Effect of Business Transformation on ERCB's Information Architecture

Prepared By:

Dwayne Popowich, Dipl. Arch., BGS, GDM, MBA
Executive Director, Planning & Architecture Group
Energy Resources Conservation Board (ERCB)
Information & Systems Services Branch (ISSB)
Planning & Architecture Group (PAG)
Enterprise Architecture Section (EAS)

August 6, 2012

Abstract. Driven by continued work on the Oilsands Regulatory Framework and the introduction of a new risk-based and play-focused Unconventional Regulatory Framework (URF), the envisioned business changes to the Energy Resources Conservation Board (ERCB) are significant and will result in deep changes to the informational aspects of the organization.

The facts, observations, and trends identified in this paper are the result of discussions with numerous thought leaders and subject matter experts both internal and external to the ERCB. The author thanks the many content contributors who have provided their valuable insights and perspectives and helped to identify the issues that must be considered throughout ERCB's transformation.

The purpose of this paper is two fold. First to explore the impact of regulatory evolution on the Information Architecture within the ERCB, using the identified topical areas expressed in Sections 1 through 9 as a guide. Secondly, to educate and inform ERCB's architectural community and its decision makers on informational issues but provide no formal direction in these informational areas. It is anticipated that decisions and direction setting on the identified topical areas will be gained through the Enterprise Architecture Council.

1. Introduction to Information Architecture

So, what is Information Architecture and how does it affect an organization and its business functions?

The term "information architecture" was first coined by Richard Saul Wurman in 1975 to describe the way information is gathered, organized and presented to convey meaning.

Wurman's initial definition suggested that Information Architecture was organizing the patterns in data, making the complex clear. Today we know it to be more than that.

To understand Information Architecture we must first define the term 'information' and establish a distinction between the terms: 'data,' 'information,' and 'knowledge.' While these words are related, they are not interchangeable.

Data is raw; it has no significance beyond its own existence and is generally recognized as a discrete textual or numerical construct (e.g. "Well" is a textual construct).

Information, on the other hand, is data that has been given meaning by associating it with a context, a set of rules and relational connection. Information is generally disseminated as a set of linked and related data components (e.g. "Oil well in Southern Alberta" is an informational statement that gives a richer understanding of the data element "well".)

Knowledge is the appropriate gathering and understanding of information, such that it becomes useful within a cognitive event. Knowledge is a deterministic process. When we apply understanding to information then we amass knowledge.

Architecture has a duality about it – it is a noun and a verb. Specifically, it is a process of design in which humanism is at the core and is the formal representation of inter-relationships into a recognized and useable form.

As such, Information Architecture is the process of assembling meaningful sets of linked data components into recognizable elements that have contextual reference and can be consumed and utilized within the human environment.

The ERCB utilizes its information to reach decisions and create knowledge to inform and advise internal and external stakeholders.

Insomuch as information is what is used to create knowledge then it is clear that knowledge and information are symbiotic and related to decision making. Information, not data, is the primary currency for intellectual exchange and is an extremely important asset within the ERCB.

However the data components that are assembled into information must be accurate, available and easily accessible. If these primary conditions are not appropriately addressed the resulting informational products can become (and often are) mistrusted and the opportunity to amass knowledge lost.

Further, Information Architecture is directly related to human actions and interactions. Therefore, understanding the human environment in which information resides and is used is critical to the development of an organization's Information Architecture.

In addition, in an information-centric organization such as the ERCB, Information Architecture underpins the Business Architecture (expressed as business capability, rules, processes, and work efforts).

2. The Unconventional Regulatory Framework (URF)

The URF, as a new risk-based and play-focused approach for accomplishing regulatory work, introduces the need for:

- An operational model that has significantly different informing, planning, authorization and compliance processes
- Greater collaboration between industry operators and the ERCB through the utilization of a central "earth model"
- Well defined indicators, thresholds and business rules to produce social, economic and environmental outcomes that reflect the impact of regulatory efforts during hydrocarbon development activities

To support this new approach the ERCB's core business processes and associated informational models must be linked to a larger set of social, environmental and economical informational landscapes; beyond what the ERCB information systems can provide today.

Therefore a requisite integration and sharing of the informational products between ERCB Branches and varying external information providers is necessary.

A play-based regulation regime that utilizes an outcome monitoring and reporting approach to establish knowledge within the ERCB itself and the industry in general are new concepts. Therefore, the URF concept contains a plethora of undefined informational constructs that range from small to very large attribute sets, (i.e. a Play Development Plan introduces a multitude of aspects – landscape conditions, water availability, air quality, etc.).

Further, the potential introduction of a play-focused regulatory approach suggests a diminished importance of pools and fields; a paradigm shift within the ERCB that may culminate as the introduction of new informational products, while driving the potential elimination or modification to existing informational components.

The URF concept builds upon a well-centric information model (i.e. the well is the link point to all other data components within the ERCB's information environment – see Appendix A).

The envisioned informational "shift", however, is to develop a broader surface-based hierarchical informational structure (ontology) that starts with a development plan, then individual project plans, numerous pads and finally large numbers of individual wells on the pads. The result would be an enrichment of the informational constructs associated with a well.

The URF concept also brings to the fore the opportunity to link the surface based informational set to a sub-surface informational environment, forming a rudimentary framework for a three dimensional digital *Earth Model*.

Of significance is the introduction of *Outcomes Monitoring*. The idea behind this approach is to monitor, on an ongoing basis, the impact hydrocarbon resource development may have on a broader social, economic and environmental scope. *Outcomes Monitoring* will require the introduction of new informational sets - thresholds, indicators and outcomes.

It is envisioned that the *Earth Model* will be enhanced with other data components (e.g. social and economics based informational sets). The nature of these enhancing data components is presently undefined. However, numerous bodies of work, including the *American Petroleum Institute's Oil and Gas Industry Guidance on Sustainability Reporting – Using Environmental, Health & Safety, Social and Economic Performance Indicators* provides insight into the numerous indicators that could be utilized.

While the evolution of the ERCB's organizational environment is, at this point, ongoing it is clear that the informational needs within the ERCB will change, driving a need for sweeping changes to the data and content management structures that are currently in place.

