



**PROPYLON**

[www.propylon.com](http://www.propylon.com)



## Statement of Work for the Indiana General Assembly

Submitted By	Sean McGrath
Phone	775 297 6310
Website	<a href="http://www.propylon.com/">http://www.propylon.com/</a>
Email	<a href="mailto:sean.mcgrath@propylon.com">sean.mcgrath@propylon.com</a>
Version	Final with editorial corrections

**Copyright © Propylon**

Company confidential, not for redistribution without authorization

Propylon, Suite 201, Technology Business Center, 2029 Becker Drive, Lawrence, KS 66047. Tel: 785 856 2675

# Table of Contents

---

1	Executive Summary.....	<u>4</u>
1.1	Background to Statement of Work .....	<u>4</u>
1.2	Summary of Findings .....	<u>4</u>
1.3	Summary of Recommendations .....	<u>6</u>
2	Introduction: Approach and Scope.....	<u>8</u>
3	Drivers and Success Metrics.....	<u>9</u>
3.1	Web-based Interaction with the Legislature.....	<u>9</u>
3.2	Paper Reduction and Paperless Committees.....	<u>9</u>
3.3	Control of the Information Published to the Website.....	<u>9</u>
3.4	Information Consistency.....	<u>10</u>
3.5	Business Rule Flexibility.....	<u>10</u>
3.6	Legacy Technologies.....	<u>10</u>
4	The Current Systems Architecture.....	<u>11</u>
4.1	Overview.....	<u>11</u>
4.2	Information Silos.....	<u>12</u>
4.3	Information Consistency.....	<u>13</u>
4.4	Information Latency.....	<u>14</u>
4.5	Compare-centric and Proofing-centric Workflows.....	<u>14</u>
4.6	Audit Trail.....	<u>14</u>
4.7	Hard-wired Business Logic and Outputs.....	<u>15</u>
4.8	Notifications.....	<u>15</u>
4.9	Miscellaneous Challenges.....	<u>16</u>
5	Root Cause Analysis.....	<u>18</u>
5.1	Enterprise Information Model.....	<u>18</u>
5.2	Normative electronic documents.....	<u>19</u>
5.3	Amendments in Context.....	<u>20</u>
5.4	Temporal, Versioning Database.....	<u>20</u>
5.5	Flexible business logic.....	<u>20</u>
5.6	Robust Integration Points and Normative Notification Channels.....	<u>21</u>
6	Proposed Architecture.....	<u>22</u>
6.1	Introduction.....	<u>22</u>
6.2	Recommended Features and Design Patterns.....	<u>23</u>
6.2.1	Information Governance.....	<u>23</u>
6.2.2	Fully Automatic Audit Trail of all workflow actions.....	<u>25</u>
6.2.3	Robust content citation and content authentication.....	<u>26</u>
6.2.4	Machine readable data feeds.....	<u>26</u>
6.2.5	Flexibility in Business Logic.....	<u>26</u>
6.2.6	Robust Notifications Framework.....	<u>26</u>
6.2.7	Bring Your Own Device Support.....	<u>27</u>
6.2.8	Load Balancing.....	<u>27</u>
6.2.9	Amendments in Context Capability.....	<u>27</u>

- 6.3 Recommendations for short term consideration.....[27](#)
- 6.4 Longer Term Recommendations.....[28](#)
- 6.5 Recommended Phasing.....[29](#)
  - 6.5.1 Phase 1.1: Website.....[29](#)
  - 6.5.2 Phase 1.2: Budget Bill and Member Services.....[30](#)
  - 6.5.3 Phase 2.1: Non-Budget Bills and Statute Management.....[30](#)
  - 6.5.4 Phase 2.2: Fully Paperless Legislature, Indiana Administrative Code Integration.....[30](#)
- 7 Appendices.....[32](#)
  - 7.1 Appendix: Current Technology Stacks - COTS.....[32](#)
  - 7.2 Appendix: Current Technology Stacks – Custom Code.....[33](#)
  - 7.3 Appendix: Distinct Sources of Information about a Bill/Resolution.....[33](#)
- 8 Statement of Work – Project Artifacts.....[36](#)
  - 8.1 Statement of Work Deliverable Documents.....[36](#)
  - 8.2 Project Artifacts.....[36](#)
- About Propylon .....[38](#)

# 1 Executive Summary

---

Propylon would like to thank the staff of the Indiana General Assembly (hereafter referred to as the Legislature) that contributed to this Statement of Work report. The help we received during this project from each of the divisions was very detailed and contributed greatly to the findings and recommendations of this report. The Legislature has built IT systems that, with the hard work of staff, get the job done. However as technology changes and demands on the Legislature, for the information it provides, increases the current systems need major upgrading/ replacement to allow staff to continue to deliver a service that they have been proud to deliver for years.

## 1.1 Background to Statement of Work

The Legislature recently conducted a successful pilot project involving the use of Apple iPads in the House Education and the Senate State Tax and Fiscal Policy Committees. The Legislature is interested in expanding this pilot into full production across all committees and more generally, using the benefits of technology to make legislative data available real-time on the web. This will help members, staff, agencies, the public and businesses to interact with the Legislature. This will also introduce efficiencies to the legislative processes and allow for the provision of new services.

The Legislature recognizes that paperless initiatives such as the iPads will create extra requirements on the existing IT systems and infrastructure. The Legislature is therefore conducting an end-to-end analysis of the existing systems/processes/infrastructure with a view to not only enabling the iPads roll-out, but also addressing some other issues with the current systems that have been identified by both staff and members.

Following discussions and demonstrations, Propylon was engaged to examine the existing systems, review where issues may exist and make recommendations for how the applications might be changed to better deliver the business goals of the Legislature. The Propylon OCDM (Output Centric Design Methodology) process was used to create a well defined scope and performance metrics for project success. The work was carried out over four weeks with a mixture of on-site meetings and off-site documentation work.

The recommendations made in this report are based on the information made available to Propylon during this process. All artifacts gathered have been stored in a data/document repository at the Legislature with shared access between the Legislature and Propylon.

## 1.2 Summary of Findings

During the analysis our team was able to identify a number of requirements for provision of new services as well as identify issues in the systems implemented today. The following summary list is a mix of current issues and new requirements:

- New requirements for information provision through technologies such as tablets/ ipads is putting an increasing strain on the current systems and IT infrastructure.
- Legislators are requesting more information to be available to the public in real-time as events are happening in the Legislature.
- Staff are restricted by the current systems in place as to the services they can provide.
- The IT staff and technicians in the divisions are consumed by maintaining the current systems. They do not have the capacity to provide additional services.
- A dependency has grown on key individuals which is a risk in the case where these staff are not available.
- The current IT systems need to be significantly upgraded/ replaced if the Legislature is going to be able to provide the new services required.

- The Legislature needs complete control over what is published to the legislative website and when it is published. The Legislature currently does not have this and it causes many issues.
- The back-end business processes need to be reviewed and succession planning needs to be implemented.
- An IT architecture that manages the digital workflow of the legislative process, and all of the data in it, needs to be implemented in a way that provides easy access for members, staff, the public, businesses and agencies. This architecture needs to respect the privilege and procedural rules of the Legislature.
- The current systems technologies are difficult to find skills to support. This is a risk to IT service provision to the Legislature in the future.
- External requirements on the Legislature to produce and make available all information in a digital format, that is easy to search and navigate, are growing.
- The provision of member interfaces that are customized to each member is a growing requirement. These types of tools will provide major benefits to members and their staff to organize and make accessible the legislative information with a particular focus on the individual member.
- There is a requirement to reduce the amount of paper throughout the legislative process. This has a significant impact on the provenance requirements of digital material.
- Critical pieces of information are being stored in non-centralized areas that do not have redundancy.
- A single consistent view of all legislative data is required across organizational boundaries.

To meet the current challenges and provide the new services required, the first step that needs to be implemented is an information architecture that traces how all information moves through the systems and defines discrete points where that information should be published and in what format(s). At the heart of the proposed approach is a Legislative Information Architecture that is independent of any application, platform or vendor. The architecture encompasses all aspects of the legislative process end-to-end. We believe that this architecture will allow the Legislature to move forward with confidence, not only based on today's known initiatives such as the IGAViewer/ iPads but new initiatives that will doubtlessly come up in the future.

The systems that are in use today are not well integrated bringing the following risks:

- The current processes rely on individuals who understand what information needs to be acted on and when, to ensure the success of the legislative process. A number of these processes should be completed systematically allowing for automatic alerts to all necessary parties as these steps are completed.
- Where systems are not integrated, it can lead to inconsistencies in information between systems (the digest is an example of this, where it can be out of synch between the chamber and LSA).
- The automatic publication of information to the intra/ internet is problematic as the rules for publication are often dependent on events happening across systems.
- The lack of integration can cause delays in the legislative processing as the staff are forced to work around the fact that the systems are not automatically sharing information.
- Additional resource requirement for re-keying and moving of data limits the amount of time staff can spend on value added services (such as addition of meta-data for improved searching).

Implementing an overall information architecture that respects privilege and procedure while ensuring the benefits of an integrated system are realized is critical for the Legislature to solve the current issues in IT systems. This will also provide the new services required by the Legislature's stakeholders, which are:- the members, staff, public, businesses, and agencies.

### 1.3 Summary of Recommendations

The recommendations set out in this section relates not only to the IT systems but also to the processes that these IT systems are being used to support. It is important for the Legislature to realize that the benefits of any new system will involve changes to existing business processes. If this is not done then old restrictions will be built into any new system preventing the benefits being delivered.

