

Лексикологический синтез слабоформализуемых документов

Б.В. Черников

Аннотация. В статье рассматривается технология формирования слабоформализуемых документов, создаваемых с помощью автоматизированного лексикологического синтеза. Экспериментальная проверка рассматриваемого способа показала возможность существенного сокращения трудозатрат (в зависимости от вида документа – от 2-3 до 5-7 раз) при создании полнотекстовых документов. Дополнительным достоинством технологии является возможность автоматизированного формирования документа с учетом установленного набора реквизитов и сокращения ошибок при вводе информации.

Ключевые слова: лексикологический синтез, слабоформализуемые документы, опорные слова.

Введение

Процесс выработки управленческих решений включает в себя несколько этапов, один из которых заключается в сборе и оценке полученных сведений. Эта информация должна быть представлена в соответствии с установленными требованиями. Необходимо также, чтобы собранные сведения были в максимальной степени удобны для восприятия на следующих этапах. Поэтому возникает проблема закрепления информации в документе, то есть выбора способа ее документирования. Процесс формирования документов различного характера является обязательным элементом при возникновении необходимости сохранения информации о каком-либо процессе или событии.

На современном этапе процесс документирования информации должен отвечать ряду требований, к которым, несомненно, необходимо отнести следующие:

- данные должны быть максимально формализованы в целях обеспечения автоматизированной обработки сведений, содержащихся в документе;
- создание документов должно занимать минимум времени при сохранении требований к

информации, необходимых для принятия управленческого решения.

Дополнительные требования к процессу документирования устанавливает принятый в начале 2008 г. стандарт MoReq2 [7], рекомендации которого, несомненно, будут приниматься во внимание при разработке систем электронного документооборота на российском рынке. Одним из требований, выдвигаемых MoReq2, является удобство процесса создания документов.

Актуальность современных исследований в области совершенствования средств и методов формирования документов возрастает в связи с реализацией в государстве Федеральной целевой программы "Электронная Россия" и началом формирования электронного правительства, которое должно быть создано как на региональном, так и на федеральном уровнях. Значение данного направления подтверждается тем, что вопросы его развития рассматриваются на самом высоком уровне, включая Федеральный Совет по развитию информационного общества и Государственный совет.

Управленческие документы в подавляющем большинстве можно охарактеризовать как слабоформализуемые, поскольку при высокой степени вариативности содержания, зависящего от

конкретной ситуации, они должны фиксировать все разнообразие деталей соответствующих сфере применения.

Слабоформализуемые документы – полнотекстовые, табличные либо смешанные документы, содержание которых существенным образом связано с произвольной, меняющейся от конкретной ситуации структурой. Это документы, обладающие достаточно высокой степенью вариативности. В связи с этим содержательная структуризация слабоформализуемых документов может требовать детализации как взаимосвязи, так и взаимной зависимости композиции текста вплоть до атомарных значений – фрагментов фраз, слов и даже частей отдельных слов.

1. Современное состояние процесса создания документов

Традиционный способ формирования текстовых документов предполагает ввод информации с клавиатуры путем прямого набора текста. В этом случае трудозатраты на формирование документов оказываются весьма существенными. Кроме того, набору текста часто сопутствует появление орфографических и синтаксических ошибок, вызванных, например, техническими погрешностями или недостаточно высоким уровнем грамотности исполнителя документа. Дополнительной проблемой часто становятся погрешности в оформлении документов, которые автор, создавая тот или иной документ, вынужден изучать, чтобы соблюсти установленные нормативы, определяющие форму и содержание документа.

В настоящее время выделяются, в основном, два вида унификации текстов: типизация и трафаретизация. В случае типизации создаются сборники типовых текстов, на основе которых формируется текст конкретного документа. Анализ показывает, что в документах существуют элементы сведений, которые могут быть представлены типовыми текстами. Раньше этот метод (метод стандартных фраз) применялся в организационно-распорядительной документации для стандартизации деловой переписки. Недостатком данного подхода является необходимость ручной корректировки типовых текстов при конкретизации управленческой ситуации.

При трафаретизации текстов информацию подразделяют на постоянную и переменную. Постоянная информация вносится в бланк документа при его изготовлении, а переменная – в процессе составления конкретного документа. К сожалению, этот способ также не свободен от недостатков, одним из которых следует отметить невозможность вариации размеров ответных текстовых полей в широких пределах, что актуально при существенных изменениях объемов фрагментов документа.

Одним из направлений, позволяющих сократить время формирования документов и снизить вероятность появления ошибок, является использование специализированных карт, содержание которых основано на применении формализованной информации. Однако подобный способ предполагает использование ограниченного диапазона элементов, входящих в состав документа. В этой связи он находит достаточно широкое применение, в основном, при формировании учетных документов. Другим направлением является использование шаблонов документов в распространенных текстовых редакторах. Основная задача шаблонов – закрепление способов представления документарной информации, однако ввод сведений при использовании шаблонов осуществляется вручную.