3. Regulatory Monitoring – Outcomes, Indicators & Thresholds

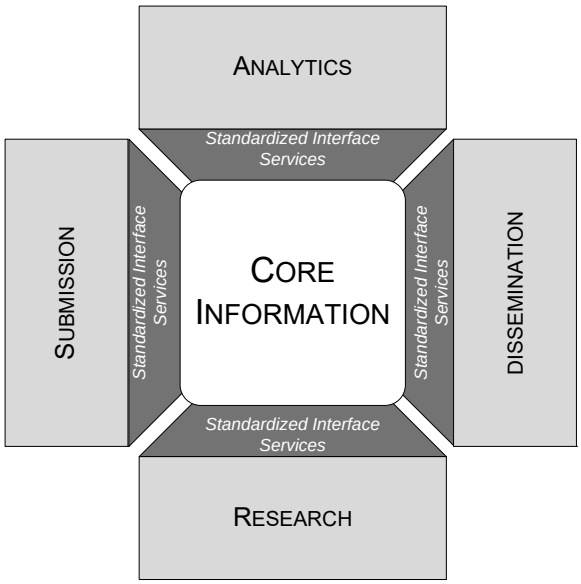
With the introduction of a risk-based play-based regulatory framework comes the opportunity to introduce a performance-based *Outcomes Monitoring* approach that will be the primary "informational engine" for an adaptable regulatory regime.

More specifically, it is envisioned that individual regulatory approaches specific to each defined *Play Development Plan* will be developed through collaborative consultation that includes the Regulator, Industry and the General Public.

The monitoring of social, economical and environmental outcomes will provide "feedback" on the impact of the defined regulatory approach.

If it is deemed that the directing regulations are insufficient (reflected in the displayed outcomes) the Regulator will then assert more stringent requirements on the play-based resource development effort.

The following diagram expresses the general informational framework of the envisioned performance-based *Outcome Monitoring Environment.*



Proposed Information Framework

Generally the envisioned *Outcomes Monitoring Environment* will consist of Standardized Interface Services that support:

- The receipt of data from various sources (Submission) that are transformed into a standard format for interim storage (Core Information)
- The processing of Core Information by applying Indicator and Tolerance Business Rules to the data set (Analytics), which will result in one or more *Results Sets*.

The *Results Sets* may be used as input into the development of new Regulations or the identification of new Outcomes or Indicators (Research) or drive the creation of Reports and Dashboards for consumption by a variety of stakeholders (e.g. the Board, Industry, Government and/or the General Public) (Communication).

The *Outcomes Monitoring Environment* will consist of three primary subject areas (see Appendix B for graphical depiction).

Each subject area would represent a separate area of

information creation or consumption, and each would serve a separate business purpose.

Subject Area: *Reference*

The *Reference Subject Area* is essentially the "engine room" of information needed to allow the ERCB to categorize and measure *Indicators* in a consistent manner, and provide the necessary structures to relate *Indicators* to *Outcomes a*nd ultimately to the Board's risk appetite by way of *Risk Categories*.

This subject area introduces the concept of *Tolerances*, which are pre-determined limits of acceptability for an *Indicator*. *Tolerances* are generally associated with one of four specializations sub-types – thresholds, objectives, guidelines, variances.

Subject Area: *Feed*

The *Feed Subject Area* addresses source data that is used as input into the *Outcomes Monitoring Environment*. This source data may be provided by external parties or from elsewhere within the ERCB. The data and its sources will be variable. Therefore the over-arching Information Architecture must accommodate flexibility in the "feed channels" that support the *Outcome Monitoring Environment*.

Preliminary analysis has identified two initial data sources – the first is structured, repeatable data from "official" sources such as Government, Government Agencies and Industry, and the second is data gathered from the General Public by means of occasional questionnaires.

Subject Area: Monitoring and Reporting

The *Monitoring and Reporting Area* is associated with the preparation and dissemination of developed output. This subject area retains and maintains data associated with *Alerts* (an event that is triggered when a *Tolerance* is breached or approached). It is envisioned that this subject area will also contain the time-based collection of *Results Sets* used to create reports and dashboards so that these can be quickly re-created for trends analysis and historical comparison.

- Dashboards

Dashboards are highly visual, interactive reports that allow the end user to see trends and status at-a-glance, in a consistent, easily consumable way. Typically, dashboards use intuitive visual cues related to everyday recognizable devices such as traffic lights, speedometers, thermometers, etc.

It is envisioned that informational output associated with *Outcome Monitoring* will be through a number of context-specific dashboards tailored to different audiences.

Audiences could include Industry, the ERCB's Board, internal and external researchers, the Government of Alberta and the General Public.

It is further envisioned that the completeness and depth of information presented via a dashboard will vary greatly depending on the target audience.

- Reports

It is also envisioned that the *Outcome Monitoring Environment* will support numerous "traditional" reporting needs that generally fall into two broad categories – Monitoring Reports and Reference Reports.

Monitoring Reports

Monitoring Reports would be designed for consumption by both internally and externally stakeholders.

It is anticipated that these reports will be static versions of the *Dashboards* and would not offer drill-down capability. Typically, the reports would be *Outcome Type* or *Geographic Location* based and further sub-categorized as required. Reports would typically be for a fixed time period (e.g. *Indicators* and *Tolerances* categorized by *Outcome* and *Outcome Type* for a given *Play or Region*).

Reference Reports

Reference Reports are for internal consumption and serve two primary purposes - ongoing improvement of the *Outcomes Monitoring Environment* and the continuous development of new *Indicators* and *Outcomes*.

Public Interface

The *Outcomes Monitoring Environment* would provide a public interface to Industry, Government Departments and the General Public, providing mechanisms to receive regular or ad hoc data feeds in a pre-defined format. These feeds may take the form of an RSS feed, a flat file, or an XML document, and could be used as input into other dashboards or reports as required.

It is envisioned that information flow between varying systems will be bi-directional and near real-time making the *Outcome Monitoring Environment* an effective and efficient feedback loop on resource development activity.

