

УДК 681.3

ЭЛЕКТРОННЫЙ ДОКУМЕНТООБОРОТ

А.И. Прокопенко, К.И. Прокопенко

Харьковский национальный университет радиоэлектроники

e-mail: anton.prokopenko@gmail.com

Электронный обмен данными - это реальность, с которой сегодня сталкивается практически каждый. Информационные системы, компьютерные сети, электронная почта – и это далеко не полный перечень средств, с помощью которых происходит обмен данными в электронном виде.

В последнее время появились и получили распространение новые инструментальные средства эффективного обеспечения управленческих процессов. В частности программном обеспечении, предназначенном для обработки управленческих документов. Здесь, прежде всего, следует упомянуть программное обеспечение классов "системы управления документами" и "системы управления деловыми процессами".

Такие системы представляют собой программные комплексы, применимые для решения ряда задач, в том числе и для построения корпоративных систем электронного документооборота. Документооборот – это движение документов между их составителями и исполнителями по информационным технологическим цепочкам. В рамках автоматизации процесса обработки документа в организации с момента его создания или получения до момента отправки корреспонденту или завершения исполнения и списания в дело должно быть обеспечено решение следующих функций:

- регистрация входящих в организацию документов, исходящих из организации документов и внутренних документов;
- учет резолюций, выданных по документам руководством организации, и постановка документов на контроль;
- централизованный контроль исполнения документов;
- списание документов в дело;
- ведение информационно-справочной работы;
- формирование делопроизводственных отчетов по организации в целом.

Использование системы электронного документооборота позволяет организовать передачу данных о ходе исполнения документов в электронном виде, что качественно меняет организацию контроля исполнения документов. Карточки зарегистрированных централизованно документов с резолюциями руководства рассылаются в электронном виде сотрудникам соответствующих подразделений. Они дополняют их резолюциями по исполнению документов, выдаваемыми руководителями структурных подразделений. По мере появления данных о ходе исполнения документов эти данные вносятся в систему.

При этом система автоматически отслеживает наступление даты предварительного уведомления о приближении срока исполнения и наступление самого этого срока. Заинтересованные пользователи системы информируются о названных сроках.

Также значительно видоизменяется процесс согласования проектов документов, в рамках которого сотрудники, участвующие в процессе согласования, получают возможность обмениваться электронными версиями согласуемых проектов. Такая технология позволяет сократить время, затрачиваемое на передачу проектов в бумажном виде.

Система электронного документооборота обязательно включает текущий электронный архив, который решает проблемы оперативного доступа к информации и наличия возможности одновременного использования документа несколькими сотрудниками. Такая форма организации хранения значительно снижает вероятность

потери информации и повышает оперативность работы за счет сокращения времени поиска нужного документа. Хранение текстов документов в электронном виде позволяет реализовывать полнотекстовый поиск, что открывает принципиально новые возможности при ведении информационно-справочной работы, например, позволяет делать тематические подборки документов по их содержанию. Использование электронного архива избавляет от необходимости создавать фонд пользования архивных документов, так как по запросу в любой момент может быть выдана электронная копия документа.

Хотелось бы акцентировать большее внимание на вопросе поиска документации: одной из функций систем электронного документооборота является, возможность хранения документов и соответственно их поиска при необходимости. Существует много средств и алгоритмов хранения документов. Также как и средств и алгоритмов их поиска в пределах системы электронного документооборота. Следует отметить, что применение тех или иных алгоритмов «напрямую» без какой-либо адаптации к области применения зачастую приводит к тому, что полученный результат далеко не всегда соответствует ожидаемому, то есть в случае системы электронного документооборота поиск документации может оказаться достаточно долгим процессом.

Один из способов повышения эффективности поиска документов является возможность поиска в некой выборке документов, вместо всей совокупности, ограниченной рамками определенной тематики. Достаточно простой, с одной точки зрения, способ повысить эффективность поиска, но с другой – формирование этих тематик и классификация документов также требуют затрат. Чтобы избежать излишних затрат человеческих ресурсов, процесс классификации вполне реально «переложить на плечи» техники, т.е. в случае электронных систем это применить некоторые алгоритмы классификации (K-NN, Naïve Byes, или что-то более интересное), что не лишит процесс временных затрат, однако потребует меньшего участия человека (что в пору финансового кризиса в частности, смею заметить, весьма экономически выгодно).

Выделение тематик документов и организация поиска с учетом конкретной тематики сам по себе процесс не сложен, однако... Когда дело касается автоматизации формирования тематик, а особенно оперативного их формирования на основе ново-полученного документа, алгоритм определения тематики усложняется.