Вопросы информационного и документационного обеспечения процессов, протекающих в деловой деятельности организаций, неоднократно рассматривались специалистами в различных аспектах. В работах [1, 2, 4, 6, 11, 12] проводился анализ систем информационного и документационного обеспечения предприятий и организаций. В основном, внимание исследователей акцентировалось на системных проблемах и организации управления и обмена документами. В то же время вопросы автоматизированного формирования текстовых документов, обеспечивающих процесс принятия управленческих решений, как и проблемы разработки информационных технологий синтеза текстовых документов, рассмотрены недостаточно глубоко и полно.

2. Сущность лексикологического синтеза документов

Естественным оказывается стремление пользователей, с одной стороны, снизить вероят-

ность появления ошибок и сократить трудозатраты, необходимые для формирования текстовых документов, а с другой стороны – сохранить легкость чтения сформированного документа на уровне составленного традиционным способом прямого ввода информации при соблюдении требований по оформлению документа.

Весьма эффективным путем удовлетворения этих запросов является применение способа автоматизированного формирования документов, основанного на принципе лексикологического синтеза [8-10]. Информация, содержащаяся в подготавливаемом тексте документа, подвергается более глубокой классификации, что позволяет при создании текста документа сократить объем неунифицированной информации, формировать текст документа, наиболее точно характеризующий каждую конкретную управленческую ситуацию.

Основой способа является тот факт, что определенная сфера человеческой деятельности сопровождается унифицированным набором документов. Любой документ, фиксирующий информацию об определенной управленческой ситуации, содержит информацию двух классов – постоянную и переменную. Из текста документа можно выделить характерную для него постоянную информацию. Постоянная информация дополняется переменной, причем в заранее унифицированных текстах документов переменная информация может принадлежать предопределенному множеству вариантов. Если это множество вариантов предварительно, путем экспертной оценки, собрать воедино, переменная информация может быть отнесена к разряду переменной унифицированной. Современные инструментальные средства позволяют осуществлять хранение этих, заранее зафиксированных, вариантов формулировок. В последующем, при создании документа, реализуется автоматическое внедрение в документ постоянной информации, а также автоматизированное включение переменной информации, выбранной исполнителем документа из сохраненного множества.

Под термином *лексикологический синтез* следует понимать в данном случае формирование текстовых фрагментов компьютерной системой путем синтеза (создания) фраз на основе исполь-

зования набора опорных (ключевых) слов, комплектуемого по результатам предварительно проведенной содержательной унификации документа, с автоматическим связыванием фрагментов и отдельных слов текста в соответствии с правилами орфографии и лексикологии. Использование понятия "лексика" в данном случае является оправданным, поскольку оно определяет "словарный состав языка, набор используемых в данном языке ключевых слов" [5]. Особенностью применения этого термина является его распространение лишь на содержание конкретного вида документа, поэтому речь следует вести только о формировании необходимого набора слов, определяющего возможность составления определенного документа.

В ходе унификации документов по содержанию формируется набор (по возможности – полный, то есть предусматривающий все возможные варианты) формулировок, которые могут присутствовать в документе. При этом необходимо учитывать разнообразие информации, которая может содержаться в каждом отдельном экземпляре формируемого документа. Следовательно, по отношению к отдельно взятому экземпляру документа сформированный набор формулировок может и должен быть избыточным, то есть содержать даже большее количество фрагментов текста, чем это необходимо при составлении единичного экземпляра документа. Сформированный набор формулировок сохраняется в упорядоченном состоянии в виде индексированной совокупности элементов (например, в виде базы знаний). Каждой формулировке ставится в соответствие основное слово, выбор которого однозначно определяет наличие конкретной формулировки в документе. Такие слова называются *опорными* и образуют основу лексикологической схемы формируемого документа. Сопоставление индексов сохраненного набора формулировок опорным словам позволит обеспечить однозначную взаимосвязь этих элементов. Полный перечень опорных слов с учетом их взаимосвязей образует лексикологическое дерево документа, "прохождение" по ветвям которого обеспечит выбор формулировок, используемых в документе. При этом выбор тех или иных опорных слов будет означать необходимость

внедрения в документ конкретных вариантов текстовых фрагментов. Фактически текст документа формируется путем выбора необходимых заготовок из числа сохраненных формулировок. Задача программных средств заключается в обеспечении необходимой связи между словами в используемых фразах, что может потребовать некоторого изменения отдельных слов в формулировках в целях их согласованного (с точки зрения правил синтаксиса) применения. К сожалению, в отдельных случаях определение полного набора формулировок текстовых фрагментов может оказаться затруднительным. В этой связи целесообразно предусматривать возможность ввода свободных формулировок прямым набором текста. Имеющийся опыт применения лексикологического синтеза слабоформализуемых документов показывает, что при достаточно внимательном исследовании возможной структуры документа объем подобных фрагментов, как правило, оказывается незначительным и не оказывает существенного влияния на производительность труда сотрудника, выполняющего задачу формирования документа. Структура процесса подготовки документа к автоматизированному формированию приведена на Рис. 1.