4. An Earth Model

It is envisioned that an integrated digital *Earth Model* will support the new regulatory approach by providing:

- A highly accurate and visual depiction of the Province of Alberta's geology at the surface and sub-surface levels
- A highly accurate and visual depiction of hydrocarbon activity (i.e. well bores, pipelines and facilities) within the geological space
- Visual support for the outcomes utilized for near real-time reporting on the current status of regulatory impacts at the local, play, regional and provincial level

A Single Geodetic Standard

To design and deploy an *Earth Model* the ERCB needs to identify and agree upon:

- A single geodetic standard for all ERCB's geospatial data so that relationships between physical location, analysis, reporting and display can be established
- A common spatial reference frame in order to map the world accurately and precisely. In a digital environment this is achieved through the use of horizontal and vertical datums, tied to the International Terrestrial Reference Frame (ITRF) that allows for a standard exchange of geospatial (mapping) information.

In the Province of Alberta the standard horizontal datum in use is North American Datum 83 (NAD83).

NAD83 has evolved over the years and reflects improvements in surveying technology and survey accuracy. Future adjustments in NAD83 are expected to be small, about the same order of magnitude as of changes in the earth's crust (i.e. centimetres); as such, can be considered relatively stable.

However, Canada's (and therefore the Province of Alberta's) standard vertical datum is the Canadian Geodetic Vertical Datum (CGVD28) defined as mean sea-level determined from data collected at five tide gauges on the east and west coasts. CGVD28 was adopted in 1935, using classical surveying techniques.

Historically, the definition of this vertical reference system has been separate from the definition of the horizontal reference system.

The current vertical system is a construct of annual survey observations that date back to 1904. Despite great care to minimize potential errors, the network of data points was established piece-meal, with data adjusted locally, resulting in significant regional distortions in published heights that over time, became further exacerbated by crustal motion.

Comparisons of these heights against the most recent geoid model indicate regional distortions of up to one metre (1 M).

While the consistency of heights at a local level are within the sub-centimetre range, the application of new technology, such as GPS, is now impeded by the inability to obtain consistent heights with the current datum.

Further, the current published heights are based on a datum that assumed the Pacific and Atlantic oceans were at the same height. In fact, the water level at Vancouver is generally higher than the water level at Halifax by 40 to 70 cm. This discrepancy causes a national-scale tilt in the published heights creating a potential negative impact to the informational underpinning of a three dimensional *Earth Model* that needs to depict surface height and sub-surface depths based on accurate vertical and horizontal datums.

A new vertical datum is being developed by Natural Resources Canada (NRCan) and is to be introduced to the scientific community in 2013. The datum is to be established using space-borne technologies that model the earth's geoid.

A geoid based-datum will enable Global Positioning System (GPS) height determinations at sub-metre accuracy, and will eliminate the long-term need for a ground control monument network. The datum will produce more accurate elevations over short and long distances, but will require a one-time adjustment of existing elevations throughout the country. In the Province of Alberta, the magnitude of the change ranges from 5 to 75 cm.

In addition, the reference surface (ellipsoid) that is expressed as latitude and longitude measurements must be mathematically transformed onto a two-dimensional (x & y coordinates) surface.

The effort is referred to as map projection; an effort that can utilize numerous coordinate reference systems. The ERCB needs to establish which reference systems it will support.

5. A Modern Reserves Classification System

Governments, Regulators, and Industry require in-the-ground estimates of oil and gas quantities to define policies, forecast supply, assess development opportunities, access capital and manage commercial interests.

Classification schemes for these quantities exist to facilitate consistent understanding and communication between users of this type of information.

In 1978 the term "established reserves" was adopted by the ERCB (Alberta), NEB, CPA (now CAPP), IPAC (now SEPAC), Saskatchewan and Manitoba for reporting estimated quantities of recoverable oil and gas as per a recommendation from the Interprovincial Advisory Committee on Energy (IPACE). The present definition of established reserves (ERCB ST-98, 2011) is:

"Those reserves recoverable under current technology and present and anticipated economic conditions specifically proved by drilling, testing, or production, plus the portion of contiguous recoverable reserves that are interpreted to exist from geological, geophysical or similar information with reasonable certainty".

Present ERCB estimates of Alberta's reserves of oil, gas, bitumen, and coal is based on the 1978 IPACE classification scheme.

The existing classification approach, however, does not explicitly recognize all of Alberta's petroleum resource endowment, particularly its unconventional gas and tight oil. As such large quantities of Alberta's petroleum resources are invisible to the regulatory process and the capital markets.

To bring visibility to all hydrocarbon-based reserves the ERCB introduced a modification in the IPACE reserves classification realm that allowed for the identification of the larger volume of established reserves that are mineable but have not been developed. This practice has been extended to include in-situ bitumen reserves as well.

In 2000, the Society of Petroleum Engineers (SPE), the World Petroleum Council (WPC), the American Association of Petroleum Geologists (AAPG), and the Society of Petroleum Evaluation Engineers (SPEE) began to develop a standardized petroleum-classification system referred to as the Petroleum Resources Management System (PRMS). The purpose of PRMS was to standardize terminology and provide "a consistent approach to estimating petroleum quantities, evaluating projects, and presenting results in a comprehensive classification framework."

Also commencing in same year, the Calgary Chapter of SPEE, in conjunction with the Petroleum Society of the Canadian Institute of Mining and Metallurgy (now part of the SPE), began working on a resource classification scheme of their own. Their definitions were taken up in the Canadian Oil and Gas Evaluation Handbook (COGEH).

The classifications were published in 2003 and became part of Canadian securities regulation.

The COGEH resource and reserves definitions are very similar to PRMS. The Canadian Securities Administrators (CSA), through National Instrument 51-101, set standards for disclosure of oil and gas reserves for companies listed on the Canadian stock exchanges. NI 51-101 references COGEH as the authority for reserves definition and determination methodology for reporting.

Though PRMS and COGEH started as separate but intertwined initiatives, COGEH (2007) Section 5 - Definitions of Reserves and Resources now acknowledges that there "is broad alignment between the COGEH and SPE-PRMS definitions and guidelines, but some minor differences remain".