IT system Recommendations:

- Implement an Enterprise Information Architecture that covers how information is managed in systems across the Legislature.
- In the Enterprise Information Architecture ensure that the information items are treated as assets that are entered once and reused where possible.
- Take complete control over the website publishing. The Legislature should have the ability to control what is published, when it is published and have the ability to roll-back changes when something is published inadvertently.
- Provide a single source for bill status information across the Legislature. Remove tracking systems that are replicating information which can lead to consistency issues. Ensure that multiple interfaces can be built to the bill status information to allow the consumption of the data from a single consistent feed.
- Put all source code for applications in the divisions under source control to ensure that quality is built into the software and is maintained as the software changes over time. This will also make the applications easier to support.
- All legislative information should be stored in a central information system with rules implemented to protect privilege and procedure. Centralizing the legislative information should also simplify supporting the applications and allow for cross-training of staff to reduce risk.
- Remove unnecessary information silos while protecting procedure and quality, and supporting security and confidentiality considerations.
- Ensure all legislative data is stored on a centralized system that has fail-over and redundancy in place.
- Use reliable system notifications (persistent queues) to support the business process. This will ensure that notifications are persisted even when there are issues and will allow for more automation of system tasks.
- Provide the data feeds of legislative information in machine readable format (recommend a combination of REST, JSON and XML) for web based standards.
- Ensure that the new IT platform has interfaces that will support the provision of new services in the future.

There are also a number of project recommendations for the Legislature on items outside of the system architecture:

- Establish a Sponsor Group with stakeholders representing each of the groups investing in the new system. This should be a leadership group that sets high level goals for the project.
- Appoint a project manager that has responsibility of delivery of a new project. This should be completed as quickly as possible as the amount of planning involved is significant.
- The OCR processes need to be more tightly integrated with OBDAR and OFMA. We recommend appointing a business process change analyst/ manager that will work with the project manager to carry out this task.
- Establish defined protocols for interaction of LSA and Caucus attorneys to help reduce issues such as statute conflicts and re-keying. A new system implementation should make this interaction systematic where possible.

- Cross-train staff where dependencies exist on individuals today. This is an important risk mitigation strategy.

## 2 Introduction: Approach and Scope

---

To complete the Statement of Work we used the Propylon OCDM (Output Centric Design Methodology) as the process for ensuring the exercise was carried out in a structured fashion and covered scope correctly. The Propylon OCDM process has been created specifically to analyze legislative projects and provide a framework for our clients to get a deep understanding of the scope of the current systems in terms of what these systems must produce, when and where they must be produced, why they are produced and how they are produced. The starting point for legislative analysis is a discovery of the “as is” situation in terms of systems and processes. In order to get an overview of the Legislature as a whole, we start with a focus on the outputs produced. These outputs fall into two categories:

- Outputs produced for consumption internally e.g. amend/repeal reports, work management reports, confidential drafts, etc.
- Outputs produced for consumption both internally and externally e.g. bills, journals, calendars, statute books, bill status, website, etc.

In parallel to gathering and categorizing of outputs, the Propylon analysts ran a number of workshops with the Legislative staff to understand the processes in use today. The findings of these workshops were documented and used as a basis for the findings in this report and making recommendations on how to move forward.

Our technical staff had a number of meetings with the Legislative IT staff to fully understand the technologies currently being used to support the legislative process. This involved detailed question and answer sessions as well as demonstrations to get a full appreciation of how the existing technologies are used to support the processes and identify current issues.

The accumulated understanding of the outputs is then combined with an understanding of the current change drivers (business goals, pain points, success metrics) to synthesize a “to be” picture. The “as is” and “to be” material is then used to produce a phased implementation plan that addresses the critical business continuity imperatives that all Legislatures operate under. Finally, a set of recommendations is included covering a wide range of areas that touch on the IT environment of the Legislature.

In order to frame the statement of work project, a high level scope determination was made at the outset. All aspects of the operation of the Legislative Services Agency (LSA), including Indiana Administrative Code, have been analyzed with the following exceptions:

- The Page Office System
- Constituent Management Functions
- Human Resource functions
- The data warehousing project related to Property Tax (OFMA)
- The redistricting/reapportionment/precincting functions (OCD)

## 3 Drivers and Success Metrics

---

In our analysis, a number of drivers and success metrics for change to the Legislature's current IT systems have been identified. These are outlined in the sections below. The list is not exhaustive for capturing each improvement or change that was suggested during the analysis but looks at rolling these up under summary headings.

### 3.1 Web-based Interaction with the Legislature

The requirements on the Legislature have been growing as the world of technology changes. The requests that LSA receive for new services is growing and the current systems are struggling to cope with the requests. The current systems are set-up to provide services as they were defined 14 years ago. The new requirements mean that each new service requirement is consuming more of the scarce human resources in the Legislature to the point where no new service can be added without effecting existing ones. The advent of tablets/ iPads is leading to even more demands as members, staff, the public, businesses and agencies expect to consume the information from the Legislature on-line as it is happening in real-time.

The ability for the Legislature to produce updates for what is happening in the legislative session in real-time (or as policy allows for it) is a key measure for the Legislature. The technology should not restrict the policy decision.

### 3.2 Paper Reduction and Paperless Committees

The iPad pilot in the House Education and Senate State Tax and Fiscal Policy Committees was very successful and there is a desire to move to full production mode. The recent report on paper usage in the Legislature suggests that a single bill currently results in 11,000 sheets of paper. The cost savings to the Legislature have been identified as significant. The Legislature wants to implement paperless committees in an efficient way that makes it easy for committee staff to publish committee information to the web.

The publishing of committee information to committee apps and websites should not be restricted to any one committee. The tools should be provided to all participating committees. The information architecture is a key enabler of ensuring this can be done in efficient way for each committee. The provision of data feeds is also important to ensure that multiple device types are supported for consuming the data. The number of these devices is constantly evolving.

### 3.3 Control of the Information Published to the Website

The staff appreciate that most people outside the Legislature depend on the legislative website as a source of information to follow what is happening during the legislative session. At some points in the session there is currently a 30 to 60 minute delay in updating the website which impacts both documents (i.e. bills) and bill status. This is a cause of frustration among staff and members of the public.

Some internal systems such as Legisoft are updated at periodic intervals rather than immediately. At times in the legislative session when things are moving fast, this delay is particularly problematic.

The Legislature needs to have the ability to control what is published to the web, when it is published, and update or remove an item where appropriate.

### **3.4 Information Consistency**

There are a large number of copies/variations/views available of critical business items such as:

- bill text
- fiscal note text
- bill sponsors
- bill digests
- bill status etc.

When faced with a discrepancy (for example, between Legisoft and the Website), determining which view to consider accurate and determining how the discrepancy came about, can be problematic. This is frustrating for staff that are under time pressure to get the job done during session as it is for members when they need quick access for decision making.

The Enterprise Information Architecture should ensure that information in multiple applications is consistent and where possible read from the one source.

### **3.5 Business Rule Flexibility**

The systems implemented today have hard coded rules that are being constantly worked around. These work-arounds are time consuming and only known to a very small number of staff. This introduces the risk of delays in the legislative process. Examples of the rules implemented include limitations in the amendment numbers and difficulties in switching between special and regular sessions. As the business changes in the Legislature the number of these work-arounds grows and consumes more staff time.

The staff in the Legislature require a system that will allow flexibility when dealing with events that are not frequent and are exception to normal work flows. The system should have tools that allow non-programming staff to configure the work where needed.

### **3.6 Legacy Technologies**

There is concern about the status of some of the technologies that are foundational to the current systems.

- Visual Foxpro
- Borland Delphi
- Corel WordPerfect Macro Language
- Folio Views
- Apache Cocoon

The Legislature needs to be in a position where the system implemented is built on a technology stack that has a greater pool of people with the IT skills to support it.

# 4 The Current Systems Architecture

## 4.1 Overview

Many of the current IT systems in the Legislature were implemented approximately 10 years ago. The core systems in place today took approximately two years to build. Since the systems were implemented, modifications, extensions and workarounds have been put in place to supplement the original systems.

The current situation can be characterized as stable but with a number of issues that require a lot of staff time to address. In other words, the systems currently work but there is a limit to how well they can work because of the systemic issues present. The problem areas have been managed through the use of process-based workarounds such as well tested “break-fix” procedures combined with significant hard work and dedication on the part of key individuals in the Legislature.

