



Российского сегмента Международной космической станции // Мехатроника, автоматизация, управление. – 2016. №1. Т.17. – С. 42–46.

К.В. Ситников, П.В. Ситников

ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА УПРАВЛЕНИЯ ЭЛЕКТРОННЫМИ АРХИВАМИ КАК ЭЛЕМЕНТ УПРАВЛЕНИЯ ЗНАНИЯМИ В СОВРЕМЕННОМ ОБЩЕСТВЕ

(ООО «Открытый код»)

Аннотация. В статье рассматриваются описание, структура и особенности интеллектуальной системы управления электронными архивами. Выделены уникальные преимущества использования интеллектуальной системы управления электронными архивами.

Ключевые слова: архив, информационная система, семантический поиск, интеллектуальная система, электронный архив, интеллектуальный поиск, база знаний.

В настоящее время в мире современных информационных технологий практически ни одна организация не может обойтись без электронного архива. Электронные архивы документации существенно экономят время сотрудников, ведь именно благодаря цифровым архивам документы всегда будут находиться под рукой у каждого заинтересованного специалиста.

Перевод документов в электронный вид – это также возможность предотвратить воздействие факторов, нежелательно влияющих на сохранность работ, которые представляют историческую и культурную ценность, и предоставить к их электронным версиям более широкий доступ.

Кроме того, внедрение электронного архива значительно повышает надёжность хранения и доступа к архивным данным всех категорий пользователей

Формирование электронного фонда документов и автоматизация процессов работы с ними позволяет осуществлять создание, управление и работу с электронными архивами государственных и коммерческих организаций, включая интеллектуальный семантический поиск и управление всеми формами представления информации, поддержку принятия решения на основе архивных данных и другие интеллектуальные возможности работы с данными.

Для достижения вышеперечисленных целей существуют программные комплексы, включающие в себя специализированные модули:

1. Универсальный электронный архив изображений, включая 3D-изображения, а также модуль интеллектуального семантического поиска по электронным архивам данных: обеспечивает хранение всех видов материалов и архивных документов с наиболее обширным диапазоном параметров, а также



управление и поиск данных по архиву с учётом семантики пользовательских запросов. Основная идея заключается в том, чтобы предоставить специалистам/руководителям возможность формулировать запрос на естественном языке и получать от системы полный перечень адресной (соответствующей семантике запроса) информации, «знаний». Кроме того, в рамках системы может быть реализован инновационный программный модуль-конвертер для повышения качества и безопасности хранения данных, позволяющий, с одной стороны, в пакетном режиме осуществлять миграцию форматов, данных от старых к новым, а с другой – определять и проводить предварительный анализ форматов данных непосредственно в процессе наполнения архива на предмет соответствия утверждённым политикам хранения данных. Пользователи системы смогут самостоятельно настраивать политики хранения данных таким образом, что старые или неподдерживаемые более форматы будут автоматически преобразовываться программным обеспечением системы в современные форматы.

2. Программное обеспечение в составе интеллектуальной системы поиска и принятия решений.

Важно отметить, рассматриваемая интеллектуальная системы предназначена для использования любыми организациями любых форм собственности, государственными, бюджетными, внебюджетными учреждениями, коммерческими компаниями для организации работы с электронными архивами данных, управления электронными архивами, а также обеспечения поддержки принятия решений за счёт интеллектуального поиска по базам данных электронного архива с пониманием системой семантики пользовательского запроса. Кроме того, внедрение электронного архива на основе интеллектуальной системы предоставляет возможность обеспечить в электронном виде:

- механизм ввода и создания электронных документов;
- просмотр документов;
- печать документов;
- быстрый поиск документов по различным параметрам, как по иерархическому каталогу документации, так и по атрибутам документа;
- разграничение прав доступа к электронным документам;
- невозможность несанкционированного удаления документов;
- надёжность хранения, резервное копирование электронных документов;
- отслеживание историй изменений документов;
- возможность интегрирования с уже существующими системами.

При этом эффективность использования архивного фонда многократно повышается за счет систематизации больших массивов информации, возможности одновременной работы различных специалистов с одними и теми же документами, быстрого создания и анализа по заданному признаку подборок документов. Также обеспечивается удаленный доступ к документам сотрудников территориально распределенных подразделений. Кроме того, решается проблема сохранения целостности документации, так как при



правильной организации хранения электронных документов невозможна их потеря, удаление и модификация.