Процесс инициируется принятием решения о необходимости автоматизированного формирования слабоформализуемого документа с помощью лексикологического синтеза. Рассмотрим содержание этапов процесса подготовки в соответствии с приведенной структурой.

1. В соответствии с задачами, решаемыми в конкретной сфере деятельности, устанавливается совокупность реквизитов, расположенных в определенной последовательности, и определяется местоположение каждого элемента информации. Определяются зоны унифицированной формы документа, предназначенные для закрепления ее в технических средствах хранения документов, а также для нанесения специальных изображений (например, логотипа организации). Устанавливается набор реквизитов, необходимых и достаточных как для идентификации автора официального документа, так и для придания документу юридической силы [3]: наименование автора документа, вид документа, дата, номер, место составления или из-

дания. В дальнейшем такие реквизиты в целях экономии времени постоянно вносятся в бланк документа программными средствами. После проведения унификации формы документа производится ее сохранение в базе знаний, содержащей унифицированные формы документов, подлежащих автоматизированному формированию.

2. Проводится классификация информации слабоформализуемого документа. В отличие от традиционной системы классификации информации (на постоянную и переменную) выделяются информационные потоки следующих типов:

- *унифицированная постоянная информация*, внедряемая в документ автоматически с использованием базы знаний; к этому типу относится



Рис. 1. Структура процесса подготовки формирования слабоформализуемого документа

постоянная (например, наименование документа) и редко меняющаяся (наименования организации и структурного подразделения, список персонала и т.п.) информация;

- *унифицированная переменная информация*, содержащая стандартизированные и формализованные данные, хранимая в базе знаний и вводимая при формировании документа путем выбора требуемых формулировок;

- *переменная вводимая информация*, представляющая конкретизирующие сведения, как правило, для конкретного экземпляра документа (например, табличные данные, отдельные фамилии и т.п.) и вводимая с клавиатуры непосредственно при подготовке документа;

- *неунифицированная информация*, содержащая свободные формулировки и вводимая прямым набором с клавиатуры.

Необходимость выделения в отдельную категорию унифицированной переменной информации вызвана тем, что формулировки (слова, словосочетания, фразы), входящие в состав информационного потока этого типа, являются платформой, на основе которой синтезируется совокупность опорных слов. Именно эти формулировки, сопоставляемые с опорными словами, должны сохраняться в базе знаний для автоматического внедрения в документ при выборе соответствующего опорного слова. Особенность переменной вводимой информации заключается в наличии условного форматирования у данного типа данных, что должно предусматривать программную проверку вводимых сведений.

3. Следующей операцией является синтез совокупности опорных слов. Опорное слово выбирается путем анализа языковых конструкций, свойственных данному управленческому документу. В перечне возможных формулировок определяются слова (или их совокупность), привязка к которым позволяет однозначно выбрать требуемую формулировку. Каждое опорное слово в зависимости от содержания конкретного экземпляра документа может соответствовать как целой фразе, так и определенному словосочетанию или даже отдельному слову в документе. Так, например, выбор опорного слова "норма" в медицинском протоколе может означать необходимость внедрения в документ целой фразы, характеризующей соот-

ветствие описываемого параметра принятым нормам. Тем не менее, в большинстве случаев опорные слова соответствуют более коротким формулировкам, присутствие которых в документе необходимо в соответствии с описываемой ситуацией. В связи с этим количество опорных слов в значительной степени зависит от вариативности документа и требуемой степени гибкости лексикологического дерева.

Взаимосвязь опорных слов в совокупности определяет последовательность обхода маршрута формирования документа. На основе анализа структуры документа выявляются основные разделы, которые должны или могут присутствовать в документе. Условные наименования таких разделов составляют основу синтезируемой совокупности опорных слов. В рамках каждого зафиксированного раздела документа выявляются составные элементы, которые должны или могут входить в состав раздела (слово, фраза, текстовый фрагмент). Для каждого подобного составного элемента определяется опорное слово (или их совокупность), выбор которого в последующем однозначно будет определять внедрение в документ соответствующего компонента. Если фрагмент текста документа содержит значительное количество строк и всегда присутствует в документе в строго определенной последовательности построения предложений, то данный фрагмент текста определяется одним опорным словом. Однако в случаях, когда текст документа формируется из предложений, не фиксированных в строго определенной последовательности, и в каждом заново создаваемом документе наблюдаются вариации построения текста, опорных слов будет столько, сколько необходимо для однозначного определения каждого конкретного предложения или словосочетания.