The most recent consolidated reference version of PRMS was published by SPE in 2007, a publication that the ERCB provided input to.

The COGEH consists of three volumes of work, the third volume published in 2007 has kept "in step" with the PRMS and includes the evaluation of coal bed methane (CBM) and bitumen reserves.

In PRMS, the totality of all the petroleum accumulations found within in a play is termed the "total petroleum initially in place (PIIP)".

Commerciality of a PIIP is defined as the meeting of all essential social, environmental, economic and legal conditions for a production project to proceed; a conceptual construct that aligns with the outcome

monitoring aspects of the *Unconventional Regulatory Framework.*

In addition, a project is commercial if the degree of commitment is such that the accumulation is expected to be developed and placed on production within a reasonable time frame – five years is the recommended benchmark, suggesting a need for information associated with commercial projects to be available for an extended period of time.

Current Status

The ERCB's Strategic Plan has directed the organization to modernization resource appraisal and forecasting methodologies in view of future resource developments in the Province of Alberta.

Many of the unconventional resources are not distinctly recognized in the current IPACE scheme thus impeding the modernization effort.

While the governing acts that mandate the work of the ERCB do not specify any given resource classification system and the term "established reserves" is mentioned only once in the Gas Resources Preservation Act, it is important to establish a common informational approach (terminology and language) in the area of reserves if open collaboration is to be established between the ERCB, Government and Industry.

A proposal for a modified IPACE scheme was presented in November 2009 to the ERCB's Regulatory Committee. Action on the proposal was suspended when the Board released its Strategic Plan and the issues it sought to address were folded into the strategic objective of modernizing resource appraisal and forecasting.

The major problem with a modified IPACE scheme is that is would continue ERCB's use of a different resource classification scheme no longer used by Alberta's oil and gas industry.

While other jurisdictions such as Saskatchewan, British Columbia, the NEB, and CAPP continue to use IPACE, the majority of the global oil and gas community follow PRMS and therefore align with Canadian reserves reporting for securities trading under National Instrument 51-101 and the Canadian Oil and Gas Evaluation Handbook (COGEH).

There is a nuance within PRMS (and COGEH) with respect to classification of commercial undeveloped reserves that will require modification before ERCB's adoption.

In Industry PRMS establishes commerciality implicitly through the assumption that the existence of a project to develop the reserves is indicative of a commercially viable reserve. Undeveloped reserves, by PRMS standards, should be in reach of existing commercial projects and be positioned to provide a reasonable rate of return within a reasonable time horizon, often suggested to be five years.

Historically the ERCB has established estimates of commercial reserves to include accumulations under development plus those deemed commercial by geological analogy, without consideration of the existence of active projects or typical industry time horizons.

That is, if all or part of an accumulation meets the geological and technical criteria for commercial development elsewhere and there is no regulatory reason for commercial production not to proceed, then the ERCB considers it as reserves even if no projects exist.

For this reason, ERCB's reported established reserves for bitumen and coal have been consistently larger than Industry projections.

The ERCB would insist on carrying this practice forward and thereby continue to include quantities of commercial undeveloped reserves for the Province of Alberta.

However, to facilitate widespread international adoption of a single classification system, PRMS sponsors do suggest their definitions and guidelines retain enough flexibility to allow for tailoring by users for particular needs, provided modifications to their guidelines are clearly identified in such cases.

It is envisioned that the ERCB will adopt PRMS and provide appropriate clarification on the ERCB practice on undeveloped reserves.

Moving forward it is anticipated that all petroleum accumulations will be categorized by petroleum system, geological unit, and geological play as well as pool or deposit. These are terms in common use to PRMS and foundational to the new *Unconventional Regulatory Framework.*

6. A Modernized Well Identification Framework

In 1978 a *Unique Well Identifier* (UWI) standard was

developed for the Industry by the Canadian Petroleum Association (now CAPP) and in the same year adopted by the ERCB and Alberta oil and gas operators.

The UWI consists of four (4) components and sixteen (16) characters that support the identification of resource type (conventional oil and gas, oil sands, water source), geographical location (referred to the land survey grid), and well event (drilling and completion.).

In conjunction with the UWI, the ERCB also applies a Well License Number (WLN) to a permitted well activity.

The WLN and the UWI are the two essential identifiers for wells, but have different purposes.

The WLN is shared by all wellbores and completions having a common surface location referred to as a Well Origin (WO).

The UWI is unique to each wellbore and has a geographic component that refers to the location of the bottom of the hole.

The main challenges with the present UWI are:

- The UWI is not a permanent identifier; it is subject to revision;
- The UWI cannot be used reliably to identify all wellbores or completions belonging to a particular WO;
- The UWI cannot identify more than nine (9) events in a well; and
- The UWI does not distinguish between wellbores and completions.

The present approach for creating a well is to have the ERCB issue an initial WLN that has a unique surface location.

The WLN applies only to the original wellbore.

If an original wellbore is abandoned, then re-entered a new WLN is issued. The result can be a single surface location that has multiple WLNs.

As a hydrocarbon based well is expanded through the completion of the original wellbore and subsequent activities that result in additional wellbores (referred to as events) each wellbore is provided a UWI.

Each UWI is initially based on the projected bottomhole location.

However, the actual bottomhole location of each wellbore is only known when a directional survey is made after the wellbore is complete.

If a survey determines a new bottomhole location, the ERCB will change the UWI.

There is a general relationship between a WLN and a UWI that, theoretically, provides stability in well identification. More specifically, generally only a single WLN represents a specific surface location.

Therefore subsequent subsurface activity can be identified through a series of UWIs that is linked to a WLN.

Unfortunately the UWI, as presently defined, cannot handle more than nine (9) events. Today, there are numerous wells that have more than nine events (subsequent subsurface activities resulting in numerous wellbores).

In these situations the ERCB allows the Well Licensee to decide on which wellbores will not be assigned a UWI. These wellbores become, in effect, invisible.

