

И.В. Ефименко,
Е.Н. Пенская

Национальный
исследовательский
университет «Высшая
школа экономики»

ОНТОЛОГИЧЕСКОЕ МОДЕЛИРОВАНИЕ В ЭКСПЕРТНОЙ ЖУРНАЛИСТИКЕ: РАЗРАБОТКА ИНСТРУМЕНТАРИЯ И МЕТОДОЛОГИИ ФОРМАЛИЗАЦИИ ЭКСПЕРТНЫХ ЗНАНИЙ

Введение

Экспертная журналистика — сравнительно новое для русской общественной жизни и культуры явление, восстановление традиций которого после многолетнего перерыва пришлось на вторую половину 1990-х годов. При этом вряд ли имеется основание считать ее целостным и завершенным феноменом.

В общем и целом экспертную журналистику можно определить как вид журналистской деятельности, построенной на апелляции к экспертным мнениям, их исследовании, сопоставлении и анализе. При этом журналист может как сам являться заказчиком экспертизы, так и использовать экспертные мнения, сведения о которых стали доступны ему каким-либо иным способом и которые были изначально высказаны в контексте, не связанном с создаваемым материалом СМИ.

Особенностью публичной экспертной деятельности является возможность доступа к экспертным мнениям и, как следствие, интерпретация их неэкспертами, что в ряде случаев приводит к их искажению, сознательному или непреднамеренному. В указанном отношении экспертная деятельность, не связанная с публичной активностью, противопоставлена экспертной деятельности, имеющей место онлайн. Таким образом, целесообразно ввести понятия объективной экспертной дискуссии и экспертной дискуссии в СМИ.

Для исследования обоих видов экспертного дискурса, их взаимосвязей, адекватности отражения в СМИ реальной экспертной деятельности, идентификации как действительно экспертного знания, так и признаков манипулирования общественным сознанием, необходимы средства знаний инженерии. Цель настоящей работы состоит в создании соответствующего инструментария, а также в формировании гибридного подхода, позволяющего формализовать работу с экспертным знанием и сюжетными линиями.

1. Экспертная журналистика и формализация знаний: теоретические и методологические основания разрабатываемого подхода. Информационная база исследования

В качестве основы для разработки гибридного подхода к формализации экспертных знаний и исследованию экспертных дискуссий используются методологии и модели, основными из которых являются:

- методология когнитивного моделирования;
- теория дискурса и модели дискурсивного анализа;
- методологии онтологического инжиниринга;
- темпоральные логики.

Методология когнитивного моделирования, предназначенная для анализа и принятия решений в плохо определенных ситуациях, основана на исследовании субъективных представлений экспертов и в настоящее время развивается в направлении совершенствования аппарата анализа ситуации [Кулинич, 2003].

Основными элементами методологии когнитивного моделирования и ее расширений являются представление совокупности сюжетов, возникающих в процессе экспертной дискуссии (см. разд. 3), в виде графа, дуги которого специфицируют отношение влияния факторов в узлах графа друг на друга, и структурно-функциональная декомпозиция, позволяющая выделить составные части наблюдаемой ситуации в виде иерархии «часть-целое». При этом в дополнение к причинно-следственным и иерархическим связям между ситуациями (сюжетами) устанавливаются семантические связи, задаваемые в рамках предметных онтологий, а совокупности сюжетов, связанных различными видами отношений, отображаются на шкале времени. Точки на шкале времени определяются на основе атрибутов текста, фиксирующего процессы развертывания экспертной дискуссии.

Ключевыми элементами *теории дискурса*, используемыми в качестве основы разрабатываемого подхода, являются глобальные и локальные темы, прагматическая пресуппозиция, риторическая схема дискурса (при этом в качестве единого макродискурса рассматривается коллекция документов, посвященных одной экспертной дискуссии) и др. [Кибрик, 2003; van Dijk, 1997].