4. Синтез лексикологического дерева предназначен для определения маршрута обхода опорных слов на основе учета их взаимосвязей. Полный перечень опорных слов с их взаимосвязями образует лексикологическое дерево, "прохождение" по ветвям которого обеспечивает выбор всех необходимых формулировок, используемых в документе. При этом выбор тех или иных опорных слов означает необходимость внедрения в документ конкретных вариантов тексто-

вых фрагментов. Структура лексикологического дерева сходна с композицией текста документа. Степень ветвления лексикологического дерева зависит от объема множества вариаций текста документа, определяемых его сложностью и различием документируемых ситуаций. В качестве опорного слова могут выступать различные части речи, определяющие сущность предписываемого действия. Генерация лексикологического дерева осуществляется при соблюдении критериев управления лексическими конструкциями. Опорное слово должно быть уникальным для конкретной конструкции, а при необходимости – уточняться другими опорными словами, иначе выбор требуемого текстового фрагмента может быть определен неверно. Уточнение одного опорного слова другим образует их иерархическую подчиненность в структуре лексикологического дерева. При этом связывание опорных слов имеет значение только в смысле порядка их следования в маршруте обхода лексикологического дерева при формировании документа. Проведение цикла выбора определенной последовательности опорных слов означает формирование экземпляра документа конкретного вида (Рис. 2, где утолщенной линией условно показан маршрут выбора опорных слов ϕ , отсекающего ряд боковых ветвей).

Пример лексикологического дерева для автоматизированного формирования акта готовности кафедры высшего учебного заведения к новому учебному году приведен на Рис. 3.

5. Формулировки сохраняются в базе знаний с индексированием, соответствующим опорным словам.

6. Формируется информационный алгоритм документа, который позволяет учесть вид функциональных связей отдельных опорных слов и текстовых фрагментов (формулировок) документа, определяет способ внедрения фрагмента в документ.

Связь опорного слова с унифицированным текстовым фрагментом задается графически с помощью заранее выбранных обозначений. Пример информационного алгоритма для документа "Акт готовности кафедры к новому учебному году", сформированный для автоматизированного формирования в высшем учебном заведении, приведен на Рис. 4.

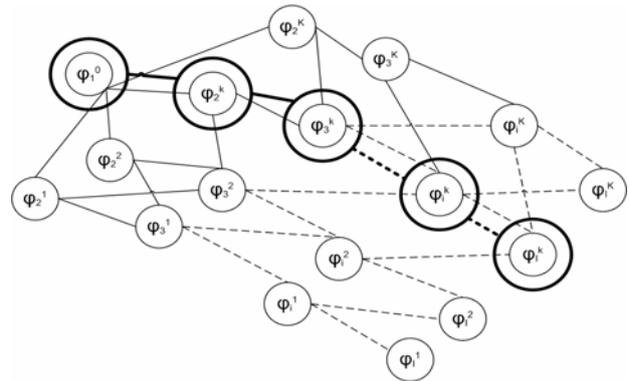


Рис. 2. Модель формирования документа при использовании дерева с отсеканием



Рис. 3. Пример лексикологического дерева

7. Синтез оперативного алгоритма (процедуры автоматизированного формирования документов) позволяет сформировать процедуру,