Further, the current identification approach does not provide for clarity around a wellbore and a completion within a wellbore. Today, Industry Operators can (and do) enter the same wellbore and re-complete to new depths or formations.

Generally a new completion creates a new UWI. As such the resulting identification can appear as a new wellbore. Additional information must be sought to determine what has transpired – a new wellbore or a new completion in an existing wellbore.

In addition if the completion is not in a new formation, or (in some areas of Alberta) is for commingled production a new UWI is not issued. In this scenario well activity is not highlighted thus introducing the potential for regulatory oversight to be avoided.

The ERCB, together with Industry, other Canadian and United States Regulators and the Professional Petroleum Data Management Association (PPDM) have begun work on a Global Well Identification Framework (GWIF). The envisioned outcome is a set of metadata standards and a formal glossary that expresses and defines all aspects of a “well”.

It is envisioned that the ERCB would utilize the

GWIF to induce appropriate internal metadata standards and data models to support a well identification that is more aligned with the modern approach to hydrocarbon well development.

7. Open Data and the Informational World

As the oil and gas industry becomes ever the more omnipresent in the day to day lives of the Province of Alberta's constituents desire to know more about the Industry and the ERCB's position on oil and gas related matters will increase.

The desire to "know more" resonates in the ERCB's general mandate to inform and advice internal and external stakeholders; a mandate that is further entrenched into the day-to-day operations of the organization through the *Oil and Gas Conservation Regulations* that establishes in Section 12.150 "The Board, at any time, shall make available to the public, from records, reports or information submitted to or acquired by it" a wide variety of oil and gas related information.

At the same time, the Government of Alberta is striving to provide public access to electronic data in open, accessible and machine readable formats, with minimal restrictions.

Together, the Government and its Regulator (i.e. the ERCB) are faced with a trend towards *Open Data* that is further facilitated by new Internet technologies, mobile applications and the emergence of data catalogues and portals.

To this end, an *Open Data Program* has been initiated by the Government of Alberta to coordinate open data efforts across the Government and its agencies.

While the Government of Alberta has not publicly announced an open data initiative, several initiatives using open data techniques have commenced:

- The Oil Sands Information Portal – established November, 2011
- GeoDiscover Alberta Portal – established in January, 2011.

The ERCB is involved with both initiatives and through the Government of Alberta's CIO Council is linked to the Open Data Committee.

The Open Data Committee and its Data Standards/Metadata Working Committee have begun to explore the Government of Alberta's utilization of the "Dublin Core Metadata Element Set" (ANSI/NISO Z39.85-2007) and the "DCMI Metadata Terms" issued 2008-01-14 as the foundation for developing metadata standards.

Within the Government of Alberta the Environment and Sustainable Resources Ministry and Energy Ministry have established metadata standards – the first using the "Dublin Core", the second a geospatial metadata standard based on the Federal Geographic Data Committee standards (CSDGM FGDC-STD-001-1998 .

Further, the Professional Petroleum Data Management Association (PPDM) includes metadata within its taxonomy dimensions. Its work in this area is also based on the "Dublin Core".

The ERCB has initiated a limited amount of metadata development work specific to geospatial metadata to achieve alignment with the standards utilized by the GeoDiscover Alberta Portal.

Internally, the ERCB has begun to utilize open data concepts through the creation of metadata within its current document management deployment.

It is envisioned that internal, inter-government and inter-industry informational exchanges will continue and expand forcing the ERCB to "take up" the numerous metadata standards that will underpin the open exchange of information and data in a world driven by an insatiable thirst for knowing and being informed.

8. Confidentiality in an Open Data Environment

What may seemingly be a contradiction to the *Open Data* movement, is the ERCB's regulated requirement to share records, reports and information with the public while being constrained by the need to withhold some informational components that are deemed confidential.

The establishment of confidentiality is established by Industry Operators as a pre-spud classification of risk.

The established well classification is reported to the ERCB at the time an application for the creation of a well is submitted. ERCB personnel make no attempt to review or discuss the classification set by an Industry Operator.

The classification of the potential well is based on the historic *Lahee Classification* approach that was first

purposed in 1944 by Frederic H. Lahee.

Over the course of many decades since, the well classifications have been modified and/or expanded to reflect technological changes in the oil and gas industry.

Today, the ERCB's extended *Lahee Classification* includes risk definitions for:

- New Field Wildcat (NFW)
- New Pool Wildcat (NPW)
- Deeper Pool Test (DPT)
- Outpost (OUT)
- Development (DEV)
- Re-entry (REN)
- Development Service Well (DSW)
- Evaluation Oil Sands (OV)
- Test Hole (TH)
- Experimental (EX)
- Other (OTH)

Of these well classifications the following are identified with "confidential status" when a new well is applied for:

- NFW, NPW, TH, OV, EX

The remaining well classifications identified below may or may not be deemed "confidential":

- OUT, DPT, DEV, DSW, REN

A "confidential status" could be established on these well types if the proposed well:

- Penetrates all zones within an area that is outside of a defined pool
- Terminates in a defined pool designated as "confidential".
- Passes through a defined pool designated as "non-confidential" and terminates below the defined pool.

The defined pools are "board ordered", meaning that the pool locations, associated boundaries and confidential/non-confidential designations are man-made and can (and do) change.

Given these conditions, it is evident that accurate informational connections between proposed wells (identified by WLN and UWIs – see Section 6 – A Modernized Well Identification Framework), defined pools and a table of formations must be maintained at all times.

In addition the accuracy of wells and supporting facilities remains important to ensure that wells are either accessing the correct pools or are clearly outside the boundaries of any designated pools and are at depths that support its confidential or non-confidential status.

The need for accuracy brings to the fore an important organizational requirement for an accurate three dimensional digital model (See Section 4 – An Earth Model) that can visually depict the relationship between surface and subsurface locations.

9. Information Management and Master Data Management

The ERCB, as an organization, is a rich informational environment that continues to receive and disseminate significant amounts of information in a variety of formats.

It is envisioned that the volume, scope and depth of its informational assets will continue to expand in the foreseeable future.