Основными методологиями *онтологического инжиниринга*, используемыми в работе, являются методология Грюнингера и Фокса [Gruninger, Fox, 1995] и *METHONTOLOGY* [Fernandez et al., 1997; Гурьянова и др., 2011]. В силу ограничений объема статьи ниже рассматриваются прежде всего вопросы, связанные с концептуализацией проблемных областей.

Темпоральные логики используются для анализа процессов развертывания экспертных дискуссий во времени.

Информационной базой исследования послужили материалы более 2000 открытых источников. Для проектирования моделей предметной области «Экспертиза» использовались коллекции документов, предоставляемые такими системами и ресурсами, как «Медialogия», «Интегрум», *Public.ru*, Национальный корпус русского языка (НКРЯ), текстовые базы ФЭП (1992–2012). Основой для разработки моделей сюжетов послужили статьи (новости, аналитика и др.) по тематике «Вопросы здравоохранения в “Стратегии 2020”», полученные с помощью набора экспертных запросов к среде Интернет.

2. Система онтологий в экспертной журналистике. Метамодель предметной области «Экспертиза»

Для построения формальной модели знаний в рамках настоящей работы создается система онтологий и моделей других классов, каждая из которых предназначена для спецификации определенных аспектов проблемной области:

- система моделей экспертной деятельности;
- онтологии видов деятельности, характеризующихся широким использованием экспертных знаний (внешние онтологии);
- онтология типов событий, являющихся основой возникновения экспертных ситуаций. При этом в рассмотрение вводятся две основные составляющие: слой фактов, в рамках которого на уровне экземпляров специфицируются события реального мира, и слой гипотез, порождаемых авторами, которые исследуют дискурс экспертов, или аналитиками, которые исследу-

ют как дискурс экспертов, так и дискурс авторов материалов, апеллирующих к экспертным мнениям. Подобный подход позволяет выявлять противоречия между фактами и гипотезами с использованием аксиоматики онтологий и правил вывода;

- система лингвистических шкал;
- шкала времени;
- совокупность лингвистических онтологий, предназначенных для спецификации механизмов трансляции экспертного знания;
- математические модели, основанные на исследовании функций доверия и правдоподобия (с использованием теории Демпстера — Шафера);
- модели сюжетов.

Далее рассматриваются отдельные составляющие представленной системы моделей.

3. Модели экспертной деятельности и модели сюжета

3.1. Модель экспертной деятельности

Разработка модели экспертной деятельности (в свою очередь, системы взаимосвязанных онтологий с использованием элементов моделей других классов) требует концептуализации нескольких подобластей знания. Перечень моделей, являющихся элементами указанной системы, представлен ниже.

Онтология «Эксперты и экспертиза: базовые понятия экспертной деятельности» специфицирует понятие *эксперт* и его отличия от смежных и взаимосвязанных понятий — *ученый, консультант, лоббист, специалист* и др. [Пенская, 2011; Jasanoff, 2003]; *функции эксперта*; основные *типы задач*, решаемых экспертом; *типы экспертов и экспертных знаний*; составляющие *статуса* эксперта; характеристики эксперта — *компетенции, ангажированность, страновая принадлежность, репутация, школа* и проч.; *институты экспертов*; *экспертные сообщества*; *фабрики мысли* и т.п.

Онтология «Экспертная ситуация» предназначена для описания различных типов коллизий и проблемных зон, требующих обращения к экспертным знаниям, а также событий, являющихся вероятными причинами возникновения экспертных ситуаций, таких как *реформы, разработка стратегий* и т.п. Данная онтология включает спецификацию ряда абстрактных и составных концептов, таких как *контекст и исходные условия* (экспертной ситуации), *идеи, конъюнктура, цели* (различных акторов), *обязательства* (различных акторов), *ресурсы, приоритеты* и т.п.

явления и исчезновения, их взаимосвязей и взаимодействий, а также событий, в которых они участвуют.