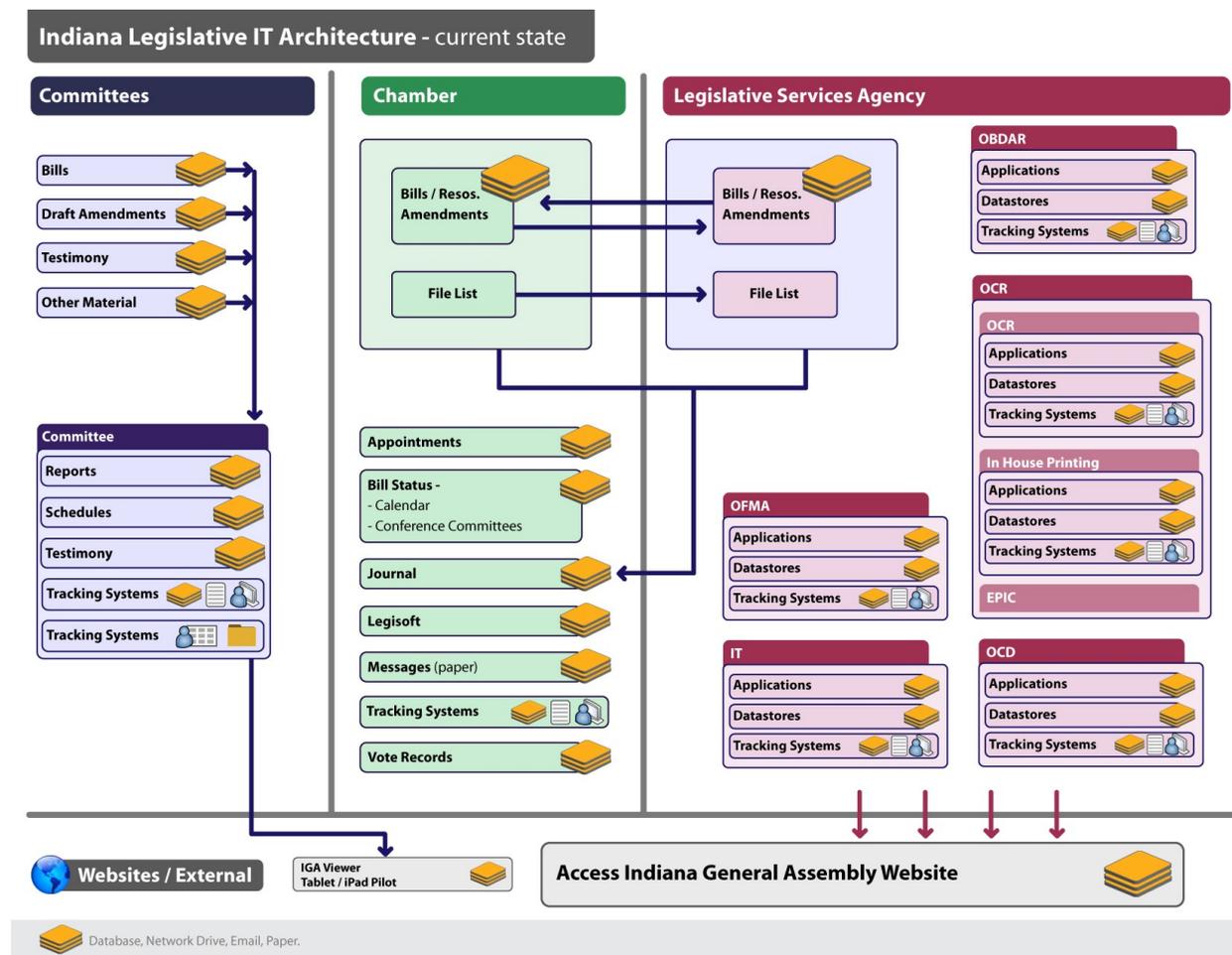


Figure 4.1: Current IT Architecture

The primary issues we have found are discussed below under the following headings:

- Information Silos
- Information Consistency
- Information Latency
- Compare-centric and proofing-centric workflows
- Audit Trail
- Hard-wired Business Logic and Outputs
- Notifications

## 4.2 Information Silos

The Bill Drafting system and the Bill Status system are two of the main epicenters of IT in the Legislature<sup>1</sup>. Although there are some integrations between the two, many of the current challenges can be traced back to these two systems, their functionality and their interaction with the main organizational divisions: OBDAR, OCR/IHP, OFMA and Chambers/Committees.

The two systems are integrated to some degree by code internal to the systems themselves but much of the practical functionality of the Legislature's current systems can be traced to "glue" systems that intermediate between the business processes and the two back-end systems, often using shared network drives as the data store.

The "glue" takes a number of forms:

- A large set of WordPerfect macros ranging from small/simple to very large and complex (approximately 500+ between OBDAR, OCR, OFMA). Many of these are mission critical.
- A set of Delphi and Foxpro Applications (Used in Bill Status, Bill Drafting and in batch processing Robots etc.)
- A set of internal reports and internal cached data formats (e.g. Source Sheets, Last Action)
- Key knowledge workers who know all the workarounds to ensure the front office is not effected
- Data re-keying and manual data moving
- Ad-hoc tracking/management spreadsheets, documents and small databases (e.g. Excel and Microsoft Access)

There are problematic "lock-step" time dependencies between many of the component pieces and the glue. Many batch processes (the robots) trigger behavior based on time-stamps of files. When the robots need to be restarted, problems occur getting caught back up to date. There can be missing/duplicated data transfers, missing/duplicated notifications to other robots which cause negative event cascades to later parts of the workflow e.g. files on the website. These issues take up a large amount of time from technical staff in the IT department as well as a few key staff in the divisions of LSA. Some of these fixes are dependent on knowledge of individuals which poses a risk to the legislative process. Should these staff members be unable to attend work for any reason then the issues will take longer to fix and could cause real disruption to legislative business.

---

1 It is important to note from the outset that the Bill Drafting System is, in fact, a lot more than just a Bill Drafting system e.g. statute management and that Bill Status is also a lot more than just Bill Status e.g. Committees/Calendars

### 4.3 Information Consistency

The current legislative workflows feature a large number of copies, conversions and variations of key assets such as bills, amendments, statute sections, etc. This is the result of many silos of information existing to manage the data as there is not one single IT architecture providing staff the ability to easily manage and version documents. In most workflow steps (from request through to enrolled bill) there are multiple copies made of WordPerfect files, PDF files, HTML files and Folio files. There are multiple entries made in database tables and tracking spreadsheets. A rough estimate is that at any one point in a typical 8 stage bill workflow, there are over 30 possibly different “views” available into the bill itself. Moreover, this causes a significant amount of time being spent in establishing which “view” should be considered accurate in the event that a discrepancy is identified.

There is very little content versioning at present. Due to the nature of the IT architecture it is easy to over-write important files accidentally. It is also difficult to keep the correct versions of related documents together, for example, the text of a fiscal note is the correct one for a given bill text. An LS can be modified (for example to make a correction) but there is no way to visually differentiate the LS as originally released by ODBAR and the LS post edit. Another example of the impact of this in a different area is when a motion is expunged in journals which can result in two copies of a journal for the same day having different contents yet with the same name.

Information consistency issues manifest in format conversions also. For example, every time a WordPerfect bill is converted to PDF or HTML or Folio Views there is scope for the introduction of variation. An example of this is the web pages being created at present for the website. This is happening outside of the control of the Legislature.

Bill information is split into a number of separate storage areas and, because of the silo nature of the current systems, these can get out of synch. Some examples:

- The source sheets that are critical to OCR’s process are generated with human intervention. In other words, it is necessary for the user to remember to re-generate the source sheet to keep it in-synch with the bill text.
- Citation reports, which are critical to the statute update workflow, are created from the Source Sheets (which may be out of date) and cached (another possible way for them to get out of date).
- Bill Author and Sponsor information is split between multiple data stores that can get out of synch with each other and also occur in the text of the bills themselves.
- Bill digest information exists in multiple places that can get out of synch.

There are some significant re-keying points in the workflow. Re-keying is problematic for information consistency for a number of reasons:

- It introduces latency
- It can introduce human error. This can be managed by introduced additional proofing cycles but this has cost.
- It consumes scarce human resources
- Some examples of re-keying in the current systems:
  - The LS identifiers introduced each day in chamber (from the file-list), are manually entered into Bill Status by OBDAR
  - Vote totals are manually keyed into Bill Status from information coming out of the vote system as PDF files
  - The material that appears on interim committee appointment letters is entered into LSA computers three times and re-entered by chamber staff

- There are at least three separate systems that need agency contact information but manage that data independently. (OBDAR, Bill Status and the Personnel tables in Bill Drafting.)

The legislative staff has devised a number of checks and balances in the system to manage the consistency issues. However this task will become more difficult as new services are added. If new services are added through new silo systems it will only increase the number of workarounds and possibilities for consistency issues.

#### **4.4 Information Latency**

The issues of information latency are closely related to information consistency above. However, it is such an important aspect of the way the current IT systems operate that it is treated separately here.

When looking at any one of the many different “views” of bills currently available in the Legislature, it is not generally possible to tell if an inaccuracy is the result of a delayed update or a true inaccuracy. Some examples:

- Updates to the website are not under the control of the Legislature and can be 30-40 minutes behind chamber release
- Some updates to Legisoft are created by batch processes creating PDFs in the background and creating copies of files. These batch processes (known as “robots”) may be simply busy, may be experiencing problems or may be just going through a normal 15 minute update cycle

There are situations that arise in which documents need to be withdrawn after they have been released. In the case of Legisoft, the Legislature has control but in the case of the website, the Legislature finds itself reliant on a third party to withdraw the document.

Note that these issues of latency and accuracy of the information published to the public apply to the existing databases (especially Bill Status) as well as the WordPerfect documents. For example, the bill status display on the website is driven by a separate bill status database from the Legislature's internal bill status database. It is sent to the external hosting company every 30 minutes.

#### **4.5 Compare-centric and Proofing-centric Workflows**

The current workflow involves a large amount of document compare and document proofing operations. There are two primary sources for compare/proofing:

- Bill amendments are created as separate documents from the bills they are amending. A merge algorithm attempts to work backwards from the amendment descriptions to make the necessary changes in the bill itself if the amendment is engrossed by IHP. Although the application that does this (the merge program) achieves 90%+ accuracy, the complete document must be proofed in its entirety because there is no way of knowing if/where the 10% merge inaccuracies may be in the output.
- Statute pull-ins into bill drafts are independent copies of statute section text. As bills go through the workflow they are amended and the text of statute sections they contain can be amended also. At any given point where it is necessary to determine if a section of statute in a bill is different from the statute database (or different from another instance of that statute section in another bill), document compare operations are required.

