database. The links to other contexts where the document is linked, or the touch-point history data as well as the revision data are maintained in the data base.

   c. Other context specific data include narratives, cost, scheduling, risk, issues, and many others.

The actual content in each of these cases are stored in the file management system with each document stored within its XML wrapper.

Ontology Engines and Business Process Automation

The traditional data model driven N tier application can automate workflows and model business processes – but not efficiently or effectively.

Ontologies which depend on the business context are situated ideally to model the business because both the raison d'être and the disparate data needed is already embodied within the business context. This makes mapping a business process a simple activity; multiple processes can operate on the context which is a prototypical object of business. Ontologies allow the complete elimination of at least two layers of consulting that currently serves to obfuscate.

The other reason why ontologies are so important is that the data model, far from being monolithic is a naturally assembled gaggle of smaller thematically relevant model that is associated with the context – the business context – rather than some vague monolithic ER (entity relationship) model. This means that data component models can behave according to business rules and easily be re-purposed and made available to disparate contexts as the publish-subscription rules permit.

Summary

Traditional fixed data model based solutions are great for transactional or manufacturing industries which is op-ex centric; but not for knowledge intensive, dynamic industries where the core line of business is about discovery, invention, developing new things and heavy on the cap-ex side. The need for active, dynamic institutional memory with the potential to support decisions with the right combination of current and past information is high and that means an alternate architecture for the information environment is required.

An alternate architecture must be based on a dynamic ontology for organizing disparate business data and information; active analytics and business applications, whether embedded or provisioned from an inventory of 3rd party and legacy applications. This combined capability, which we call an ontology engine, should power the information environment. The ontology engine, as a platform, should be capable of being triggered from outside by other application using a variety of web services and APIs. It should also be capable of providing services to the external legacy environment as well other ERP or transactional functions. Toolkits for rapidly configuring solutions, and pre-integrated templates of industry solutions should be provided so that client organizations can rapidly automate their key business processes and get people into the habit of being able to focus on the business of their business while the data and information they generate are self-organized; and they are able to discover authorized information contextually.

An Ontology based environment like Orchestra creates true institutional memory with unending possibilities of discovery and innovation based on a range of analytics that can work on the data and information from the line of business processes and projects; the activities of people; and the decisions they make. People collaborate contextually building their social networks; while the same contexts participate in choreographed and orchestrated processes building their expertise profiles and the institutional skill pool spectrum.

Orchestra's ontology of the business grows to be a high fidelity virtual replica of the company and its activity.

About PointCross

PointCross is a global provider of advanced strategic business solutions to knowledge-rich markets, including the pharmaceutical industry.

Our Integrated Drug Development Suite (IDDS) specifically addresses the pharmaceutical industry's key R&D concerns. Likewise, our Integrated Exploration and Production Suite (IEPS) addresses key needs of the upstream Oil and Gas business. At the heart of these solution frameworks is Orchestra+Solo™, an adaptive, contextual knowledge environment and personalized client that orchestrates core business processes. Among its diverse capabilities, IDDS and IEPS deliver:

- ☑ Single point of access to contextualized tacit and structured knowledge across the enterprise, with search and guided navigation within and across contexts;
- ☑ Flexible, fool-proof IP security based on contexts and roles, determined by business rules;
- ☑ Predictive analytics for critical business data;
- ☑ Secure multi-party workflows for knowledge sharing and corporate social networks within and across companies;
- ☑ Solutions for core business processes, deal flow, acquisitions and licensing, e-discovery, audit, compliance, and more;
- ☑ Scalable architecture and development toolkits for additional capabilities.

PointCross represents a new way of doing business. We deliver business ready solutions in 1/10th the time and a fraction of the costs compared to standard technologies.

We are headquartered in the California Bay Area of the US.

For more information, visit us at www.pointcross.com and call us at (650) 350-1900.