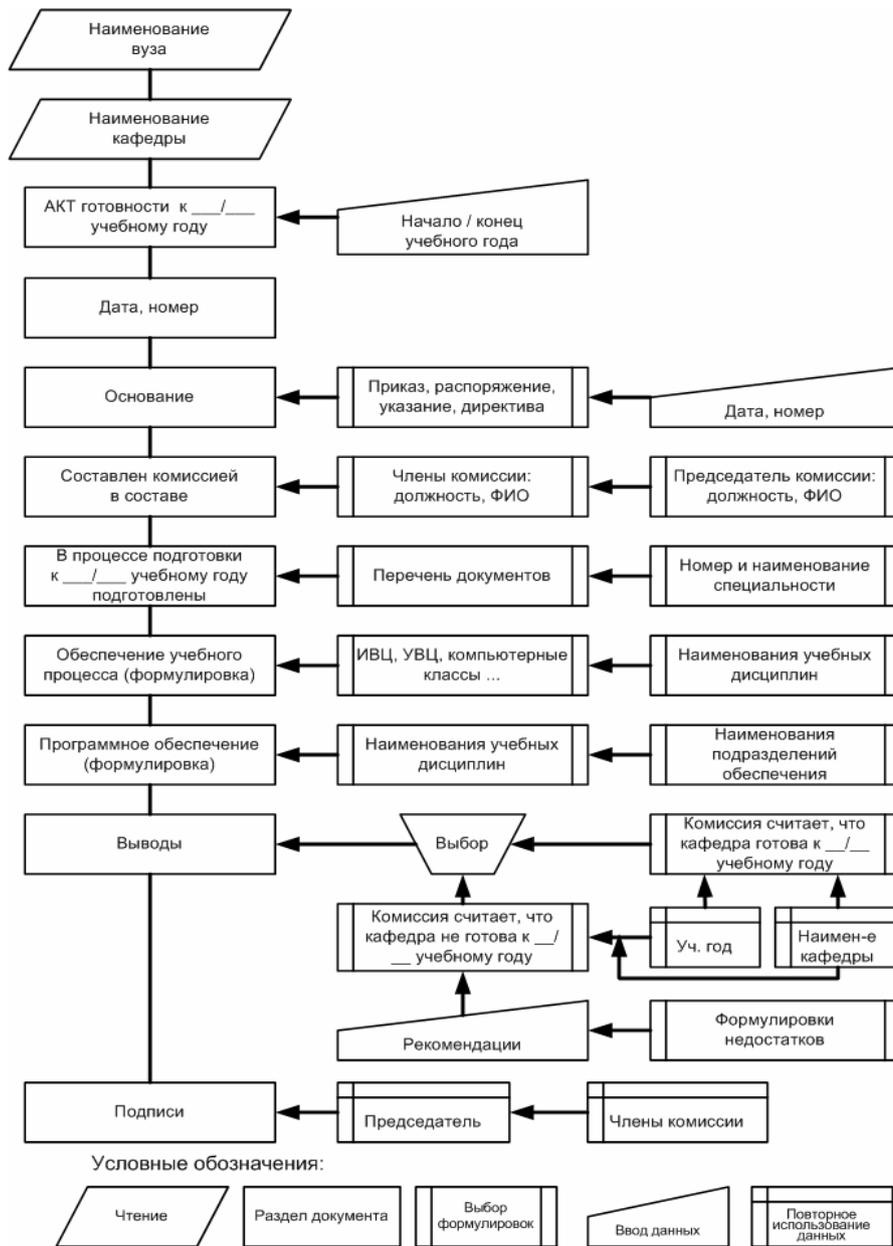


Рис. 4. Пример информационного алгоритма

представляющую собой основу для последующего создания соответствующей компьютерной программы.

В процедуре учитываются взаимосвязи блоков и фрагментов будущей программы, определяется последовательность работы пользователя при автоматизированном формировании документа. Задачей разрабатываемой процедуры является управление процессом последовательного выбора опорных слов. Кроме этого,

процедура организует пошаговое наращивание текста документа путем присоединения функционально связанных текстовых фрагментов (формулировок). Пример процедуры автоматизированного формирования документа приведен на Рис. 5 для слабоформализуемого документа "Акт готовности кафедры к новому учебному году".

8. Создание компьютерной программы формирования документа в интерактивном режиме.