The ERCB is subjected to confidentiality rules (See Section 8 – Confidentiality in an Open Data Environment) that require the organization to hold back (from public consumption) data components on some oil and gas activities for a specified period of time (generally one year).

Further, the ERCB has a formal Records Management approach (and supporting policies and procedures) driven by the following enabling legislation:

- Administration Procedures Act
- Electronic Transactions Act
- Freedom of Information and Protection of Privacy Act
- Government Emergency Planning Act
- Government Organization Act
- Alberta Records Management Regulation (AR 224/2001)
- Historical Resources Act

In addition the ERCB is required to be compliant with:

- The Alberta Evidence Act
- Alberta Court Practices
- Sedona Canada Principles (discovery on electronically stored information)

In consideration of the continued evolution of the

ERCB's informational assets and the increasing desire by Government, Industry and the General Public to gain access to more information, the ERCB has commenced development of an Information Management Strategy for the purpose of bringing together all of the varying information management responsibilities under one formal policies-based construct.

It is envisioned that the development of ERCB's Information Strategy will be lead by a cross-functional IM Committee. The Committee's primary purpose is to delivery, into the ERCB, formal IM Governance (expressed as policies, standards, controls and metrics).

More specifically, the envisioned Information Management approach will incorporate a three-tiered IM Policy Framework that will have the following policies:

- Information Management Policy
- Data Management Policy
- Electronic Document Management Policy
- Records Management Policy
- Messaging Management Policy
- Legal Hold Policy
- Digital Imaging Policy
- Information Disposal Policy

The IM Committee recognizes the need for the following:

- Alignment with an Information Architecture
- The identification of information and data owners
- The initiation of metadata and master data initiatives

10. International Standard Data Models

General

The ERCB has embarked upon an effort to identify appropriate internationally standard data models for the organization and management of envisioned geological, earth resources (subsurface) and surface based data objects.

Initial research has found a wide range of models – some being a part of the suite of international standards from the ANSI or ISO. Others developed and used by government agencies such as the United States Environmental Protection Agency or Environment Canada.

In addition, varying data models have been developed by international consortia such as the Commission for the Management and Application of Geoscience Information (CGI) and industry-specific standards such as the Public Petroleum Data Model from the Professional Petroleum Data Management Association (PPDM).

Almost all data models are focused on a particular domain (surface, subsurface, geology, etc.).

Initiatives like EarthCube and the Open Model Interchange have begun in support of a common interchange for varying data models and the development of standard interdisciplinary data models.

Environmental Data Standards

In 2009 the Open Geospatial Consortium (OGC) jointly published a standard for observations and measurements with the International Standards Organization (ISO) and was described as follows: "This standard specifies an XML implementation for the OGC and ISO Observations and Measurements (O&M) conceptual model (OGC Observations and Measurements v2.0 also published as ISO/DIS 19156), including a schema for Sampling Features."

Around the same time, a team of researchers published an Observations Data Model (ODM) for use with relational database technology (Horsburgh et al., 2008).

Their model was developed with consideration for OGC's best practices for observations and measurements. The ODM is used by the Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUASHI) for their Hydrologic Information System (CUASHI-HIS).

Beran et al., 2008, published a data model for environmental observations that addresses some of the deficiencies of Horsburgh et al. (2008) model and arose from the collaborative work of Microsoft and RENCI (Renaissance Computing Institute) on environmental data modeling.

RENCI's Sensor Data Bus project produced a complete data model for environmental data, referred to as EnviroDB.

Alberta Environment and Water (AEW) have been assembling environmental data into a data warehouse over the last five (5) years. The warehouse comprises a transactional relational database, and a set of data

marts that are denormalized extracts of the data warehouse, tailored for specific application and business uses.

Economic and Social Data Standards

Generally economic and social data exist as statistics that provide insight into such aspects as the Gross Domestic Product (GDP) for a country or employment rate for a Province.

One standard has been widely adopted for the management of statistical data, the Statistical Data and Metadata Exchange Standard (SDMX) ISO-TS 17369.

SDMX is a standard design template that describes statistical data and normalizes their exchange. It is designed to reduce the development, maintenance and operational costs within an organization through the following:

- Logical unification of data stored inside and across national and international organizations through defining the common logical data model, harmonization of the statistical metadata (e.g. code lists) and use of prescribed objects (e.g., schemes, data structure definitions)
- Application of common logical model and related standards effects in reduction of diversity among statistical data production processes and related business process.
- Sharing of standard, generic software and IT infrastructures allowing automatic production, processing and exchange of data and metadata files among statistical organizations.
- Use of standard software and standard data model allows machine to machine communication what in turn minimizes manual interventions and human errors.
- Discovery and unification of distributed data shaped according to standard logical model."

In addition, SDMX is the standard preferred by the UN Statistical Commission for the exchange and sharing of data and metadata.

Geological Data Standards

There are a number of geological data standards that could address the data management needs of the ERCB, including the Public Petroleum Data Model from the Professional Petroleum Data Management Association (PPDM).

The PPDM data model was developed by an industry consortium and includes schema for most hydrocarbon exploration and production data.

However, the Alberta Geological Survey within the ERCB, generate and maintain special geological datasets including geochemistry, hydrogeology and rock property data – schemas that do not exist in the PPDM data model.

Natural Resources Canada (NRCan) maintains a National Rock Properties Database (NRPD) using a data model that includes rock property data from wireline geophysics and laboratory analysis of borehole and surface samples.

NRCan also publishes geochemical data through their Geoscience Data Repository.