Для области экспертной журналистики целесообразно выполнить модификацию указанного подхода. Для решения задач прикладного характера в данной области высокий уровень детализации при описании актантов событий (действующих лиц, акторов) нецелесообразен. В качестве основных акторов выступают:

- эксперт:

внутренний (участник рассматриваемой дискуссии, например, член рабочей группы по разработке документа «Стратегия 2020»),

внешний (эксперт в обсуждаемой проблемной области, не являющийся участником экспертной дискуссии, но высказывающий мнение по вопросам или ведения и содержания самой дискуссии, или области в целом);

- общество;
- государство (власть);
- автор (лицо или группа лиц, являющихся источником материалов, по которым необходима апелляция к экспертным мнениям).

При этом основными элементами сюжетных линий, развивающихся относительно той или иной экспертной дискуссии, являются не актанты, а сами сюжеты, и для каждого из них выполняются следующие действия:

- указываются взаимосвязи рассматриваемого сюжета с другими сюжетами, имеющими место как в тот же момент времени, так и в других точках временной шкалы;
- обеспечивается описание сюжета в терминах основных концептов и атрибутов онтологии сюжетных линий (см. ниже);
- указывается источник появления сюжета (в рамках экспертных дискуссий сюжеты возникают и развиваются не «объективно», а с подачи тех или иных акторов, главным образом, внешних и внутренних экспертов или авторов).

Основными концептами онтологии развития сюжетных линий (элементами модели сюжета) являются представленные ниже.

Макросюжет. Основная тематика экспертной дискуссии, например «Состояние здравоохранения в Российской Федерации».

Сюжет. Элемент макросюжета, тематика, обсуждаемая в контексте основной экспертной темы, составляющая проблемной области, например, «Детская смертность» в рамках макросюжета «Состояние здравоохранения в Российской Федерации». Сюжеты образуют иерархии различной глубины. Семантика иерархической (в нестрогом понимании) связи: «X является под-

типом *У*» или «*Х* является составляющей *У*». В целях упрощения модели целесообразно рассматривать отношения указанных классов как квазисинонимы и ввести для них единое обозначение «*Х* относится к *У*». Пример: сюжет «Изношенность медицинского оборудования» по отношению к сюжету «Плохое состояние лечебных учреждений». Каждый сюжет фиксирует некоторый фактор, влияющий на проблемную область макросюжета непосредственно или через цепочку иерархических связей. Сюжет, расположенный выше по иерархии относительно сюжета *Х*, назовем супертипом сюжета *Х*¹; *а сюжет, расположенный ниже по иерархии, — подтипом.*

Объективная оценка сюжета. Характеристика (атрибут) сюжета, оцениваемого одинаковым образом независимо от источника оценки (с точностью до исключений, которыми можно пренебречь). Обычно синонимом является оценка сюжета обществом, рассматриваемым как единое целое. Возможные значения: «положительно», «отрицательно», «нейтрально», «неоднозначно/неизвестно». Сюжет, оцениваемый положительно, специфицирует фактор, оказывающий положительное влияние на проблемную область макросюжета. Пример положительно оцениваемого сюжета для тематики «Состояние здравоохранения в Российской Федерации» — «Здоровый образ жизни».

Лингвистические шкалы предназначены для качественной оценки текущего состояния развития определенного сюжета и/или его влияния на другой сюжет. Примеры шкал: «слабый-сильный», «маленький-большой», «рано-поздно».

Направление изменений — направление движения точки, специфицирующей состояние развития сюжета или характер его влияния на другой сюжет, по лингвистической шкале. Атрибутом направления изменений является характеристика «Оценка» (возможные значения: «положительно», «отрицательно», «нейтрально», «неоднозначно/неизвестно»). Составляющая «Направление изменения» особенно важна для оценки сюжетов, которые сами по себе являются нейтральными, но приобретают положительный или отрицательный характер при изменении местоположения на определенной лингвистической шкале. Обычно движение по шкале вправо (возрастание) характеризует положительную тенденцию для сюжетов, оцениваемых положительно или нейтрально, и отрицательную тенденцию для сюжетов, оцениваемых отрицательно. Однако в ряде случаев это правило нарушается.