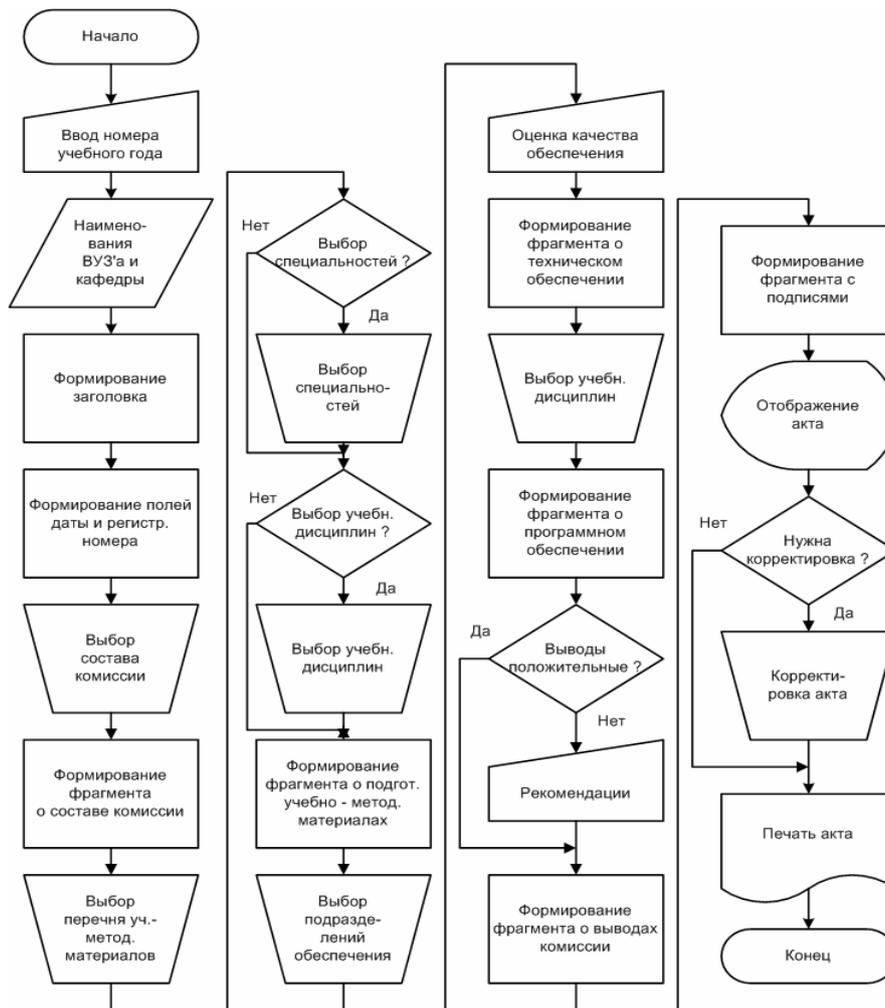


Рис. 5. Пример оперативного алгоритма формирования документа

3. Эффективность лексикологического синтеза документов

Проверка работоспособности лексикологического синтеза проводилась при формировании организационно-распорядительных документов (ОРД), документов лечебного учреждения (ДЛУ), документации высшего учебного заведения (ДВУЗ), производственно-технологических документов (ПТД) машиностроительного предприятия и других.

Анализ результатов разработки и использования программ автоматизированного формирования слабоформализуемых документов на основе лексикологического синтеза показывает, что объем информации, вводимой прямым на-

бором с клавиатуры, для конкретных экземпляров документов незначителен по сравнению с унифицированными формулировками, внедряемыми на основе выбора опорных слов.

Результаты реструктуризации информационных потоков документов для некоторых систем документации приведены в Табл. 1 для традиционной схемы формирования документов и автоматизированной схемы с использованием лексикологического синтеза.

В традиционной схеме формирования документов отсутствует унифицированная переменная информация, а общий объем содержащихся в документе сведений складывается только из постоянной, переменной вводимой и неунифицированной информации, причем доля неунифицированной информации достаточно велика.

Табл. 1. Объем информационных потоков (в процентах) по категориям данных и схемам формирования слабоформализуемых документов

Вид документа	Тип информации							
	Унифицированная постоянная		Унифицированная переменная		Переменная вводимая		Неунифицированная	
	Традиционная схема	Автоматизированная схема	Традиционная схема	Автоматизированная схема	Традиционная схема	Автоматизированная схема	Традиционная схема	Автоматизированная схема
Организационно-распорядительные документы								
Приказ (распоряжение) о наложении взыскания	15	27	–	53	40	8	45	12
Приказ (распоряжение) о переводе работника на другую работу	13	18	–	62	55	8	32	12
Приказ (распоряжение) о поощрении работника	14	26	–	49	39	11	47	14
Приказ (распоряжение) о предоставлении отпуска	15	23	–	56	36	9	49	12
Приказ (распоряжение) об увольнении работника	16	29	–	53	42	8	42	10
Приказ (распоряжение) о приеме работника на работу	17	31	–	55	38	6	45	8
Трудовой договор (контракт)	15	32	–	55	33	6	52	7
Документация лечебного учреждения								
Выписной эпикриз	6	21	–	59	12	7	82	13
Документация консилиумных совещаний	6	17	–	53	46	14	48	16
Заявки на проведение обследования пациента	18	23	–	72	32	1	50	4
Представление пациента на врачебную комиссию	7	16	–	61	18	5	75	18
Протокол ультразвукового обследования пациента	9	23	–	65	24	7	67	5
Протокол эндоскопического обследования пациента	9	22	–	63	12	9	79	6
Свидетельство о болезни	8	15	–	54	37	19	55	12
Документация высшего учебного заведения								
Акт готовности кафедры к новому учебному году	4	47	–	24	26	2	70	27
Отчет о практиках	8	80	–	7	54	5	38	8
Отчет о работе ГЭК	9	61	–	15	46	15	45	9
Отчет о работе кафедры в учебном году	3	56	–	7	17	3	80	34
Протокол заседания ГЭК	18	32	–	46	27	–	55	22

Именно неунифицированная информация предопределяет, в основном, объем вводимых вручную сведений, увеличивая трудоемкость создания документов. В то же время проведенная унификация обеспечивает существенное сокращение доли этого типа данных при одновременном возрастании объемов потоков постоянной и переменной унифицированной информации. Таким образом, в автоматизированной схеме с использованием лексикологического синтеза присутствуют все четыре типа информационных потоков.

Примеры графического отображения результатов реструктуризации информационных пото-

ков некоторых формируемых документов приведены на Рис. 6, а в виде лепестковой диаграммы для исследованных групп документов – на Рис. 7.