These can be resource intensive processes that add time to the production of important documents. The staff are limited by the systems they have at hand today to fulfill these functions.

#### **4.6 Audit Trail**

At present, the Legislature's staff spend a significant amount of time manually managing, tracking and manually determining provenance of bills, statute sections, fiscal notes, etc. The tracking/provenance systems are separate silo systems in the Legislature. There are a multiplicity of such systems and some individuals/offices create private

tracking systems using, for example, Excel spreadsheets. Although these systems are run by individuals and possibly stored on their local machines, they are key to the process running smoothly. This is a risk to the Legislature as single dependencies should be avoided where possible.

## 4.7 Hard-wired Business Logic and Outputs

As business processes have changed over the last decade the volume of workarounds has increased. Examples include:

- The 99 amendment limit in the original design is now problematic
- Amendments to Resolutions was not envisaged in the original system
- Vehicle bill procedures have changed
- Handling of legislative deadlines such as bill introduction/hearing cut-off days is problematic as the logic sits inside Bill Status
- Special session setup, switching and handling has evolved and is problematic

For some changes in business processes, no workarounds are in place. Some examples:

- Sometimes, bills need to go to the “dog house” without a fiscal note. When this happens the tracking system lists them as “in review” as opposed to “released but fiscal note pending”. Thus, reports can be inaccurate reflections of the status of a bill.
- Requests may come from the Senate with “Committee on Rules” as the author but OFMA needs the name of a person for whom the the notification letter should be sent.
- Members only have one e-mail address slot in the Bill Drafting system but many members have multiple e-mail addresses.
- OFMA cannot easily send copies of fiscal notes to all sponsors of a bill because the tracking system only makes available the original author.

Some outputs that have changed over the years are created in a hard-wired fashion (i.e. direct production of final PDF) that makes “tweaking” the outputs problematic. Some examples:

- Merging of calendar days from PDF-based reports coming from Bill Status
- Handling of the conference committee grid
- New custom bill status codes

These are only examples of rules that were implemented that require more flexibility as processes change. The staff have learned to live with the limitations and work around them but the requests for more new services and access to information is putting a strain on how much time they can spend on workarounds.

## 4.8 Notifications

There is currently no system receipting/notification protocols at important workflow checkpoints. For example, the only way for OFMA to know that a fiscal note should now be on the website is to check to see if it is there. If it is not there, it could be because the website service provider has it but has yet to get it onto the website. It could also be that they have not yet received it because the batch process (robot) has not yet run or because the robot has a problem.

A large number of different notification mechanisms are being used in the Legislature for business events such as “new committee report” or “request bill file transfer to chamber”. Some examples:

- Point to point e-mails and phone calls
- File drops into shared folders
- Automatic folder watchers (robots)
- printing documents to specific print queues
- staff monitoring chamber audio
- FTP (to Website and Chamber of Commerce)

The caucus staff (Attorneys and Fiscal Analysts) create many of their own tracking/notification systems in Excel. Leadership staff track conference committees, chamber staff keeping their own tracking lists to follow concurrence and enrolling with the governor's office, OBDAR sometimes uses spreadsheets to track the status of requests etc.

There are no audit trail records created as a byproduct of robot activity. For example a robot watches the viewfn/ folder to determine what fiscal notes have changed and thus need to be processed. However, OFMA have no way of knowing when the material is picked up or when it gets to Legisoft or to the Website.

There is significant use of "out of band" phone calls and point-to-point e-mails to trigger workflow movement.

OFMA fiscal analysts sometimes are unaware of amendments because Bill Status and Bill Drafting systems get out of sync.

#### **4.9 Miscellaneous Challenges**

- The requests for new services to be provided is putting pressure on a staff that is spending significant time working around the limitations of the current IT systems.
- There is increasing pressure on turnaround time for post-session work because the date to open up pre-filing is getting earlier each year.
- The line and page processing workflows (i.e. the Merge program, the Excel part of the Budget bill etc.) are heavily dependent on tight control over line and page numbers. The word processor configuration is sensitive to the exact details of the printer being used at the time a document is displayed on the screen. This "printer sensing" creates significant dependencies on precise control over fonts, page margins, printer drivers and physical printer types.
- Special session handling is semi-manual at present.
- Aspects of the current systems get slower at peak times (e.g. when the chamber file-lists get long) or getting slower as the overall volume of bills in the system gets higher (e.g. the conflicts report) or the bills get bigger (e.g. budget bills). This can have a negative snowballing effect. For example, a second reading amendment may need to be created and distributed to interested parties immediately, but the engrossed bill is not yet available. The scope for error in this is significant but the pressure to do it might also be very significant (e.g. session deadline), potentially creating a significant issue to be cleaned up later.
- There is a high percentage of clean-up code in the overall custom code base. An estimated 30% of the source code in the OCR system exists to clean up precise layout and content of documents so that downstream processes will function correctly.
- No statistics are available at present on what parts of the very large surface area of custom application code is actually being used.
- Workflow procedures become complicated at points where the legislative day differs from clock/calendar time. For example:

- Chamber days that extend past midnight yet might need to record events at 11:59
- Formal adoption times for changes to bill digests. The change may have been formally adopted on one day yet recorded into the system a day later.
- Color images in handbooks require special treatment as EPIC is primarily paper focused (black and white) at present
- Indiana Registers are based on an XML based application but the Indiana Administrative Code is WordPerfect based. The existing system is a partial implementation of what was originally envisaged. There are some significant issues with the current design, specifically around managing the changes to the code. Moreover, the partial nature of the implementation has resulted in much of the pain associated with XML-based systems without the gain.
- There are problems with special character handling in the Indiana Administrative Code. When these occur, manual changes to the XML and possibly the style sheets are required. In the past there has also been issues which appear to be a bug in the Apache Cocoon/FOP software versions used. It is not known what parts of the system might break as part of an upgrade to address this bug.

## 5 Root Cause Analysis

The seven primary issues identified earlier in this report can be addressed through the introduction of a Legislative Information Model/Architecture. Underlying this architecture are a number of guiding principles drawn from the field of Enterprise Architecture and industry best practice. These are discussed later in the section related to recommended features of any new system the Legislature might consider.

Here we will concentrate on the aspects that directly pertain to addressing the seven issues identified:

Issue	Remedies
Information Silos	Creation of an Enterprise Information Model, covering naming conventions, normative copies, permanent identifiers etc.
Information Consistency	Identification of normative electronic document for each critical legislative information object (bill, committee report, statute section, member, committee etc.), at each stage in the workflow.  Managing bill information and meta-information as a single management unit
Information Latency	Electronic point-in-time citation. Replication based website. Robust notification channels
Compare-centric and Proofing-centric Workflows	Amendments in context. Metadata annotation of documents such as bills.
Audit Trail	Temporal, versioning database
Hard-wired Business Logic and Outputs	Flexibility business logic layers using edge-based design.
Notifications	Robust asynchronous integration points and normative, fault-tolerant notification channels.

### 5.1 Enterprise Information Model

Much of the work of a Legislature revolves around a small number of key information objects such as bills, statute sections, committees etc. By taking control over the naming and electronic citing of these key information objects, many problems associated with silos, tracking and notifications can be addressed.

The naming system in the Information Model should cover all the primary nouns that interact in the Legislative workflows. Examples include:

- Bills
- Statute Sections
- Members

- Committees
- Amendments
- Bill Status Action Codes
- etc.

It should include all the information objects that appear in mission critical mappings such as:

- members ↔ committees
- sponsors ↔ bills
- staff ↔ committees
- statute sections ↔ bills
- etc.

At present, multiple sets of identifiers exist for these information objects, spread across the silo applications and across different workflows.

Ideally, the naming conventions would have the following attributes:

- Short and URL friendly
- Readily understandable
- Globally unique

## 5.2 Normative electronic documents

It is inevitable, given the way computers work, that there will be copies/variants made of critical information objects as the move through legislative workflows. At present in the Legislature, there are more copies/variants being made than would be ideal, however, there will always be some amount of copy/variation involved. Examples of inevitable copy/variant creation operations:

- When material moves from confidential areas to shared areas, copies are inevitably made
- When bill material is converted to web pages or when committee schedules are converted to iPad calendars, variants are inevitably made
- When sponsor information is loaded into a database for fast sorting/retrieval. copies of information present in the bill text are inevitably made

The goal of the Information Model is not to eliminate all copies/variants but to firmly establish which digital object will be considered correct, in the event of a discrepancy. We use the term “normative” for this concept. The Information Model should also allow for reporting on provenance to allow the full audit trail of the document to be systematically tracked throughout it's life-cycle.

Moreover, as the trend towards paperless operation of Legislatures accelerates, it is becoming increasingly important that it is possible to identify what copies of what digital artifacts are considered authentic in the legal sense. For example, in the sense of the Uniform Electronic Legal Materials model law from the Uniform Law Commission.

Closely related to the question of normativeness of content is the current split of bill information into document components and database components. Some information is present only in the document e.g. the text of a bill. Some is both in the document and in the database tables e.g. sponsor information and bill digest. Some is only in the database e.g. requester identifier. By managing all data related to a bill as a single management unit, all the issues related to the parts getting out of synch with each other can be addressed.