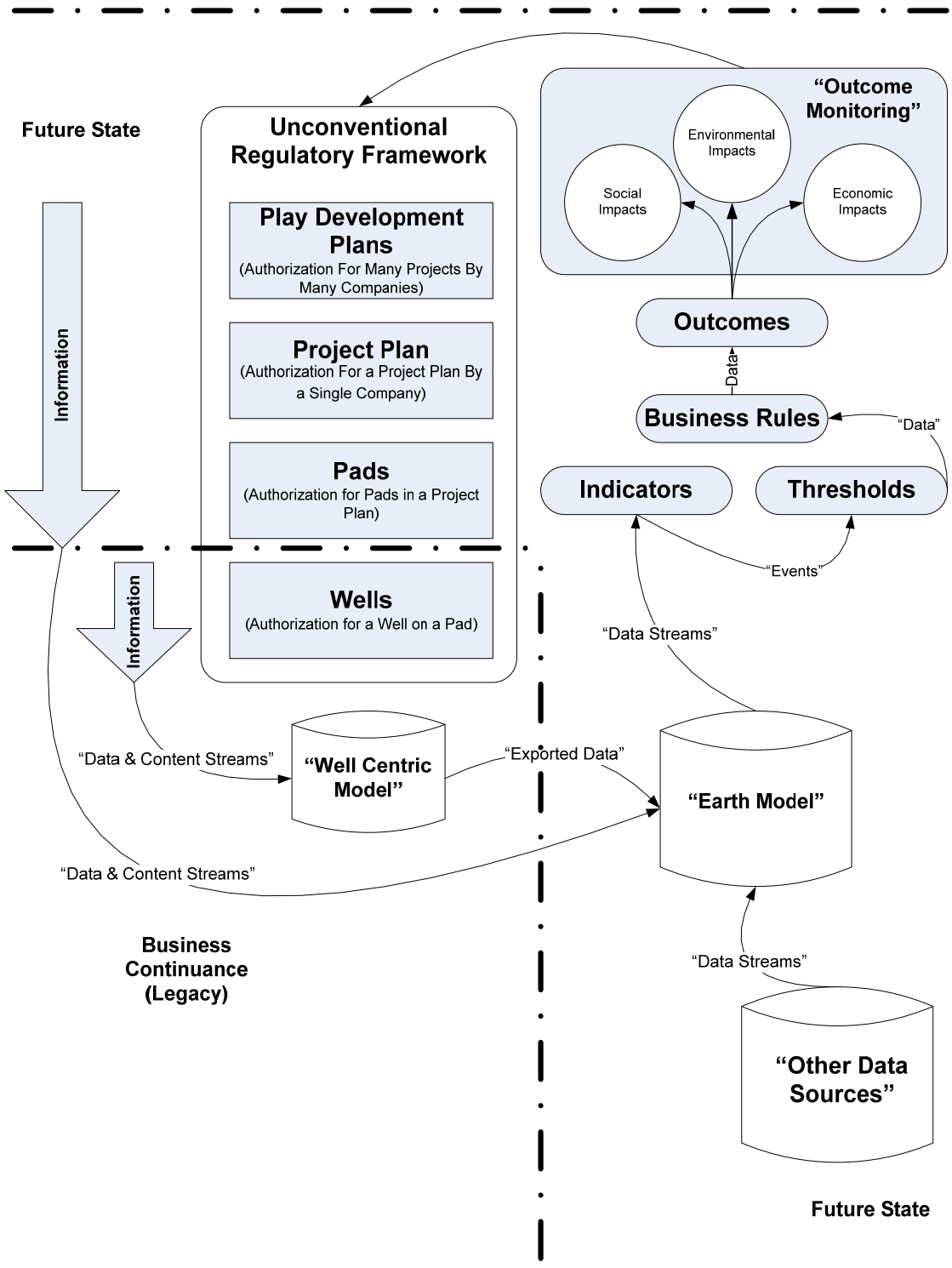
The ERCB will continue to monitor these efforts as it develops its Information Architecture.

11. Conclusions

Based on the issues and directions identified in this whitepaper it is concluded that the ERCB's informational environment will:

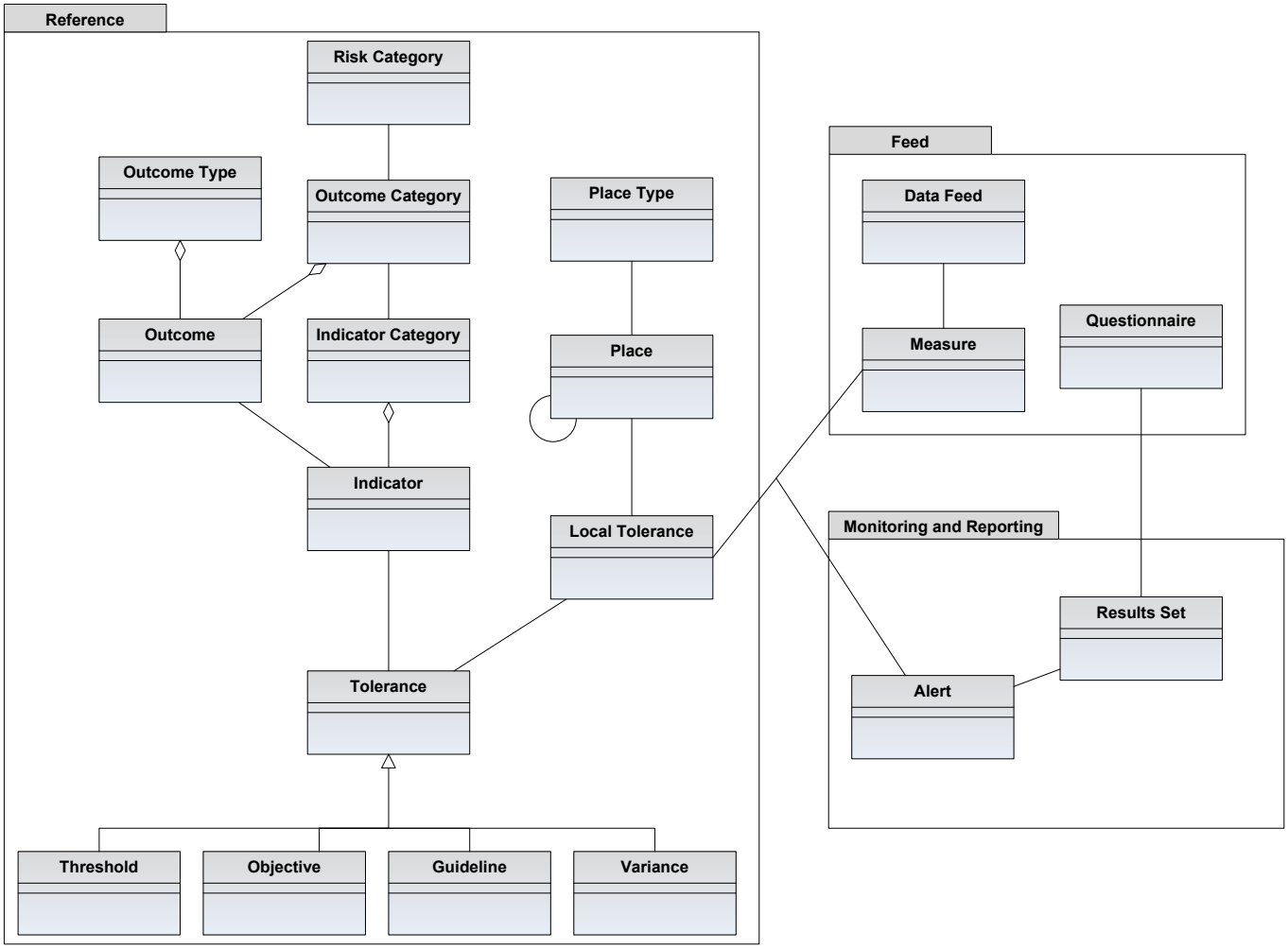
- Be required to support the development and dissemination of new informational products such as outcomes, indicators and thresholds
- Include and potentially centre on a three dimensional earth model based on a single geodetic standard that provides the accurate visual depiction of geologically based surface conditions, sub-surface activities and a near real-time relationship of pads, wells and digital dashboards that provide outcome reporting
- Support a modern reserves classification approach that is more aligned with the evolving resource endowment within the Province of Alberta
- Be impacted by changing industry-based standards associated with, but not limited to, well identification, reserves classification, risk classifications and the establishment of confidentiality pursuant to a revised Lahee Classification approach
- Be more open and transparent in support of ERCB Branch, Industry, Government, and General Public informational exchanges
- Utilize more standard data models
- Continuously change in a more rapid way

