

ИНФОРМАЦИОННАЯ ПОДДЕРЖКА ПРОЦЕССА ОБРАБОТКИ ДОКУМЕНТОВ В СИСТЕМЕ ЭЛЕКТРОННОГО ДОКУМЕНТООБОРОТА

При принятии решений по документам руководителю часто требуется провести параллельный анализ других материалов, связанных с рассматриваемым [1]. Поиск этой информации и формирование информационного поля рассматриваемого документа необходимо осуществлять в системе электронного документооборота (СЭД) в процессе технологической обработки поступившего документа таким образом, чтобы к этапу принятия решения руководителю был предоставлен необходимый набор информационных материалов с желаемой степенью релевантности.

В существующих системах документооборота подобный режим функционирования не поддерживается [2,3]. Поиск необходимых документов осуществляется традиционными методами на основании формируемых пользователем поисковых предписаний в момент обработки документа. Целесообразность и эффективность этих методов для ИСС несомненна. В СЭД, при формализованном представлении документов и легко устанавливаемых регламентированных связях между ними, поисковое предписание в основной его части достаточно просто автоматизировать, что сокращает затраты времени на запрос. Более существенно то, что в случае групповой обработки документа в рамках традиционного способа поиска необходимых материалов единый технологический процесс обработки разрывается на отдельные, обособленные этапы. При этом утрачивается преемственность и взаимодополняемость стадий обработки, а информационное поле документа формируется каждый раз заново и не сопровождает его при перемещении по технологической цепочке.

Для устранения отмеченных недостатков предлагается технологическая модель формирования информационного поля документа, в рамках которой система анализирует информационные материалы перед их включением в информационное поле рассматриваемого документа, осуществляя этот анализ в процессе поиска с помощью идентификационных карт, формируемых посредством механизма предварительной динамической идентификации документов или идентификации в процессе обработки.

Технология поиска для формирования информационного поля включает в себя анализ информации на предмет релевантности идентификационной карте рассматриваемого материала и его содержательной части. Запрос на поиск информации, необходимой для подготовки принятия решения по проблемной ситуации, содержит статическую и динамическую составляющие. Статическая часть, связанная непосредственно с данной проблемой, в максимальной степени формируется автоматически, пополняясь на этапах обработки проблемной ситуации. Динамическая часть

в большей степени подбирается системой по специальному запросу, составленному пользователем, и частично, в автоматическом режиме, учитывая направление деятельности данного структурного подразделения.

Каждому документу ставится в соответствие его идентификационная карта, в которой находятся его идентификационные признаки. Данные, идентифицирующие информацию, вносятся в карту, как правило, автоматически, в процессе обработки информации.

Структура идентификационной карты разбита на блоки структурной, адресной и тематической идентификации.

В первом блоке отражается разбивка всей имеющейся в СЭД информации на структурные уровни, объединяющие крупные блоки информации, характеризующие групповыми функциональными признаками.

Во втором блоке отражаются адресные данные по поступающей информации. Применяется метод фасетного индексирования. В качестве идентифицирующих элементов используются ключевые слова. Выбор идентификационных ключей производится из установленного перечня для каждого ключевого поля и может зависеть от идентификационной цепочки, сформированной на уровне структурной идентификации.

Третий блок характеризует тематическую направленность информации. В поля блока вносятся ключевые слова. В тексте документа в режиме тематической идентификации автоматически или пользователем выделяются слова, термины, характеризующие этот документ. Если автоматически, - то ключевые слова из идентификационных перечней, сформированных ранее. Если пользователем, - то, дополнительно, из текста документа.

Идентификация информации производится в процессе регистрации и обработки на одном или нескольких АРМ. Часть данных по идентификации информации может вноситься специально при поступлении информации в СЭД после регистрации адресных данных.

Идентифицированная, согласно изложенной процедуре, информация заносится в идентификационную базу данных. Информационное поле документа создается в рамках двух этапов:

1) Первый этап - формирование базового информационного поля рассматриваемого материала в результате поиска в автоматическом режиме с помощью идентификационных карт.

2) Второй этап - пополнение базового информационного поля документа в результате:

а) поиска по ключам, когда используются ключи из блоков идентификационной карты документа.

б) контекстно-зависимого поиска, когда используются ключевые слова из текста рассматриваемого материала;

Два последних способа поиска выполняются с участием пользователя. Из полученного перечня материалов для дальнейшей работы формируется группа фактически релевантных (соответствующих информационной потребности).

При поиске по ключам могут использоваться варианты поиска либо с помощью всех трех блоков идентификационной карты информационного материала, либо в любом другом возможном сочетании. При использовании блока тематической идентификации учитывается то, что существует промежуточный пользовательский файл ключевых слов, создаваемый на этапе работы пользователя с СЭД, с которым работает конкретный пользователь при поиске информации. Это те базовые ключевые слова, которые были внесены пользователем в качестве терминов, идентифицирующих тематику поступающей и интересующей его информации или документов.

В случае реализации первого варианта контекстно-зависимого поиска используется текст рассматриваемого документа. Из отмеченных слов, характеризующих данную информацию, выделяются базовые, по которым подбирается перечень информационных материалов и ранжируется по количеству ключевых слов в документе. Из этого перечня, далее, выбираются те материалы, где задействованы также остальные слова, отмеченные по тексту.

В случае реализации второго варианта контекстно-зависимого поиска используется пользовательский файл ключевых слов. Вариант используется, когда пользователь не работает с конкретным документом или ему нужны материалы, которые нельзя отразить только с помощью текста документа.

Таким образом, предложенный алгоритм позволяет разработать процедуру идентификации информации для создания идентификационных баз данных и реализовать эффективную технологию информационного обеспечения функционирования руководителя при обработке документов в СЭД во взаимодействии с подразделениями и вспомогательными службами организации в рамках единого технологического процесса.

1. Сумароков Л.Н., Тимофеева О.В. Методы организации работы руководителя. - М: Издательство стандартов, 1990.

2. Михайлов С. DOCFLOW 95: задачи дня - документооборот.//PC Magazine/Russian Edition, 1996, № 5.

3. Афанасьев А. Частные реализации систем документооборота.// Открытые системы, 1997, №1.