Normative concerns also apply to the statute-related metadata for bills currently created in the Source Sheets. Adding a section of statute to a bill is a copy/paste operation by a typist at present. There is no automatic associ-

ation between a bill and the statute sections it relates to. The Source Sheets serve to capture this relationship but because they are separate documents they can get out of synch.

### 5.3 Amendments in Context

Although a certain amount of comparing and proof-reading is inevitable and proper in a Legislature the tools to support the process can be greatly improved by leveraging amendments-in-content. Simply put:

- Amendments to bill text can be created in such a way that engrossment is fully automated and 100% accurate in most cases
- Statute sections can be tracked in bills in such a way that the text of the statute is locked-down during the bill workflow, easing the burden on the proofing process

Moreover, basing bill amendatory cycles on amendments-in-context makes it straightforward to publish bill amendments in a form that many legislators/staff prefer to read. i.e. redline/strikeout directly in the text of the bill rather than “on line 7, page 6...” Summaries that remove unmodified material, distilling bills down to just the changed portions etc.

Amendments in context would also simplify the problem of performing updates to the OFMA spreadsheet used in budget bills as analysts would be able to tell at a glance, what the changes are. Amendments in context would also help fiscal analysts who spend a lot of time trying to determine what has changed in a bill in order to update its fiscal note.

### 5.4 Temporal, Versioning Database

In order to have a comprehensive, authoritative audit trail, it needs to be an automatic by-product of legislative workflows. For example:

- As a bill moves around a workflow, the history of its movements should be automatically tracked so that bill status becomes a report as opposed to a self-contained silo
- As the text of a statute section changes, a comprehensive history of the changes should be automatically generated so that it is always possible to explain how the text came to say what it says
- As a fiscal note moves through its internal workflow from first draft to approved version, a comprehensive history of the changes should be automatically generated so that it is always possible to explain how the text came to say what it says
- As any mission critical document in the Legislature changes over time, there should be an automatic snapshotting/versioning occurring so that any and all previous versions of the document can be recovered in the future.

### 5.5 Flexible business logic

There are two primary aspects to flexibility in the context of business logic in legislative workflows. Firstly, outputs such as calendars and committee schedules which are currently produced directly as PDFs can be produced in a “soft” form that would allow edits post generation but prior to final publication.

Secondly, business rules that are likely to change over time should sit within applications that connect to, but are distinct from, the core temporal versioning database. That way, changes to the business logic can be made without impacting the data layer.

The legislative staff should have abilities to reconfigure certain workflows when events happen that are outside of the norm. The legislative process can frequently throw up events that are outside of the norm.

## **5.6 Robust Integration Points and Normative Notification Channels**

The current batch processing mechanism known as “robots” is sound, conceptually, but requires the addition of fault-tolerance and load-balancing.

These need to be fault-tolerant because they are mission critical. It needs to be possible to monitor and react to health issues before they progress to impacting service. By making them load-balanced, two important benefits are derived. Firstly, individual instances can have failures with no disruption to service. Secondly, multiple instances can perform work in parallel in times when the workload is high.

Both can be achieved architecturally by the addition of asynchronous message queues. Simply put, robots would then behave like e-mail readers in the sense that messages simply queue up for them in the event that they are busy or offline. No messages are ever lost.

# 6 Proposed Architecture

## 6.1 Introduction

The proposed architecture implements a central IT system for managing legislative information in a way that solves a number of the existing IT issues as well as providing the platform for the provision of the new services required in the Legislature. The proposed architecture anticipates that a number of new information services will be required that are not yet defined. It will support the provision of these services through implementation of an extensible framework and allow the Legislature IT staff to take ownership over building out these services in the future.

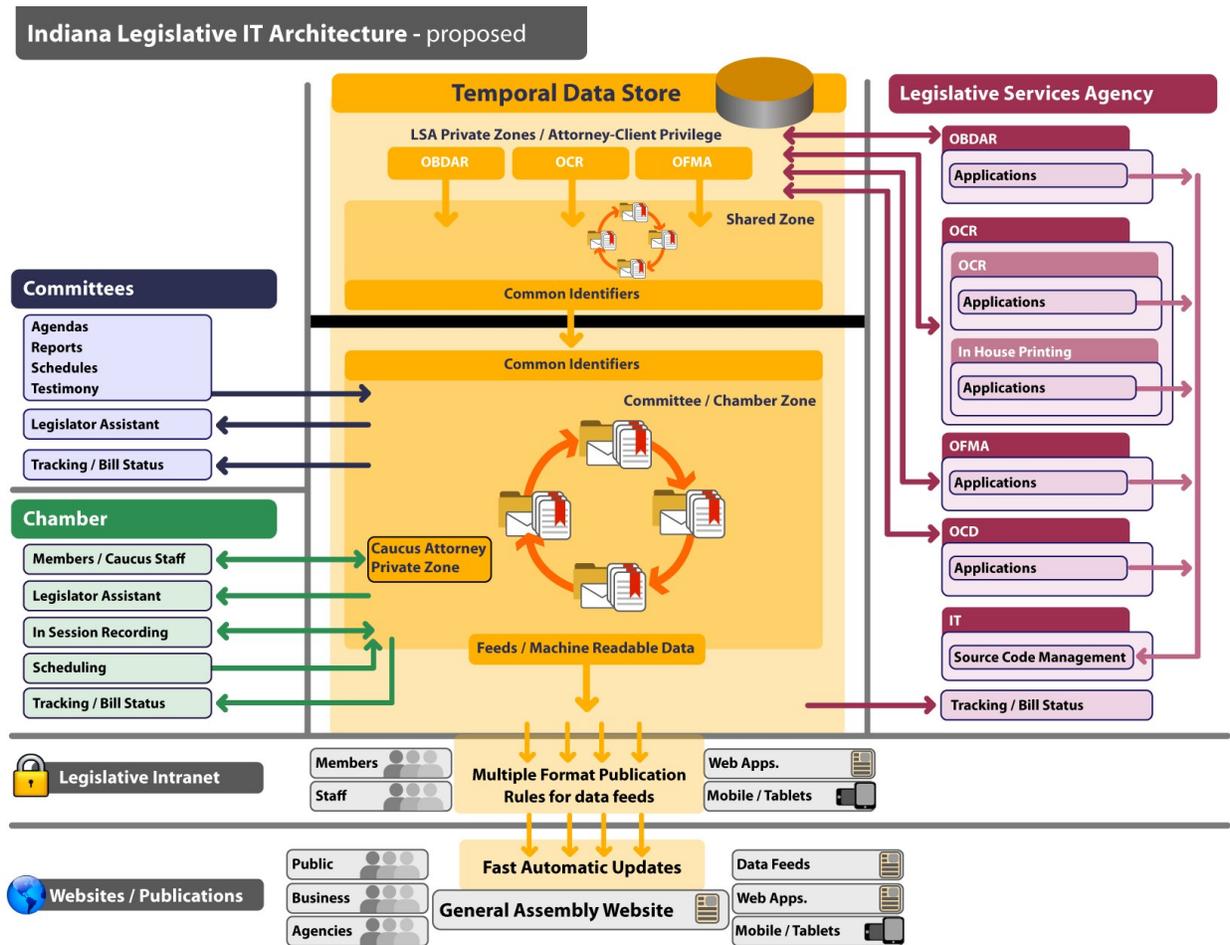


Figure 4.1: Proposed IT Architecture

## 6.2 Recommended Features and Design Patterns

It is recommended that any proposed implementation of a system conforming to the above diagram be evaluated for support of the following Features and Design Patterns.

### 6.2.1 Information Governance

- Ownership of key legislative assets and applications should belong with the Legislature as a whole, not with any specific office function or application. This is especially true for assets that “flow” through the various offices and are thus shared assets out of necessity e.g. Bills.
- Naming conventions for all critical assets should be centrally managed. A single set of identifiers is required for any events/documents that need to be considered as a whole for reporting/workflow purposes.
- Control of privileged access to the assets of the enterprise needs to be managed and single-points-of-failure in people, processes and technology need to be carefully mitigated.
- Ownership and access to confidential information should be managed without the creation of silos that allow work to flow across access boundaries without latencies, re-keying etc.
- Implement “hand-shake procedures” at the office workflow boundaries so that the enterprise-wide tracking system can tell where an item is, when it arrived, when it departed, how long it was at any give point in the flow etc.
- Discrete pieces of information needed in a Legislative system should be entered exactly once. i.e. no duplication of data entry.
- The overall flow of information through the system should be controlled by policy. A clear separation between policy and operational concerns should be in place so that concepts such as turnaround-times, publishing release speeds and so on can be set in policy, rather than have the technology dictate these policies.
- Managing a corpus of law is a very significant QA/QC undertaking. There is no more important corpus of content in the State and yet, out of necessity, proposed modifications to it are made in parallel, under pressure during short time periods in Legislative sessions. This creates a classic Time/Cost/Quality triangle. It is very challenging to keep all three dimensions at the right levels at the same time i.e. keep time and cost down while keeping quality high. The best place to ensure that trade-offs occur in the right place is at the information governance level:
  - Insofar as practical and excluding information that is necessarily confidential, everybody in the Legislature should have access to as much information as possible about:
    - current master copy of the statute corpus
    - current proposed amendments extant across all bills
    - statute sections already changed this session by enrolled bills
    - bills that have not yet had a hearing, etc.
    - pending technical corrections
  - The more individuals looking at the corpus and the update cycle as it progresses through the legislative session, the more chance that potential problems will be found quickly. In an ideal world, post-session publishing work would be largely devoid of any work related to corpus inconsistencies.
- The Legislature's IT systems have many, many moving parts that are inter-related in complex ways. A small change to a WordPerfect macro in OBDAR could have knock on effects in OCR or OFMA. The critical co-de-

dependencies make it imperative that all software components – from WordPerfect Macros to Database reports, client side and server side, be under full formal source code control and life cycle management.