Основными преимуществами внедрения интеллектуальной системы управления электронными архивами являются:

- обеспечение высокого уровня управления информационными ресурсами электронного архива;
- обеспечение удобного и интуитивно понятного использования информационного ресурса;
- возможность интеграции с уже существующими информационными системами;
- возможность интегрирование с имеющимися базами данных документов;
- оптимальные требования к серверному оборудованию;
- возможность хранения электронных документов различного типа общим объемом до сотен терабайт;
- использование в качестве автоматизированного рабочего места «тонкого» web-клиента, возможность входа с различных устройств, минимальные требования к рабочим станциям пользователей;
- удобство работы с системой, «дружественный» интерфейс;
- комплексное обеспечение информационной безопасности;
- предоставление возможности по аттестации системы на предмет соответствия требованиям по информационной безопасности, работы с персональными данными и конфиденциальной информацией.

В архитектуру рассматриваемой интеллектуальной системы входят следующие элементы:

- индексная база данных (на свободных решениях Oracle), поддерживающая автоматизированный импорт данных из других баз данных;
- серверная часть подсистемы (Java), включающая в себя сервер приложений и Web-сервер;
- автоматизированное рабочее место пользователя подсистемы (Web-клиент), позволяющее работать с подсистемой с любых устройств, имеющих web-браузер и подключение к сети интернет.

Так же, за счёт использования облачной инфраструктуры станет возможным обеспечение поддержки в актуальном состоянии всей значимой информации, размещенной на облачной платформе. Реализация интеллектуальной системы управления электронными архивами на основе технологии облачных вычислений позволит обеспечить следующие возможности:

- возможность не устанавливать копию программного обеспечения на рабочие места конечных пользователей – доступ к программному обеспечению осуществляется через веб-браузер (веб-доступ) или «тонкий» клиент;
- поддержка всех распространённых веб-браузеров, включая: Microsoft Internet Explorer, Mozilla Firefox, Apple Safari, Google Chrome, Opera, Maxthon, SeaMonkey и другие;



– кросс-платформенность, т.е. работа в разных операционных системах, с разными СУБД, на аппаратных платформах разных производителей.

Все вышеперечисленное свидетельствует о том, что организация единого электронного архива на основе рассматриваемой интеллектуальной системы управления электронным архивом позволит обеспечить электронный документооборот и актуализацию имеющейся электронной информации, а также непрерывность организационных процессов, связанных с использованием архивных документов.

Использование интеллектуальной системы управления электронными архивами будет способствовать достижению следующих научно-технических целей:

- Разработка научно-технического задела по перспективным технологиям в области информационно-телекоммуникационных систем, исследования и разработки по которым осуществляются в соответствии с направлениями стратегической программы исследований Технологической платформы «Национальная программная платформа».
- Создание научно-технических основ построения интеллектуальной программной платформы для создания систем поддержки принятия решений на основе электронных архивов данных.
- Формирование новых знаний в области управления корпоративным контентом за счёт применения передовых технологий Semantic Web и онтологий, а также проектирование и разработка улучшенных программных решений для создания электронных архивов данных.

Указанные цели являются одинаково значимыми как для организаций реального сектора экономики (крупных промышленных или производственных предприятий и др.), так и для органов государственной власти и бюджетных учреждений, работающих с большими массивами данных и/или предоставляющими гражданам государственные услуги в электронной форме.

Кроме того, в основе интеллектуальной системы управления электронными архивами лежат следующие технологии и принципы, способные улучшить функциональные, потребительские, стоимостные и другие показатели в сравнении с имеющимися аналогами:

- Современные технологии Semantic Web (Web 3.0). Технологии семантического веб используются при создании модуля интеллектуального семантического поиска данных по архиву, обеспечивающего возможность пользователям формировать запрос к базе данных электронного архива на естественном языке и на выходе получать лишь ту информацию и «знания», которые максимально точно и полно соответствуют семантике пользовательского запроса.
- Принципы краудсорсинга, реализуемые средствами разрабатываемого программного обеспечения и обеспечивающие возможность привлечения ресурсов «толпы» к переводу архивов в электронный вид.
- Интероперабельность данных при работе и организации хранения



данных в электронных архивах. Реализуется путём создания модуля-конвертера, позволяющего, с одной стороны, в пакетном режиме осуществлять миграцию форматов, данных от старых к новым, а с другой – определять и проводить предварительный анализ форматов данных непосредственно в процессе наполнения архива на предмет соответствия утверждённым политикам хранения данных. Пользователи системы смогут самостоятельно настраивать политики хранения данных таким образом, что старые или неподдерживаемые более форматы будут автоматически преобразовываться программным обеспечением системы в современные форматы.

Заключение

Организация интеллектуальной системы управления электронными архивами позволяет обеспечивать формирование электронного фонда документов и автоматизация процессов работы с ними, а так же позволяет осуществлять создание, управление и работу с электронными архивами государственных и коммерческих организаций, включая интеллектуальный семантический поиск и управление всеми формами представления информации, поддержку принятия решения на основе архивных данных и другие интеллектуальные возможности работы с данными.

