

МОДЕЛЬ ГИПЕРТЕКСТОВОЙ ОРГАНИЗАЦИИ ИСТОРИКО-ИССЛЕДОВАТЕЛЬСКОЙ ИНФОРМАЦИИ

Предлагаемая модель основывается на гипотезе о деятельности субъекта в плане оперирования информацией в процессе постановки и решения исследовательской задачи, которая сформировалась под влиянием ряда работ, собственных наблюдений и бесед с коллегами. Суть гипотезы заключается в том, что с информационной точки зрения процесс познания состоит в накоплении информации фрагментами и объединении и связывании образов (моделей) этих информационных фрагментов.

Фрагмент информации либо вычленяется исследователем из имеющегося в его распоряжении массива информации, либо порождается им самим. Рассматривая фрагмент по частям и воспринимая его содержание с различных точек зрения, с одним и тем же информационным фрагментом, исследователь может связывать множество образов, по разному отражающих этот фрагмент.

Для борьбы с “проклятием размерности”, которое выражается в огромном количестве информационных фрагментов и их образов, накапливающихся в процессе решения исследовательской задачи, применяются такие приемы как декомпозиция и структуризация.

Первый прием заключается в объединении образов в некоторые совокупности, исходя из самых произвольных соображений познающего субъекта. Второй прием состоит в связывании образов друг с другом также по воле субъекта.

Наша цель — предложить модель комфортной и рациональной организации информации, как вовлекаемой в исследование извне, так и возникающей в процессе его. Исходя из рассмотренного представления об информационной деятельности познающего субъекта, по нашему мнению, наиболее подходящей является гипертекстовая (гипермедиа) организация исследовательской информации вообще и историко-исследовательской¹ в особенности.

¹ *Перевертень В.А.* Понятие информации в историческом исследовании // “Информационный бюллетень Ассоциации “История и компьютер”. М., N 14, 1995, с. 17 — 20.

Предлагаемая модель гипертекстовой (гипермедиа) организации исследовательской информации несколько отличается от “классического” представления о гипертексте. В ее основе лежат следующие понятия: информационный объект (ИО), образ ИО, объект гипертекста (ОГТ), ассоциация и связь.

Информационный объект — это фрагмент информации, который может объективно существовать или порождаться как единица информации, произвольно выделяться из имеющегося информационного массива, формироваться познающим субъектом. На содержание, объем, форму представления (текст, аудио, видео или комплексная), структурную организацию (линейный текст, гипертекст, формуляр, книга, база данных, база знаний и т.д.), носитель (бумага, компьютерная среда, мозг человека и др.) информационного фрагмента, понимаемого как информационный объект, не накладывается никаких ограничений.

Образ ИО — это обозначение одного или нескольких аспектов содержания ИО, это некоторые признаки, отражающие содержание ИО, это вторичная информация, являющаяся моделью содержания ИО. Понятие образа ИО охватывает и понятие знака, и понятие поискового образа документа, и понятие вторичного документа.

Образ ИО, дополненный уникальным идентификатором его и указателем на ИО (уникальным идентификатором ИО), к которому этот образ относится, называется *объектом гипертекста*. ОГТ выступают в качестве узлов гипертекстовой сети и играют роль “представителей” ИО в гипертексте. С одним ИО может быть связано несколько ОГТ (один — первичный, а остальные — вторичные), представляющих его с разных сторон. При такой организации сам ИО, будучи включенным в гипертекст, является пассивным компонентом гипертекста и существует независимо от него.

Ассоциациями мы называем группы ОГТ, которые исследователь формирует, исходя из объективных свойств образов ИО или своих субъективных соображений. Каждая ассоциация имеет уникальное имя. Один и тот же ОГТ может входить в несколько ассоциаций.

Под *связью* в предлагаемой модели гипертекстовой организации информации подразумевается симметричное бинарное отношение, устанавливаемое между двумя ОГТ относительно определенных ассоциаций, в которых связываемые ОГТ находятся. Именно в таком определении связей заключается существенное отличие нашей модели от модели “классического” гипертекста.

В формализованном виде модель гипертекстовой организации информации можно определить как пару

$$HT=(HTS, HTP),$$

где HTS — множество допустимых гипертекстовых структур;

HTP — множество допустимых операций над этими структурами.

Пусть NH — множество идентификаторов ОГТ, IM — множество образов ИО, ID — множество идентификаторов ИО, IO — множество ИО, NA — множество имен ассоциаций. Допустимые гипертекстовые структуры (HTS-структуры) определяются соответствующими отношениями, заданными на перечисленных множествах, и ограничениями целостности. В предлагаемой нами модели HTS-структура — это кортеж:

$$HTS=(HTO, IDIO, ASS, LNK, HTSL),$$

где HTO — множество ОГТ; IDIO — отношение идентификатор ИО — ИО; ASS — отношение ассоциации; LNK — множество связей; HTSL — ограничения целостности гипертекстовой структуры. $HTO \subseteq NH \times IM \times ID$, $IDIO \subseteq ID \times IO$, $ASS \subseteq NH \times NA$, $LNK \subseteq NH \times NA \times NH \times NA$, при этом $\pi_1 HTO \leftrightarrow \pi_{2,3} HTO$, $\pi_1 IDIO \leftrightarrow \pi_2 IDIO$, элементы отношения LNK представляют собой кортежи вида (h_1, a_1, h_2, a_2) , где $h_1, h_2 \in \pi_1 HTO$, $a_1 \in ASS(h_1)$, $a_2 \in ASS(h_2)$, причем $h_1 \neq h_2$ и для каждого кортежа в LNK имеется симметричный ему кортеж (h_2, a_2, h_1, a_1) .

Начальное состояние HTS-структуры соответствует утверждению: $HTO = \emptyset \ \& \ IDIO = \emptyset \ \& \ ASS = \emptyset \ \& \ LNK = \emptyset$. В состояниях, отличных от начального, она должна удовлетворять ограничениям целостности HTSL, которые заключаются в следующем утверждении:

$$\pi_1 IDIO = \pi_3 HTO \ \& \ \pi_1 ASS = \pi_1 HTO \ \& \ \pi_{1,2} LNK \subseteq ASS.$$

В множество операций HTP, допустимых над определенными выше гипертекстовыми структурами, входит 15 операций трех типов: операции формирования гипертекста, которыми определяются возможные изменения состояния HTS-структуры; операции выборки элементов гипертекстовой структуры, предназначенные для

отбора тех ее элементов, с участием которых будут выполняться некоторые операции или которые будут просто просматриваться либо редактироваться; операции выделения части гипертекстовой структуры в соответствии с заданными условиями.