There is currently no formal build process and no access to testing/staging/live environments for developers. There is no rigorous component level or application level versioning. This is accentuated by the fact that some of the most critical algorithms are “just WordPerfect macros”. Word processors as a whole, make it difficult to put proper software engineering disciplines in place. It is easy for power end-users to roll-their-own macros using “Tools --> Record Macro”. This is excellent for one-off activities but what tends to happen is that some of these one-off macros/reports become mission critical over time. However, they never transition to formal management and formal build/test/regression cycles.

- File naming conventions that remove important file-application association information should be avoided as part of the information governance model. For example, internally in OFMA “.fn” is used for fiscal note but the file itself may be a pdf or a WordPerfect file. The use of a naming convention to indicate workflow steps is not in itself particularly problematic but the underlying file type should be preserved also to allow operating systems such as Windows and iOS to know what to do with the file. i.e. “fiscal\_note.fn.wpd”. There are similar issues in ODBAR with .mn (meeting notice) files.
- A single role management system should form part of the information governance model. Currently a mixture of Microsoft Active Directory roles and Oracle Roles is used, combined with application-specific role-based access control logic.
- We recommend that the information governance model take into account the increasing importance of geo-spatial technologies as a tool for policy formation and legislation analysis. For example, the use of electronic interactive maps as tool for fiscal notes and tax/pension handbook created by OFMA.
- We recommend that the information governance model include an Inventory of all custom code that was created by external as well as internal entities, with a view to ensuring it is properly managed, backed up, versioned etc. Some examples:
  - Adobe Acrobat Plug-Ins used in Legisoft, originally created by an external contractor
  - Folio Views query links, fsr scripts created with help from Information Solution (Jerry Junker, Kentucky)
  - Configurations of Google Appliance used on Administrative Code website
  - The queuing software created by ICON for use by the EPIC/IHP groups.
  - The Apache Cocoon environment used in the Indiana Administrative Code (XML schemas, cocoon pipelines, style sheets, transforms, FOP configurations)
- We recommend a clear separation between historical and current data. For example, Bill Status is currently historical back to 1994. The ODBAR system has old drafts back to 1989. OFMA has Excel budget files back to 2011 and Lotus 123 budget files back to 2001 . From a data management perspective a balance needs to be found between access to historical artifacts, preservation, and work-in-progress data volumes.
- The information governance model can address questions of normative data. An example at present is Bill Subject indexes. Although they are clearly “attached” to bills they are not in the bill objects themselves at present (they are in the Bill Status system only). In the Information Model, bill text and all bill metadata – to include subject index terms – should be in the same information object in the temporal asset store.
- The Legislative Information Center and the House/Senate information centers should, as much as possible, be working off the same normative reference resources to ensure information consistency.
- Re-using identifiers should be avoided if they are a potential source of citation confusions/errors. Note that this will involve creating multi-level identifiers. i.e. referring to a house bill as a combination of (ses-

sion, year, number). That way, bill numbers can reset because the session/year combination serves to uniquely pick out a set of numbers.

- A permanent URL naming convention should be created so that names of key assets such as bills, statute sections, committees etc. are:
  - Short
  - Meaningful
  - Permanent

Invariably, this involves trade-offs and this is best dealt with as part of the information governance model.

- We recommend that the information governance model fully leverage hyperlinks. In many respects, legislative materials are ideal for the web because of the value that can be readily added to a corpus via hyper-linking. The web pages produced from the WordPerfect files do not contain hyperlinks.
- The Information Architecture needs to ensure it can accommodate platforms such as Facebook, Twitter and also supports Web 2.0 machine readable data formats such as RSS/Atom, vCalendar, vCard etc.
- We recommend that the information governance model prefer thin to thick client application development. LSA have approximately 500 desktop PCs. The management overhead of installing/maintaining/synchronizing software on this many desktops is significant. Overheads can be reduced by using as much thin client application approaches as possible. Moreover, the clear direction of the IT industry – and thus many of the IT suppliers Legislatures are dependent on is toward thin client, public/private cloud computing.

### 6.2.2 Fully Automatic Audit Trail of all workflow actions

The information processing in a Legislature can be conceptualized as an accounting system in which

- the statute at the start of a biennium is the equivalent of an “opening balance”
- a legislative session is the equivalent of an “accounting period”
- actions on bills processed during a session are the “ledgers”
- the codification of statute is the equivalent of “rolling forward” new balances

As with any accounting system, the audit-trail – the ledgers – constitutes the record of what happened. It is critical that the audit trail is comprehensive and automatic. It should not be the case that users could forget to update it or a developer's applications could fail to update it or update it incorrectly.

Just as in an accounting system, any errors that do happen (e.g. bill referred to wrong committee) should be left in place in the ledgers but with so called “contra-entries” added to reverse the action. This maintains the integrity of the audit-trail as the definitive record of what happened – including any clerical steps taken along with the business-level steps.

In the digital world, a number of words in common use are somewhat misleading and they can cause significant problems in Legislature:

- The word “copy” rarely can be understood to mean “exact copy”
  - A WordPerfect document may look different on my screen/printer/word-processor than it does on yours even though we are looking at exactly the same digital file contents
  - A PDF/RTF/Folio Views “copy” of a WordPerfect document may be substantially different from the original depending on a whole variety of factors

- The word “send” can rarely be understood to mean what it does in the physical world. Every time a file is “sent” (via e-mail or network drive or ftp) the reality is that a new “copy” (see above) is made.

The creation of variants of a document such as a bill or a statute section is impossible to avoid. It is quite simply an unavoidable side-effect of the way computers work. Consequently, it is vital to have a clear understanding of which “copy” is considered authoritative for all audit-trail steps of critical information objects such as bills.

### **6.2.3 Robust content citation and content authentication**

- The audit-trail should be leveraged as a source of provenance information so that the authenticity of electronic materials can be established.
- It should be possible to cite any document asset in the Legislative repository based on any desired point in time.
- Citations should, as far as practicable, take the form of hyperlinks so that the corpus can be navigated using RESTian API principles.

### **6.2.4 Machine readable data feeds**

Rather than only publish rendered documents such as PDF or HTML, machine readable forms should also be generated so that new applications can be created quickly and so that Web technologies can be leveraged e.g. JSON, XML, vCalendar etc.

Examples of candidates for machine readable “feeds” include:

- Bill status
- Committee Schedules
- Chamber Calendars

### **6.2.5 Flexibility in Business Logic**

In Legislatures, it is best to assume that any existing “rules” are temporary and seek to soft-code these in applications based at the edges of the core data model – not embedded within it. Some examples:

- Special sessions
- New type of concurrent resolution
- Add a fourth bill stage in the Senate
- Create a new status for a bill to be in
- Change the rules for ordering author/sponsor information in journal entries

### **6.2.6 Robust Notifications Framework**

All non-confidential events in the Legislature should have a single reliable event source that applications can monitor. This event source should be tied to the audit-trail so that the information repository and the notifications can not get out of synch. The Information Architecture should explicitly specify the notification flows and separate the official methods from any informal notification methods.

Robots (batch processes) need to be asynchronous with message queues. The goal is not to design a system that is 100% reliable at the parts level. The goal is to make the systems as a whole, 100% reliable through fault-tolerant design i.e. redundancy, horizontal scalability and load-balancing.

### 6.2.7 Bring Your Own Device Support

It should be possible for any web-enabled device to interact with the Legislature's applications. Either via web-native thin client applications, mini-apps or remote desktops.

### 6.2.8 Load Balancing

Many legislative workloads lend themselves to a "divide and conquer" approach. For example, if there are N bills that need to be processed for Chamber introduction, each of the N can be processed independently – by machines as well as by users – because there are few inter-dependencies.

### 6.2.9 Amendments in Context Capability

Application programs that merge descriptive amendments back into bills cannot be made 100% reliable because of the complexities of reverse engineering amendment actions from human-oriented amendment descriptions. Figures provided by the Legislature suggest about a 97-98% accuracy rate. This is an excellent rate but a 100% proofing is always required because it is not possible to know where the 2-3% of inaccuracies may lie.

100% accuracy in merging (engrossing) can be achieved if the workflow is re-ordered so that amendments are entered in-line into the bills and the amendatory instructions generated from the in-line forms.

The positive impact of this on the amount of work In House Printing (engross/enroll) has to do would be very significant with benefits ranging from turnaround time to reduced risk or errors occurring. The tools for staff in OCR could greatly help them to do their work.

## 6.3 Recommendations for short term consideration

The following are miscellaneous recommendations that can be considered in the short term and are independent of initiating any large project.

