

Семантическая информация и когнитивная информатика

© В.П. Седякин

МГТУ им. Н.Э. Баумана, Москва, 105005, Россия

Рассмотрены методологически важные для информационно-аналитической деятельности вопросы интерпретации информации, существования «несемантической» информации, а также соотношения терминов «смысл», «цель» и «целевая информация».

Ключевые слова: DIKW-модели, данные, информация, анализ текстов, семантика, аналитика

Введение. Прежде всего рассмотрим определения понятий значения и смысла слова. В основополагающей работе советского психолога Л.С. Выготского [1] были введены следующие определения. Значение слов автор определил как «объективно сложившуюся в процессе истории систему связей, которые стоят за словом» [1]. Выготский поясняет это определение на примере слова «чернильница». Это слово сложилось в общественной истории и обозначает нечто, имеющее отношение к краске (черн-), орудийности (-ил-), вместилищам (-ниц-). Таким образом, оно не только указывает на определенный предмет, но и подвергает его анализу, вводит его в систему объективных связей и отношений. Значение есть устойчивая система обобщений, стоящая за словом, одинаковая для всех людей, причем эта система может иметь только разную глубину, разную обобщенность, разную широту охвата обозначаемых им предметов, но она обязательно сохраняет неизменное ядро — определенный набор связей.

С позиции современной методологии науки значение — это конвенциональное соглашение, принятое по поводу одинакового понимания данного слова в данном языке. Толковые словари, по сути, являются сборниками таких соглашений и описывают их, причем значений у одного слова может быть несколько (явление полисемии).

С рассмотренным понятием часто сопоставляется слово «смысл». Под смыслом, в отличие от значения, понимается индивидуальное значение слова, выделенное из этой объективной системы связей; оно состоит из тех связей, которые имеют отношение к данному моменту и данной ситуации. Поэтому если значение слова является объективным отражением системы связей и отношений, то смысл — это привнесение субъективных аспектов значения соответственно данному моменту и ситуации. Общепринято понимание семантики как раздела языкознания, изучающего значение единиц языка.

В силу этого не существует никаких конвенций по поводу смысла слов. Однако при обсуждении несколькими субъектами смысла слов или текстов чаще всего достигается его одинаковое понимание. Важно отметить, что в какой-то момент происходит переосмысление слов. Кроме того, при более широком рассмотрении контекста, как это определил Л. Витгенштейн в концепции «языковых игр» [2], происходит уточнение смысла слов. М. Мазур в работе [3] приводит пример, как с расширением или изменением контекста грамматически правильное, но семантически бессмысленное предложение «Зеленая свобода преследует думающий дом» приобретает смысл.

Не касаясь здесь различных пониманий смысла применительно к отдельному предложению, текстам или деятельности, отметим лишь, что понятие смысла в тексте или деятельности часто отождествляется с главной мыслью, целью или ожидаемым результатом. По отношению к информационно-аналитической деятельности, целью которой является поиск критически важной информации для решения поставленной перед аналитиком задачи, следует отказаться от использования термина «смысл текста», поскольку он обусловлен проблемой контекста и может восприниматься по-разному в зависимости от опыта аналитика и различных условий возникновения текущей ситуации (контекста). В рамках поставленной перед аналитиком задачи следует говорить о поиске им «целевой информации», т. е. критически важной информации, поиск которой ведется аналитиком по доступному ему массиву текстов.

Далее рассмотрим вопросы, связанные с семантической информацией и ее отношениями с социальной и вербально-логической информацией.

Вербально-логическая информация и другие виды информации. В монографии [4, с. 88] рассматривается попытка А.Н. Колмогорова применить аппарат математической теории связи К. Шеннона для анализа русских поэтических текстов. Интуитивно понимая информацию как «снятую неопределенность», Колмогоров и его последователи полагали, что чем неожиданнее использование слова в поэтическом тексте, тем оно информативнее. С точки зрения современной психологии и когнитологии субъект в своем восприятии опирается на неосознаваемые им предположения («паттерны»). Поэтому можно полагать, что Колмогоров и его последователи под «неожиданностью» подразумевали степень «субъективной предположительности». Чем она меньше, тем информативнее следующее слово. Если сведения совпадают с предположениями, то они малоинформативны. Вполне очевидно, что субъективная предположительность восприятия затрудняет использование равновероятно-событийной модели, которая лежит в основе теории Шеннона, для анализа русских поэтических текстов.