Анализ полученных результатов показывает, что внедрение автоматизированного способа лексикологического синтеза в практику формирования слабоформализуемых документов обеспечивает смещение групповых центров аккумуляции компонентов из областей неунифицированной информации в сторону унифицированных элементов, что предопределяет возможность автоматизированного выбора унифицированных формулировок в автоматизированной схеме. Увеличение в несколько раз

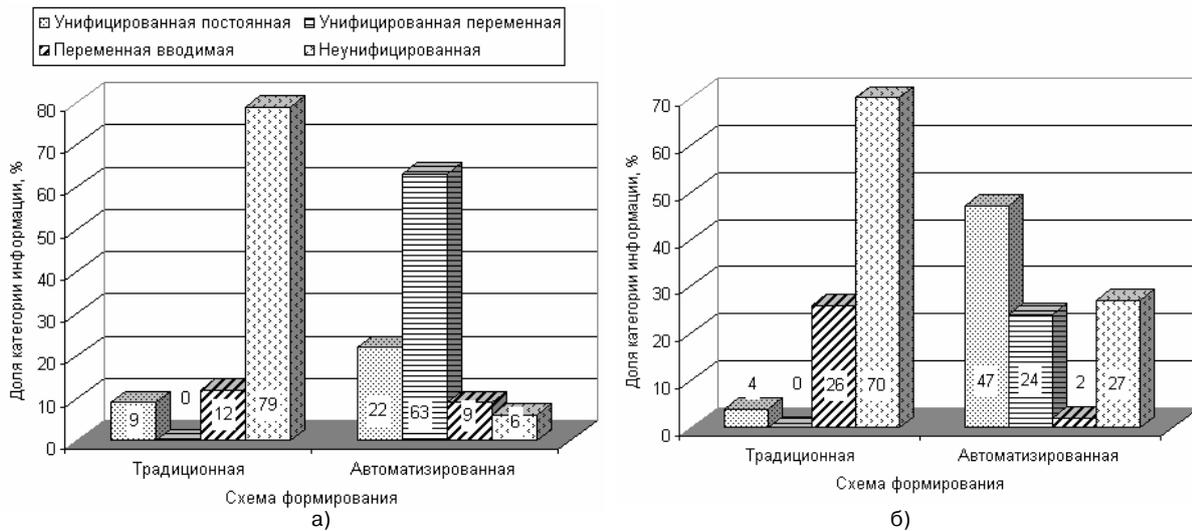


Рис. 6. Структура информации в документах

а) акт готовности кафедры к новому учебному году
 б) протокол эндоскопического осмотра пациента

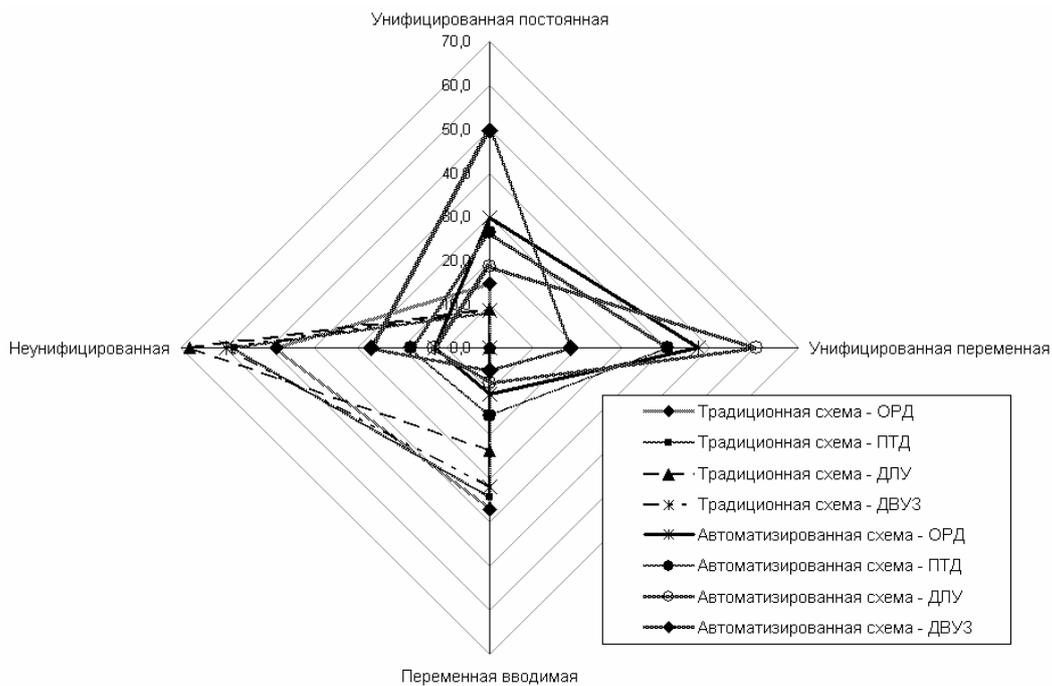


Рис. 7. Диаграмма результатов реструктуризации информационных потоков по подсистемам документов

объема унифицированной информации предопределяет возможность перевода информации, вводимой с клавиатуры, в разряд выбираемой на основе опорных слов, что выполняется исполнителями намного быстрее, чем набор фрагментов документа с клавиатуры.

При неизменности схем прохождения документов доминирующим показателем эффектив-

ности процесса автоматизации подготовки документации является объем трудозатрат, необходимых для формирования того или иного документа. Показатели трудозатрат при формировании документов приведены в Табл. 2.