- We recommend requesting detailed website usage statistics from the external website. These can give valuable insights into what parts of the website are used and what are not used: volumes of traffic, peak times, geographic location etc. The usage data will also give insights into what the URL breakage exposure is, if the Legislature decides to move to a different website.
- We recommend full preventative monitoring of all automated agents i.e. robots. There are 5-6 critical robot processes at present but there is no easy way to monitor their health or be informed of malfunctions. It may seem that adding something to monitor these robots just creates one more thing to watch but once in place and trusted, it will significantly reduce what needs to be watched. Moreover, it is highly likely that some robot failures have warning signs that, if detected, could be resolved prior to any service outage taking place.
- Many of the sensitivities to printers and printer device drivers in the current systems are caused by the way the Word Processor implements its WYSIWYG by sensing the current printer configuration and adjusting the layout of words on the page accordingly. We recommend transitioning to a WYSIWYG configuration that is printer/device driver independent. As well as simplifying line/page handling (a critical part of legislative amendatory cycles), this will significantly reduce the likelihood of two printouts of the same file producing different results for different printers inside and outside the Legislature.
- We recommend that the Indiana Code DVD issues related to missing/malformed bookmarks be addressed upstream i.e. not fixed in the PDF files. This is because fixing the bookmark problems involves modifying the PDF files themselves. This is problematic in terms of the audit-trail and looking ahead, will be problematic for authenticity and digital signatures.

- We recommend that all PDFs produced for consumption internally and especially for consumption externally be created using PDF/A. This is especially important for the Indiana Administrative Code which is already electronic only.
- We recommend role based naming of email addresses, folders etc. Rather than e-mail individuals or name workflow folders after individuals, we recommend the use of role-based naming. This is especially useful for e-mail notifications where multiple people can be trained in a particular role so that if one person is not available, others will get the same e-mail notifications. The existing Microsoft Exchange environment in the Legislature can readily be used to set up e-mail aliases for this.
- There exists currently a large number of different contact-management related documents are created by hand, especially by ODBAR. There is an Address Book application that could be considered as a central application for this.
- We recommend a complete spider of the current website for the Legislature and an analysis of its contents. It may be the case that over the years files have accumulated on the external servers that are not present in the Legislatures data center or backups.
- It is possible – though rare – in the current system for the same number be allocated to multiple PD or LS documents. Given the tracking difficulties that this can create, we would recommend addressing this independently of any other initiatives.
- The vote systems remain Windows XP based whilst the rest of the Legislature has transitioned to Windows 7. We recommend the use of Windows 7's XP compatibility mode to make these Windows 7 based at the base OS level. It may be possible to use Virtualized Desktop technology to put the two voting mini-data centers under the same backup/DR regime as the rest of the Legislatures IT infrastructure.
- OFMA drive C backups created centrally but this does not happen in OCR. Some users in OCR keeping important material on their drive Cs. We recommend either extend the OCR model to all desktops (i.e. all backups centralized) or recommend all users to not keep anything significant on drive C.
- Not all servers are virtualized currently. We recommend full virtualization of servers in order to be able to better leverage multi-core servers and drive down server costs by raising server utilization rates. Also, a fully virtualized server environment will facilitate disaster recovery.
- It is understood that at present, backup/disaster recovery is performed between the main data center in the LSA office building and the Legislature building. We understand that a staff member also takes a backup offsite periodically. We recommend a disaster recovery strategy that involves moving data/applications out of the reach of single-area incidents such as tornadoes. We also recommend a “cold-iron restore” be performed once a year from backup media to ensure that critical applications/data can be restored within a reasonable amount of time.
- We understand that there are plans to upgrade Oracle to Oracle 11G. We would recommend caution in that the term “upgrade” can be misleading for any non-trivial computer system component. The lack of testing/staging/live and the lack of regression testing in the current environment makes end-to-end testing of the existing suite of apps on top of an upgraded Oracle 11G difficult to perform and it is not clear what the value add of the upgrade is.

## 6.4 Longer Term Recommendations

- We recommend enriching the structure and content of fiscal notes. There is a lot of valuable information in fiscal notes under a variety of thematic headings that can be harvested for cross-the-board reporting on bills.

- Creating accurate citations to Indiana Code from Administrative Code/Registers is problematic because the hyperlinks being created are always going to point to the Indiana Code as it is today, not at the time the cite was created in the Administrative Code/Register. We recommend that this be addressed by taking a consolidated view of how the Indiana Code corpus relates to the Indiana Administrative Code corpus.
- The voting systems are standalone entities at present with different code bases per chamber. The Legislature has a Cisco IP phone infrastructure with jacks on each chamber desk. IP Phones are becoming as powerful as laptops/tablets and this creates an opportunity to consider using IP Phone stations as voting devices in the future.

## 6.5 Recommended Phasing

There are multiple ways that the project could be implemented. The following phasing is one option based on the over-riding concerns of business continuity and bearing in mind the drivers and success metrics. We recommend that the project as a whole be split into two separate projects, each of 15 month duration:

- Project 1 : Intranet, Website, Budget Bill and Member Services
- Project 2 : Non-Budget Bills and Statute Management

### 6.5.1 Phase 1.1: Website

- The outputs currently sent to the external website provider are populated into the temporal database in a set of workflow folders corresponding to “published” phase of the workflows.
- An Intranet Legislative Portal is created that replaces Legisoft with a browser based environment. It connects to the temporal database to generate pages of information related to bills, committees etc. dynamically. i.e. as new material becomes available from the internal systems they are immediately visible on the website.
- Member, Committee and Bill pages created based on the Information Model naming conventions.
- WordPress files are processed into richly hyperlinked HTML and integrated with bill status pages, member pages, committee pages. All pages based on Permanent URL patterns. PDFs hyperlinked from the HTML pages.
- The internal site is replicated to create the Internet-facing website. The policy rules for defining what is published and when it is published is made configurable by state staff.
- RSS and ATOM-based feeds are created for Bills and Committees
- A RESTian API is created to allow machine readable content to be consumed
- The Intranet and Internet sites are created to be Tablet/Smartphone/WebTV friendly i.e. BYOD
- Some internal cleanup is performed in preparation for Phase 1.2 related to
  - macro suite consolidation (300+ macros at present. Estimated 20-30% overall size reduction),
  - creation of standard modules for saving documents and metadata
- Some internal “quick win” changes are instituted:
  - Load balancing of conflicts reporting to reduce the 24 hour interval
  - Load balancing of file list processing for faster turnaround time in busy session days
  - Generation of “soft” output formats for calendars and schedules rather than PDFs from existing Bill Status
  - Intranet WebForm to streamline submission of content to LSA from Members and caucus staff

- Digest consistency
- The robots related to external publishing are revisited to trigger based on the presence of actionable content rather than at period intervals. E.g. bill status content dump is currently set to 30 minutes.

### **6.5.2 Phase 1.2: Budget Bill and Member Services**

- The temporal database from Phase 1.1 is extended by adding Workflows, Author/Edit and Publishing methods for handling all aspects of the Budget Bill from drafting to amending to enrolling. Integration points created for:
  - Bill Status Integration
  - statute pull-in and post-session publication
  - Intranet/Internet publishing
- The Budget Bill process should tightly integrate spreadsheet and document technologies to provide users an easier way of managing changes to the bill. The Budget Bill process should also allow the amendments to be entered in context so it is easy to read the direct impact of proposed changes.
- All network drive mappings are changed so that all content is written into the Temporal Database: bills, fiscal notes, statute etc.
- All WordPerfect macros and applications modified to store what is currently document-oriented metadata “on top” of the WordPerfect files. E.g. Source Sheets
- All tracking systems and reporting sub-systems are reviewed and as many as possible changed to utilize temporal database queries. Some will be removable completely because of the availability of normative copies for all document-oriented assets.
- Website enhanced with Paperless Committees applications (HTML5 based, iPad devices). Makes extensive use of temporal database for retrieving normative assets and reporting on audit trails
- Legisoft features replicated on the Intranet and the current Legisoft application retired
- Temporal database is used for core Bill Status functions. Committee/calendar features of current Bill Status system replaced. This is then used as a single source of bill status information for all applications and workflows.
- The voting system could be replaced as part of this phase.

### **6.5.3 Phase 2.1: Non-Budget Bills and Statute Management**

- Statute book converted to XML-based open data format
- All OBDAR/IHP/OCR functions related to bill drafting → statute codification replaced with XML-based open data format applications.
- Amendment cycles modified to utilize bill amendments in context. Amendment-in-context documents added to the Intranet.
- Bill drafting subsystem is statute-aware and can “lock” statute sections, thereby streamlining document compare workflows, post-session publishing, conflict detection etc.
- Corel WordPerfect usage minimized/eradicated except for IAC

### **6.5.4 Phase 2.2: Fully Paperless Legislature, Indiana Administrative Code Integration**

- The Indiana Administrative Code is converted to XML based open data format

- The Indiana Register XML format is migrated to the same XML notation being used for all Bills, Statute sections, the IAC etc.
- Existing IAC/Register website replaced with the consolidated Intranet/Internet websites.
- Point-in-time hyper-linking IAC <-> IC enabled.
- Normative copies established for Registers and the Indiana Administrative Code
- Digital signing and UELMA for the entire corpus
- Fiscal Notes enhanced with:
  - cross-fiscal note automatic reporting replacing some of the manual processes used in tracking today
  - Geo-Notes for interactive and static GIS based analysis
- Committee video streams integrated with committee agendas and hyperlinked

