

# МЕТА-ПРЕОБРАЗОВАНИЯ СИСТЕМ ДОКУМЕНТОВ С ГИПЕРТЕКСТОВОЙ РАЗМЕТКОЙ

М. Ю. Колодин

Санкт-Петербургский институт информатики и автоматизации РАН  
СПИИРАН, 14-я линия ВО, д. 39, Санкт-Петербург, 199178

<myke@iias.spb.su>

---

УДК 681.3

Колодин М. Ю. **Мета-преобразования систем документов с гипертекстовой разметкой.** // Труды СПИИРАН. Вып. 3, т. 1. — СПб.: Наука, 2006.

**Аннотация.** Рассматриваются описание и преобразования документов как многоуровневых расширяемых гипертекстовых систем. Такой подход позволяет повысить уровень представления и эффективность обработки данных и ведения проектов прежде всего в научных исследованиях и разработке программного обеспечения. — Библ. 7 назв.

UDC 681.3

Kolodin M. Y. **Meta-transformations of systems of documents with hypertext markup.** // SPIIRAS Proceedings. Issue 3, vol. 1. — SPb.: Nauka, 2006.

**Abstract.** Description and transformations of documents as multilevel extensible hypertext systems is discussed. This approach allows for increase level of presentation and efficiency of data processing and projects maintenance, first of all, in scientific research and software design. — Bibl.7 items.

---

Особенности современных научных исследований в информатике и разработке программного обеспечения, прежде всего, высокая сложность и динамичность процессов, возрастающие требования к их технологичности и надежности получаемого результата, приводят к необходимости систематизации, теоретической разработки основ и практической реализации методов обработки документов [1, 2, 4, 6].

Одним из перспективных подходов в этом направлении могут стать «мета-преобразования», определяемые как управляемые многоэтапные преобразования многоуровневых систем гипертекстовых документов («мета-систем»), включающие не только исходные их наборы, но и преобразователи, метрики, схемы управления. Под «гипертекстовыми документами» понимается широкий класс информационных объектов (тексты, документация, веб-сайты, программы), которые допускают запись в виде размеченных иерархически и перекрестно структурированных текстов [5].

Основа изучаемого подхода к преобразованиям систем — их многоуровневое рассмотрение и построение [5, 7]. Любой их элемент можно представить как часть более общей структуры, в которой он участвует как активный или пассивный объект, определить его свойства, сформулировать правила преобразований. Наибольший эффект достигается, когда удается использовать единообразные описатели и преобразователи для разных систем и их уровней. Таким образом, получаются многоуровневые структуры, сложность частей которых в сочетании со сложностью связей между ними ниже сложности исходной совокупности документов.

Для описания преобразования документов теоретически достаточным является рассмотрение их совокупности в сочетании с процедурами обработки как состоящей из активной («программы») и пассивной («данные») частей. Практически полезным оказалось выделение понятий преобразователя

(субъекта) и контекста преобразования (как набора внешних и внутренних условий и значений, при которых выполняется преобразование) в дополнение к исходному и результирующему объектам. В процессе преобразований объекты изменяются; таким образом, мы имеем дело с целенаправленным управлением преобразованиями динамических произвольно-уровневых систем гипертекстовых документов. Это хорошо соотносится с принятыми в системном анализе схемами («вход – выход – инструмент – ограничения»), а также с современными исследованиями по семантическим сетям, добавляя единообразие и гибкость в описания и преобразования систем [1, 5, 7].

Разработанная специальная нотация [5, 7] помогает как более точно и полно описать уже используемые системы документов и их преобразователей, так и более целенаправленным, регулярным, конструктивным образом подойти к созданию новых, имеющих заданные характеристики.

Первоначально большая часть работ проводилась на языке Forth [1, 3, 4], как наиболее подходящем для непосредственной реализации мета-свойств и многоуровневости, затем — на языках Perl, Parser, Tcl/Tk, Python, Lisp, т. е. на интерпретируемых языках, позволяющих построить программу и выполнить ее средствами самого языка [6, 7].

Мета-описания на XML и подобных ему языках — основа форматирования и преобразований в рассматриваемом подходе. Часто применялась полуформальная схема: от частично форматированного исходного документа (формата) переходили к четкой разметке на XML, после чего из него получали все нужные результирующие документы во всех требуемых форматах. Успешно применялась и полностью формальная схема работы, с предварительной проработкой всех форматов и преобразователей.

Практические приложения разрабатываемой технологии уже дали положительные результаты: реализованы простой, но гибкий документооборот в научной организации, управление проектами и документами в небольшом коллективе программистов, подготовка сборников научных трудов, разработка информационных (не художественных) динамических веб-сайтов, многоэтапные многоформатные многоуровневые преобразования систем документов с четкой формальной или полуформальной разметкой, описания и преобразования распределенной гипертекстовой информации (как локально, так и в интернете) [2, 3, 5, 6, 7].

Полученные результаты позволяют усовершенствовать работу с документами различного формата и назначения, прежде всего — в научных исследованиях и при разработке программного обеспечения. В дальнейшем предполагается расширить область применения подхода на технологические и организационные задачи, усилить его оптимизационные свойства.

## Литература

1. *Ершов А. П.* Избранные труды. Новосибирск: Наука, Сиб. издательская фирма, 1994. 416 с.
2. *Баранов С. Н., Колодин М. Ю.* Феномен форта // Системная информатика. Новосибирск: Наука, Сиб. издательская фирма, 1995. С. 193–271.
3. *Kolodin M. Y.* Perspective metaforthness // International conference “EuroForth-99”, St.Petersburg, 1999. P. 6–9.
4. *Баранов С. Н., Ноздрунов Н. Р.* Язык форт и его реализации. Л.: Машиностроение, 1988. 157 с.
5. *Колодин М. Ю.* Преобразования многоуровневых систем активных многоформатных документов в расширяемых гипертекстовых форматах // Труды VI Международной

электронной научной конференции «Современные проблемы информатизации в технике и технологиях-2001». Воронеж: ЦЧКИ, 2001. С. 45–48.

6. *Колодин М. Ю.* Разработка, реализация и сопровождение веб-сайта научной организации. // Труды СПИИРАН. 2002. Вып. 1, т. 1. СПб.: СПИИРАН, 2002. С. 217–223.
7. *Колодин М. Ю.* Произвольно-уровневые гипертекстовые системы. // «Современные проблемы информатизации в системах моделирования, программирования и телекоммуникациях». Сб. трудов. Вып. 9 (по итогам IX международной открытой научной конференции). Воронеж: Научная книга, 2004. С. 358–359.