Следствием снижения трудозатрат, безусловно, будет высвобождение дополнительного времени персонала. Учитывая повторяемость

Табл. 2. Динамика трудоемкости формирования документов

Вид документа	Время формирования, ч		Сокращение трудоемкости, %
	Традиционная схема	Автоматизированная схема	
Документация лечебного учреждения			
Выписной эпикриз	4,6	1,3	71,7
Документация консилиумных совещаний	1,8	0,7	61,1
Протокол первичного осмотра пациента в лечебном отделении	3,7	0,8	78,4
Протокол эндоскопического обследования пациента	0,6	0,2	66,7
Свидетельство о болезни	5,9	2,3	61,0
В среднем по группе документов			67,8
Документация высшего учебного заведения			
Акт готовности кафедры к новому учебному году	3,2	1,8	43,8
Отчет о практиках	4,2	2,5	40,5
Отчет о работе Государственной экзаменационной комиссии	3,6	1,9	47,2
Отчет о работе кафедры в учебном году	30,0	10,0	66,7
Протокол заседания Государственной экзаменационной комиссии	1,8	0,7	61,1

формируемых документов, снижение уровня трудозатрат будет пропорционально возрастать по сравнению с показателями, соответствующими созданию единичного экземпляра, что при накоплении результатов может быть оценено в единицах условного сокращения персонала. При этом высвободившееся время работники смогут направить на выполнение других задач, не связанных с процессом создания документов.

Заключение

Эффективность использования лексикологического синтеза обусловлена существенным снижением трудозатрат персонала при автоматизированном формировании слабоформализуемых документов. Так, например, время, необходимое для составления типового трудового контракта, сокращается с 2,4 часа при прямом вводе информации с клавиатуры до 1,1 часа при использовании компьютерной программы, основанной на предложенной методике. Документы кафедрального уровня подготавливаются в 2-3 раза быстрее. Выигрыш во времени при подготовке организационно-распорядительных документов и документов медицинской направленности может достигать 5-7 раз. Значительный выигрыш во времени сопровождается, помимо всего, повышением качества документа, которое проявляется в большей детализации текста при внедрении описательных фрагментов и невозможности "пропуска" отдельных

фрагментов документа. Кроме того, обеспечивается последовательность и логика изложения текста, а от исполнителей документов не требуется знание требований по составлению документов, поскольку документ создается на основе установленного формуляра-образца, сохраняемого в базе знаний программного комплекса.

Литература

1. Бессонов С.В. Оптимизация электронного документооборота в корпоративных системах: Дис. ... канд. экон. наук. – М.: Фин. академия при правительстве РФ, 2001. – 175 с.
2. Бойцов Л.М. Синтез системы автоматической коррекции, индексации и поиска текстовой информации: Дис. ... канд. техн. наук. – М.: Моск. акад. рынка труда и информационных технологий, 2003. – 138 с.
3. Документы и делопроизводство: Справ. пособие / Сост. М. Т. Лихачев. – М.: Экономика, 1991. – 271 с.
4. Крошилин А.В. Разработка и анализ интеллектуальных программ информационного поиска в вычислительных сетях на основе универсальных алгебр: Дис. ... канд. техн. наук. – Рязань: ГРТА, 2003. – 140 с.
5. Першиков В. И., Савинков В. М.. Толковый словарь по информатике. – М.: Финансы и статистика, 1991.–543с.
6. Силантьева Е.Ю. Исследование и разработка методов и алгоритмов автоматизации документопотоков в корпоративных системах: Дис. ... канд. техн. наук. – М.: МГТУ им. Н.Э.Баумана, 2003. – 134 с.
7. Спецификация MoReq2. Типовые требования к управлению электронными официальными документами. – М.: РОО "Гильдия Управляющих Документацией", 2008. – 287 с.
8. Черников Б.В. Принцип лексикологического синтеза в технологии создания текстовых документов // Секретарское дело, 2000. – № 1. – с. 47-49.

9. Черников Б.В. Способ автоматизированного лексикологического синтеза документов // Патент России № 2253893, 2005.
10. Черников Б.В. Технологии подготовки документов на основе кибернетических методов. – М.: Финансы и статистика, 2009. – 208 с.
11. Baeza-Yates R., Navarro G. Block Addressing Indices for Approximate Text Retrieval // Journal of the American Society on Information Systems, 2000, #51. – P. 69-82.
12. Merseguer J., Campos J., Mena E. A Performance Engineering Case Study: Software Retrieval System // Performance Engineering. State of the Art and Current Trends (Heidelberg). – 2000. – Vol. 2047. – P. 37-332.

Черников Борис Васильевич. Начальник отдела общесистемных решений ООО "АНТ-Информ", Москва. Окончил Рижское ВВАИУ им. Я.Алксниса в 1973 году. Кандидат технических наук, старший научный сотрудник, доцент. Автор более 150 научных работ, в том числе 5 изобретений и 20 монографий. Область научных интересов: проблематика автоматизированного формирования документов. E-mail: bor-cher@yandex.ru.