¹ Такое понимание не в точности соответствует определению супертипа в традициях онтологического инжиниринга.

Неразработанная область специфицирует сюжет, который является отрицательной составляющей или «белым пятном» (характеризуемым наличием фактора неизвестности) в рамках в целом положительного сюжета более верхнего уровня, например «Отсутствие источников финансирования» для сюжета «Повышение заработной платы работникам социальной сферы».

Субъективная оценка сюжета — характеристика (атрибут) сюжета, его оценка определенным источником, наиболее значимыми из которых являются «Эксперт» и «Автор».

Другие ключевые составляющие онтологии сюжетных линий: *объективная экспертная дискуссия, экспертная дискуссия в СМИ, акторы, контекст сюжета, сценарий, следствие, влияние* и др.

Пример анализа сюжетов для отдельного документа, содержащего экспертную дискуссию, представлен на рис. 2. Визуализация динамики изменения контекста сюжета «Детская смертность» для совокупности документов (в силу большого числа экземпляров концептов отображается только часть из них) представлена на рис. 3.

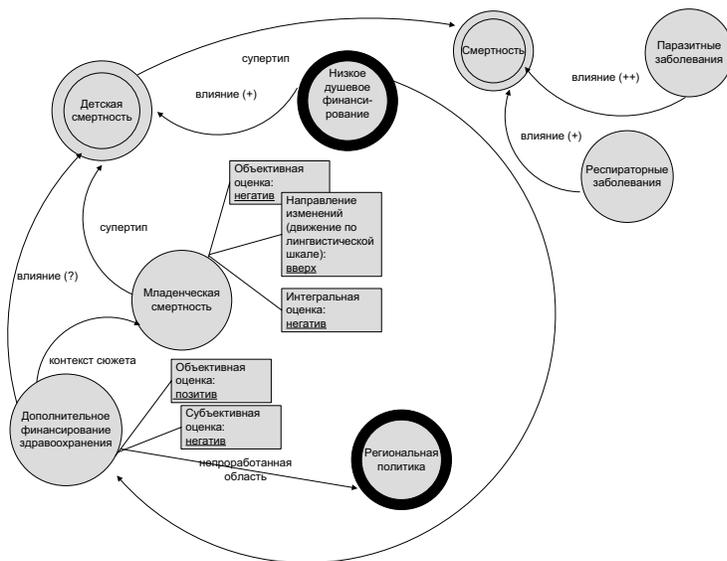


Рис. 2. Визуализация сюжетов для отдельного документа, содержащего фрагмент экспертной дискуссии

3.3. Операции над сюжетами

В процессе исследования были выявлены следующие основные операции над сюжетами, имеющие место при развитии экспертной дискуссии:

- *появление* (зарождение). При этом сюжет может быть введен в рассмотрение непосредственно или через его составляющие;
- *обобщение* — переход к обсуждению супертипа сюжета;
- *детализация* — переход к обсуждению подтипа сюжета;
- *перемещение фокуса* — переход от обсуждения сюжета X к обсуждению сюжета Y , не являющегося его подтипом или супертипом, но относящегося к тому же макросюжету, что и сюжет X ;
- *смена темы* — переход от одного макросюжета к другому через сюжет, непосредственно или опосредованно относящийся к обоим макросюжетам;
- *фиксация результатов* — эксплицитная констатация завершения дискуссии по сюжету по причине получения финальных результатов или осознания невозможности их получения;
- *исчезновение* (вымирание). Относится к случаям, когда сюжет исчезает из дискуссии и при этом не имеет места ни одна из перечисленных операций.

Исследование операций над сюжетами в рамках развития конкретной дискуссии является основой для формирования аналитических выводов.