## 7 Appendices

### 7.1 Appendix: Current Technology Stacks - COTS

Oracle Database	
Visual Foxpro v7	
Google Search Appliance	
Apache Cocoon	
Apache FOP	
Corel Wordperfect X5	
Folio Views	
Oracle Roles	
Oracle Forms	
FTP	<a href="ftp.iga.in.gov">ftp.iga.in.gov</a>
WebDav	
Windows Media Files	IHETS (Video)
OpenText Web/RedDot	Web Application and Content Management System used by NIC to host the Legislature
Perl	Used by OpenText Web/RedDot CMS
WebTrends	User Tracking used by OpenText Web/RedDot
Wp2Html	<a href="http://www.brackenbeds.co.uk/Wp2Html/">http://www.brackenbeds.co.uk/Wp2Html/</a> Used by AccessIndiana to convert WordPerfect files to basic HTML.
Oracle AppServer/Java J2EE	Java JSP used by Access Indiana Oracle App Server used in Bill Drafting Application (with Oracle Forms, Oracle Roles and auto-generated JSP pages+servlets) Apache Tomcat used by Indiana Administrative Code
Corel WordPerfect PDF generator	Part of WordPerfect X5
Adobe InDesign	Used for OFMA-produced handbooks : Tax Handbook, Pensions Handbook
Google Maps	Used for Redistricting application
CityGate AutoBound	Used in "Find your Legislator" part of Web Site
ESRI ArcView	Used in OCD
USC Web Service	United States Census. Used in "Find your Legislator" part of Web Site
Microsoft Office (Word, Excel) 2007 and 2010 versions.	House rules created in Word .docx. OFMA prepare the numbers-oriented portion of the budget bills in Excel. Bubble attorneys, fiscal analysts make extensive use of Excel for ad-hoc tracking systems Property Tax Report
Microsoft Access	IRC, Citations Affected Report, Duplicate Citations Report, OCD OCD Use for GIS data Daily conflicts report (Access front end to Oracle backend)
Microsoft Publisher	Used by OCR
IRC	Distinct voting system software and hardware in House and Senate
Microsoft Windows 7	Staff desktops. Member Laptops.
Microsoft Outlook	
Microsoft Exchange	
Microsoft Active Directory	
Lockheed Martin, Internet Quorum	Used by caucus staff for constituent management.
VMWare ESXi 4.1	Server side virtualization
Windows 2008 and 2008 R2	Server side operating systems
Linux Red Hat Variants (from v5 to 6)	Server side operating systems
iFiles	Used by Committee staff as part of iPad prototype.

Microsoft SQL Server	Used in the Data Warehousing project.
Microsoft ASP.NET	Used for iPad prototype and for District Lookup application (find my legislator).
Citrix XEN Enterprise	Used as part of iPad initiative
Oracle Stored Procedures	Used in Bill Status
PCL5E	HP printer composition engine and drivers
CISCO RightFax	Sometimes LSA fax letters to agencies. Some fax of doghouse letters to Members. Usage declining.
TIFF	Used for page scanning of vote sheets in committee, testimony.
Web Services	Call tracker and District Lookup share a couple of Web Services
Microsoft SQL Express	Used in Call Tracker and District Lookup
KML	Files produced by OCD used in District Lookup
DeltaCompare	Used instead of WordPerfect compare in some situations but uses the rtf exports from WordPerfect. <a href="http://www.workshare.com/products/wsdeltaview/">http://www.workshare.com/products/wsdeltaview/</a>
WordPerfect SDK	
USC	United States Census web service. Used in Find Your Legislator.
Google Geo-Locater	Used in Find Your Legislator.
Cisco Agent Desktop	Part of Call Tracking for incoming phone calls
Internet Explorer Browsers version 8 and 9	
Magenta Research	Aspects of the Senate audio/visual environment
Adobe Acrobat Distiller	Used by IHP.
Crystal Reports	Used by OCD
Windows Registry	Used as a form of cache for session-level information.
Borland Delphi	Used in Bill Drafting applications and in Legisoft
HP EVA 6400 Network Storage	
Oxygen XML Editor 7.2	Used for the Indiana Register
GoodReader	PDF Reader for iPads
Zenprise	Mobile Device Management

## 7.2 Appendix: Current Technology Stacks – Custom Code

- WordPerfect Macros
- Foxpro Applications e.g. Bill Status and Robots
- Folio FSR and build scripts
- WordPerfect SDK Tools
- Adobe Acrobat Plug-ins created by PDFDream
- Printer Queuing applications created by ICON

## 7.3 Appendix: Distinct Sources of Information about a Bill/Resolution

For an average bill that passes through the 8 typical phases of the Indiana legislative process, there are many, many potential sources of information that may not be identical. This can cause significant confusion at all points in the workflow.

The main distinct sources of information are:

- Bill Text WordPerfect File x 8
- Bill Text PDF of each WordPerfect with watermark for website x 8
- Bill Text PDF of each WordPerfect for internal use (e.g. doghouse) and Legisoft x 8

- Green sheets (OCR. paper .Initialed by everyone who works on it.)
- Fiscal Notes x 8
- Folio Views infobase of Session Laws
- Folio Views infobase of Public Laws
- Bill Status Database Tables
- Tracking Database Entries
- Access Indiana Bill Status (different from internal)
- Paper jackets maintained by chamber staff
- Journals PDF
- Journals WordPerfect
- Journals Folio
- Author/Sponsor information in Bill Drafting
- Author/Sponsor information in Bill Status
- Source Sheets
- The Digest at the front of the bill
- The Digest in Bill Drafting
- Merge Mockups
- Bill Summary File (OCR)
- HTML pages created from WordPerfect by Access Indiana
- Chamber Tracking Spreadsheets
- caucus Staff Tracking Spreadsheets
- Comparison of Conference Committee Reports
- Conference Committee Grids
- Duplicate Citation Report
- Enrolled Acts Report
- Delta Compare Reports (processing the RTF – not the WordPerfect)
- Bill Certifications
- Chamber Filing Book
- State Records Commission PDFs

The reasons for potential differences between the variations can be split into a number of categories:

- Synchronization latencies e.g. website update delay
- Lossy conversions e.g. HTML pages produced from WordPerfect Files

- Re-keying errors
- Re-keying latencies
- Silo data model mis-matches : e.g. Digests

## 8 Statement of Work – Project Artifacts

### 8.1 Statement of Work Deliverable Documents

All deliverable documents are located at [https://docs.lis.state.oh.us:9880/docs/sow\\_final\\_delivery](https://docs.lis.state.oh.us:9880/docs/sow_final_delivery)


### 8.2 Project Artifacts

Documents gathered during the analysis – all located at <http://10.100.50.40/repos/repo/> on the VPN (<https://vpn.iga.in.gov>)

Folder	Details
\received_documents\Macro Documentation	Details and descriptions of macros used by OBDAR, OCR, OFMA
\received_documents\Macro Documentation\Auto Convert	Instructions for using Auto Correct
\received_documents\Macros\obdmacros	OBDAR macros
\received_documents\Macros\ocrmacros	OCR macros
\received_documents\Macros\ofmmacros	OFMA macros
\received_documents\OBDAR\Manuals	OBDAR drafting manual
\received_documents\OCR\Manuals	OCR process manuals
\received_documents	Bill status ERD
\ocdm\archive_output_lists	Staff created output lists (later merged into main output list)
\ocdm\sample_documents\chamber_samples	Chamber sample documents
\ocdm\sample_documents\chamber_samples\tracking_h_maj_caucus	Caucus sample documents

\ocdm\sample_documents\chamber_samples\h_clerk_docs	House clerk sample documents
\ocdm\sample_documents\Interim	Interim committee sample documents
\ocdm\sample_documents\Jeff\Samples	Sample documents from Jeff Ford
\ocdm\sample_documents\OBDAR	Sample documents from OBDAR
\ocdm\sample_documents\OBDAR\Administrative	Administrative sample documents
\ocdm\sample_documents\OCR	OCR sample documents
\ocdm\sample_documents\OCR\House&SenateJournal-SamplePages	Journal samples
\ocdm\sample_documents\OCR\Manuals	OCR manuals
\ocdm\sample_documents\OFMA	OFMA sample docs
\ocdm\sample_documents\Varnie\Samples	Chamber samples from Varnie Karmo

## About Propylon

---

Propylon is a world leader in legislative and regulatory data processing and document management. With a large pool of expert consultants and analysts, and a world class software development team dedicated to legislative and regulatory domains, we have considerably broader and deeper expertise than any other vendor in the market. We pride ourselves on having some of the leading world domain technical authorities on our team and contribute to development of global standards relevant to legislative data processing.

Our software solutions cover the different departments in the Legislature providing an integrated information architecture for the legislative staff to enter information once and get the most value out of it in multiple places. Our drafting solutions introduce efficiencies such as automatic engrossment that allows for the return of amended bills more quickly. Our Chamber solutions integrate chamber actions to allow for generation of the Journal, Calendar and Bill Status information reducing risks of data entry error and cutting down on the end of day legislative processes. Our Research tools provide efficiencies in the integration of data from multiple sources for the production of the research divisions publications. The results of integrated information management are better tools for staff, members and the public. Staff can easily generate reports of the information they need, legislators can be provided their own portal to access all legislative information and the public website can be automatically updated as required.

Propylon has introduced cost saving for Legislatures by heavily reducing print and composition costs through the provision of electronic camera ready publications. Also, the consolidation of applications and services has reduced support and maintenance costs. Efficiencies of the application allow staff to be redeployed to higher value tasks such as better provisioning of electronic material to members and the public.

Propylon's work in the state Legislatures of Kansas, North Dakota, and Pennsylvania as well as the Irish Parliament, Irish Office of the Attorney General, Welsh Assembly and with legal publisher LexisNexis demonstrates our ability to deliver innovative and cost effective technical solutions for authoring, managing and publishing complex legislative documents.