Одним из возможных способов определения тематики, является способ, при котором определённый набор тематик задаётся изначально с неким базовым набором слов принадлежащих исключительно или в большинстве случаев к соответствующей тематике. А затем, в процессе определения тематик новых документов, словари пополняются новыми словами, предположительно относящимися только к заданной тематике. Например, на основе результатов классификации документов одним из уже известных методов кластеризации.

Рассмотрим алгоритм на примере. Алгоритм формирования тематик может иметь следующие пункты:

Обучение (рис. 1):

1. Классификация документов методами K-NN или Naïve Byes для выделения тематик.
2. Построение словарей для каждой тематики, в рамках заданной предметной области (ограничивается обучающей выборкой)
3. Сохранение словарей в хранилище данных.

Классификация (рис. 2):

1. Анализ документа, расчет функции принадлежности его лексем к имеющимся в хранилище данных – тематикам.

2. По результатам функции принадлежности к той или иной тематике определить наиболее вероятную тематику.

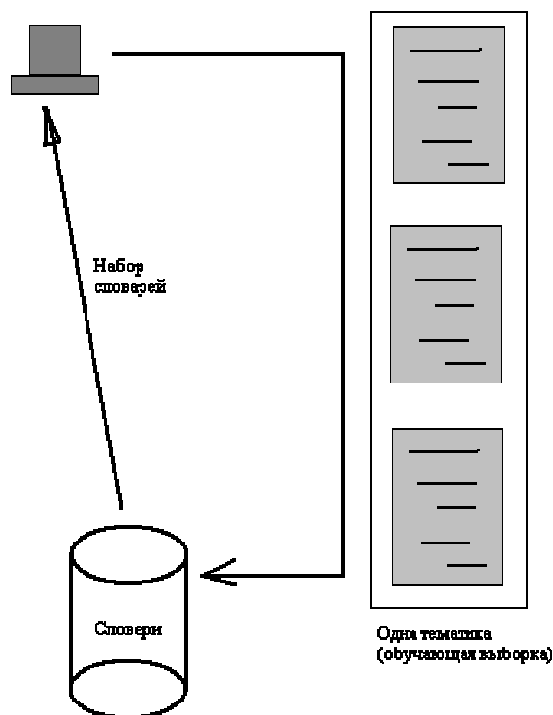


Рисунок 1 – Схема обучения

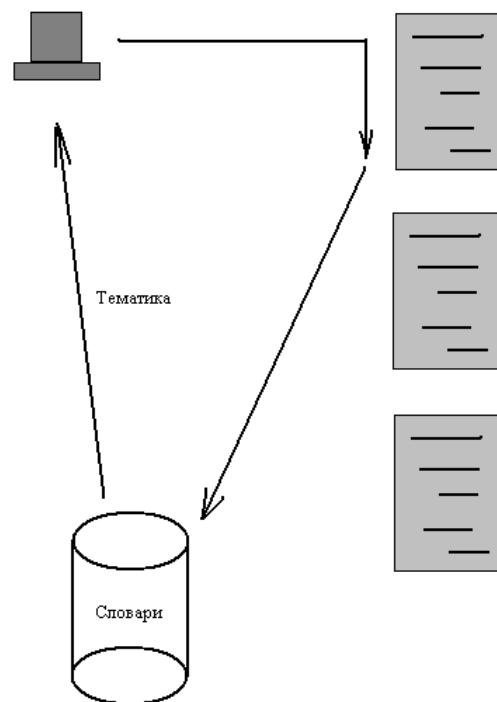


Рисунок 2 – Схема классификации

Как можно заметить приведенный алгоритм достаточно просто, однако применение подхода оперативного формирования тематик очевидно повысит производительность любой системы работа которой в определенной степени привязана к поиску документов того или иного типа. Данный подход вполне может применяться как для систем электронного документооборота, так и для Web-ресурсов или документов отдельно взятого Web-ресурса.

Список используемой литературы

1. Автоматизированные информационные технологии в экономике. Под ред. Титаренко. М.: ИНФРА-М, 1998. – 258 с.
2. Кузнецова Т.В. Делопроизводство (Документационное обеспечение управления). — М.: ЗАО «Бизнес-школа «Интел-Синтез», 2000. - 818 с.
3. Остелло М. Влияние новой информационной технологии на управленческий процесс // Проблемы теории и практики управления. – 2002. – №6. - С.49.
4. Печникова Т.В., Печникова А.В. Практика работы с документами в организации: Учеб. пособие для вузов. – М.: ЭМОС, 1999. – 208с.
5. Шафрин Ю. Информационные технологии, - М., ООО " Лаборатория базовых знаний", 2000.