Пример типового шаблона (сочетания объектов и связей различных типов) для операции «смена темы», специфицирующего возможную манипуляцию и сформированного в результате анализа коллекции документов по тематике «Здравоохранение и «Стратегия 2020», представлен на рис. 4 (экземпляры объектов заполнены на примере одного из документов, содержащего фрагмент экспертной дискуссии).

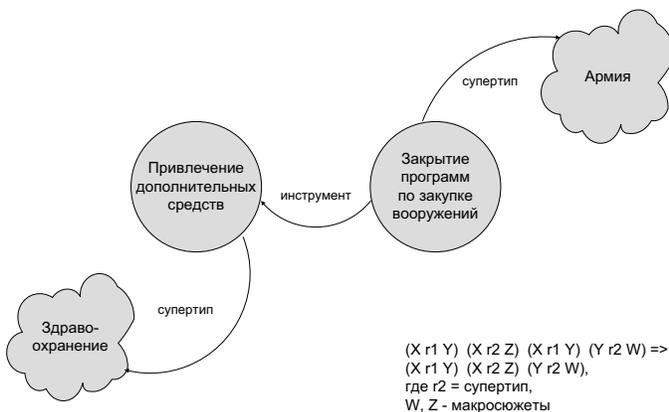


Рис. 4. Шаблон потенциальной манипуляции

Заключение

Результатом работы, представленной в настоящей статье, является создание гибридного подхода и инструментария для формализации экспертных знаний, а также исследования экспертных дискуссий и анализа сюжетных линий.

Предполагается, что результаты исследования будут применимы и получат дальнейшее развитие в области гуманитарных технологий, прогнозирования, экспертной журналистики и анализа дискурса.

На стадии разработки находятся следующие аспекты исследуемой проблемной области:

- возможности вывода на знаниях и типы правил вывода, наиболее релевантные для исследуемой предметной области, в частности возможности создания аппарата для автоматизации генерации гипотез о взаимосвязях разнородных событий, об апелляции к экспертным знаниям с целью манипулирования адресатом и др.;
- применимость предложенного подхода к различным типам сценариев в сфере работы с экспертными знаниями, модели соответствующих сценариев с точки зрения предлагаемой методологии.

Литература

Гурьянова М.А., Ефименко И.В., Хорошевский В.Ф. Онтологическое моделирование экономики предприятий и отраслей современной России. Ч. 2. Препринт. М.: Изд. дом ВШЭ, 2011.

Кибрик А.А. Анализ дискурса в когнитивной перспективе: дис. ... докт. филол. наук. М., 2003.

Кулинич А.А. Методология когнитивного моделирования сложных плохо определенных ситуаций // Вторая Междунар. конф. по проблемам управления. Избр. труды. М., 2003. Т. 2. С. 219–226.

Онтологический редактор Protégé <protege.stanford.edu, 2012>.

Пенская Е.Н. Эксперт, экспертиза, экспертность. Медийный фактор формирования идеологических и языковых концептов // XII Междунар. конф. М.: Изд. дом ВШЭ, 2011.

Fernandez M., Gomez-Perez A., Juristo N. METHONTOLOGY: From Ontological Art Towards Ontological Engineering // Workshop on Ontological Engineering. Spring Symposium Series. AAAI97. Stanford, USA, 1997.

Gruninger M., Fox M.S. Methodology for the Design and Evaluation of Ontologies, Workshop on Basic Ontological Issues in Knowledge Sharing. Montreal, 1995.

Jasanoff S. Breaking the Waves in Science Studies: Comment on H.M. Collins and Robert Evans, "The Third Wave of Science Studies" // Social Studies of Science. 2003. Vol. 33. No. 3. June. P. 389–400.

Van Dijk, Teun A. (ed.). Discourse as Structure and Process. Discourse Studies: A Multidisciplinary Introduction. L.: Thousand Oaks; New Delhi: SAGE Publication, 1997. Vol. 1, 2.