Appendix A



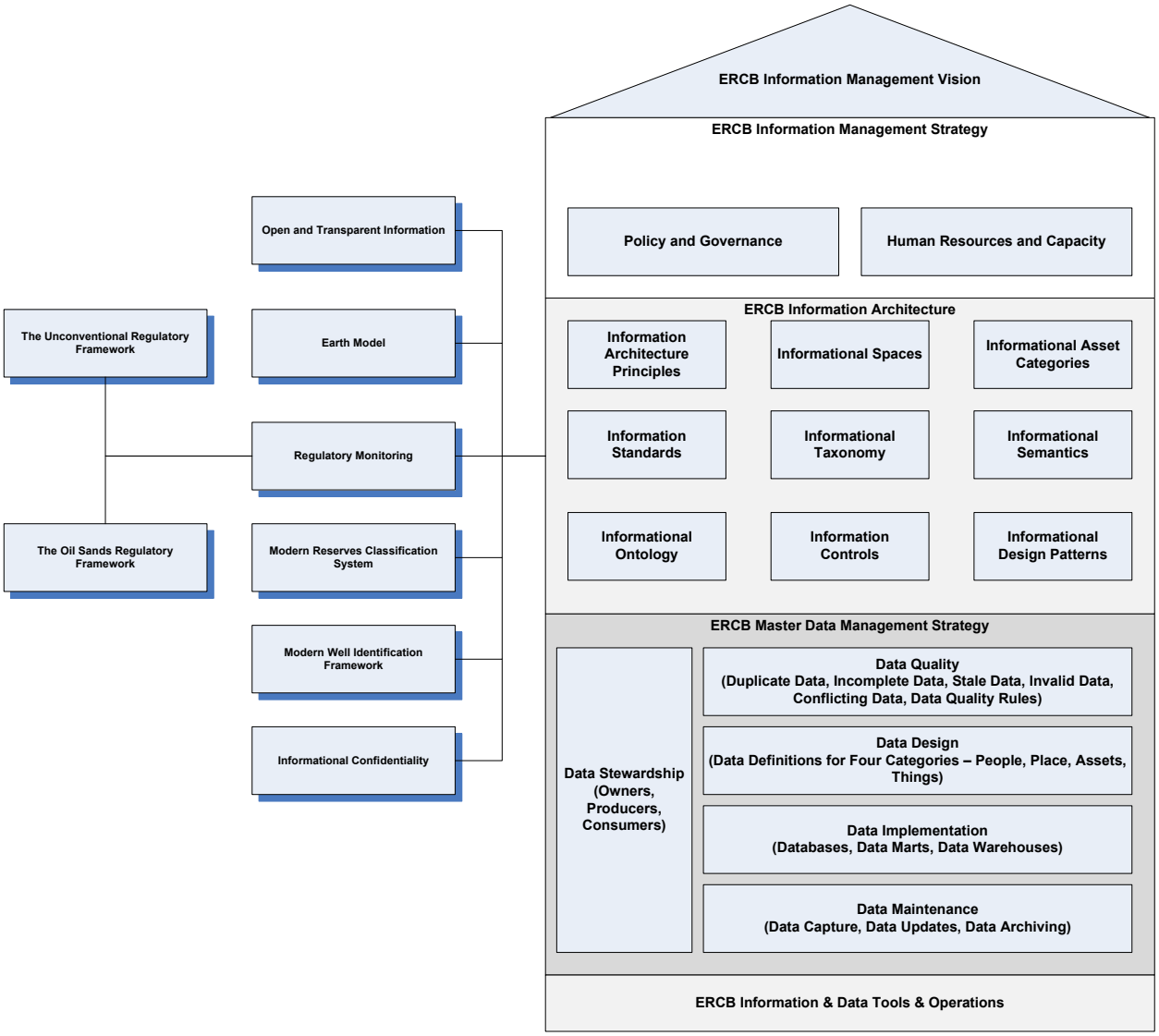
General URF Informational Model

Appendix B



Outcomes Monitoring Conceptual Data Model

Appendix C



Information Management, Information Architecture & Master Data Management Alignment Model