О.П. Солдатова, Д.З. Иваев

ИССЛЕДОВАНИЕ ЭФФЕКТИВНОСТИ АЛГОРИТМОВ ОБУЧЕНИЯ МНОГОСЛОЙНОГО ПЕРСЕПТРОНА ПРИ РЕШЕНИИ ЗАДАЧИ КЛАССИФИКАЦИИ

(Самарский национальный исследовательский университет
имени академика С.П. Королёва)

Целью данной работы является изучение эффективности алгоритмов обучения многослойного перцептрона при классификации данных, решение задачи классификации данных на основе модели многослойного перцептрона с одним или несколькими скрытыми слоями, сравнение полученных результатов при различных алгоритмах обучения сети.

Сигнал, подаваемый на вход i -го нейрона k -го слоя, проходит через сумматор, в котором выполняется взвешенное суммирование элементов входного сигнала:

$$u_i^{(k)} = \sum_{j=1}^m x_j w_{ij}^{(k)}, \text{ где } m - \text{число нейронов } k-1 \text{-го слоя; } i = 1..n_k.$$

В случае входного слоя $x_j - j$ -ый компонент вектора входных данных; m - размерность вектора x ; n_k - число элементов в слое k ; $w_{ij}^{(k)}$ - соответствующий весовой коэффициент.

Выход каждого элемента рассчитывается по следующей формуле:



$$y_i^{(k)} = f(u_i^{(k)}), \text{ где } f(u_i^{(k)}) = 1/(1 + e^{-\beta x}) \text{ (здесь } \beta = 1).$$

Этап обучения нейронной сети заключается в подборе таких значений весов $w_{ij}^{(k)}$ для всех слоев сети, чтобы при заданном векторе x , получить на выходе значения сигналов y_i , позволяющие определить принадлежность данного вектора к одному из классов. Для решения этой задачи используется градиентный алгоритм наискорейшего спуска с методом обратного распространения ошибки и эвристические алгоритмы RPROP и Quickprop [1].

Смысл градиентного алгоритма наискорейшего спуска с обратным распространением ошибки заключается в вычислении параметра коррекции δ_i для каждого элемента нейронной сети $\delta_i = (y_i - d_i) \frac{\partial y_i}{\partial u_i}$ (в случае выходных нейронов), где d_i - ожидаемое значение на i -м выходе, или $\delta_i^{(k)} = (\sum_j \delta_j^{(k+1)} w_{ji}^{(k+1)}) \frac{\partial y_i}{\partial u_i}$ (для нейронов скрытых слоев k), и коррекции весов по формуле $w_{ij}^{(k)} = w_{ij}^{(k)} + \Delta w_{ij}^{(k)}$, где $\Delta w_{ij}^{(k)} = -\eta \delta_i^{(k)} u_j^{(k-1)}$. Приведенные вычисления выполняются для всех векторов обучающей выборки.

Алгоритмы эвристического типа представляют собой в основном модификацию методов наискорейшего спуска или сопряженных градиентов. Подобные модификации широко известных алгоритмов связаны с внесением в них некоторых изменений, ускоряющих процесс обучения.

Алгоритм QuickProp содержит элементы, предотвращающие заикливание в точке неглубокого локального минимума.

Вес w_{ij} на k -ом шаге алгоритма изменяется согласно правилу:

$$\Delta w_{ij}(k) = -\eta_k \left[\frac{\partial E(w(k))}{\partial w_{ij}} + \gamma w_{ij}(k) \right] + \alpha_{ij}^k \Delta w_{ij}(k-1)$$

Коэффициент γ , обычно имеющий малую величину (порядка 10^{-4}), приводит к уменьшению весов вплоть до разрыва соответствующих взвешенных связей.

Константа η_k - в данном алгоритме может иметь ненулевое значение η_0 (как правило, $0.01 < \eta_0 < 0.6$ на старте процесса обучения, когда $\Delta w_{ij}(k-1) = 0$

или когда $\Delta w_{ij} \left[\frac{\partial E(w(k))}{\partial w_{ij}} + \gamma w_{ij}(k) \right] > 0$, либо нулевое значение - в противном случае. Фактор момента коэффициент подбирается индивидуально для каждого веса по правилу:

$$\alpha_{ij}^k = \begin{cases} \alpha_{\max}, & \text{где } \beta_{ij}(k) > \alpha_{\max}, \text{ где } S_{ij}(k) \Delta w_{ij}(k-1) \beta_{ij}(k) < 0; \\ \beta_{ij}(k) & \end{cases}$$

причем:

$$S_{ij}(k) = \frac{\partial E(w(k))}{\partial w_{ij}} + \gamma w_{ij}(k), \quad \beta_{ij}(k) = \frac{S_{ij}(k)}{S_{ij}(k-1) - S_{ij}(k)}, \text{ а}$$