Как было отмечено в работе [4, с. 89], понимание первостепенной роли предположений (психологических паттернов) для восприятия информации позволяет определить условность отношений в терминосистеме «данные — информация — знания», рассмотренной в так называемой DIKW-диаграмме [4, с. 74]. Субъект интерпретирует данные на основе своих знаний (сложной системы паттернов) и с учетом характеристик физиологического состояния. Именно это позволяет субъекту решать задачи полисемии и контекста и определять расхождения с предположениями (неожиданности информации).

Рассмотрение терминосистемы «данные — информация — знания» позволяет вернуться к вопросу о соотношениях несемантической и семантической информации. О.В. Зеркаль в статье [5] отмечает, что классики отечественной информатики Р.С. Гиляревский и А.В. Соколов отрицают существование несемантической информации как таковой и полагают ее одним из видов социальной информации. Восприятие субъектом информации — сложный процесс, который отражается в самой терминосистеме «данные — информация — знания». Современное восприятие социальной информации не может замыкаться на одном лишь понятии «информация»: без ее связи с данными и знаниями невозможно представление об информации как о результате процесса восприятия. В настоящее время из-за бурного развития информационных технологий происходит обновление и обогащение информационной терминологии. Именно поэтому из термина «информация» сформировалась терминосистема «данные — информация — знания».

Отмеченная выше необходимость интерпретации данных субъектом на основе его знаний (сложной системы паттернов) позволяет вновь поставить вопрос, что такое семантическая информация. Не отрицая достижений А.А. Харкевича и других авторов, рассматривавших понятия полезности и ценности информации, отметим лишь, что оба они вмещаются в понятие актуальности. Субъект в своем восприятии данных, как уже было отмечено, опирается на имеющиеся у него знания. В его сознании происходит своеобразная «фильтрация». Воспринимается только актуальная информация и только та информация, которую субъект готов принять. Таким образом, любая вербально-логическая информация, воспринимаемая субъектом, должна быть признана потенциально семантической. Оговорка о вербально-логической информации здесь не случайна. В соответствии с концепциями биологической информации, выдвинутой В.И. Корогодиным [5, с. 161], и вербально-логической информации, homo sapiens наследуют от высших животных так называемую поведенческую информацию, которая выражается в мимико-жестовом и тональном языках. В силу вышеизложенного следует согласиться с мнением Гилярева-

ского и Соколова и признать, что любую вербально-логическую информацию можно считать семантической и социальной.

Л. Флориди «определил подход, базирующийся на рассмотрении понятия “информация” с позиции анализа и исключения альтернатив (актуализации одной из возможностей, по В. Дембски) как *модальный подход* в семантической оценке информации» [5]. Популярность работ Флориди не случайна. Выдвинутые им проблемы информации и информатики актуальны для философии и методологии современных информационных наук [6]. Целый ряд рассмотренных Флориди проблем разрешен в работах российских и советских авторов [4]. В книгах Гиляревского сформулированы основные признаки понятия информации при функционально-кибернетическом подходе [7]. В работах Д.М. Дубровского выявлены три основных подхода к определению понятия информации, обсуждавшиеся еще во время знаменитой дискуссии 1970-х годов о сущности информации [8]. В этой дискуссии принимали участие крупнейшие советские ученые А.А. Харкевич, В.И. Сифоров, А.И. Берг, Р.С. Гиляревский, Д.И. Дубровский, Е.К. Войшвилло, В.В. Вержбицкий, Ю.А. Шрейдер и др., был признан общенаучный статус понятия информации и междисциплинарный статус информатики. После дискуссии в СССР и России опубликовано множество работ по информационной тематике, однако далеко не все из них были содержательными в научном отношении. Значительная часть авторов оказались под влиянием ненаучных взглядов А.Г. Спиркина, считавшего информацию категорией, равноправной энергии и материи. В конце 1990-х годов в России сформировалась весьма влиятельная лженаучная школа информациологов. Преодолению этой лженауки во многом способствовали исследования Флориди, которые позволили вернуть на передовые научные позиции публикации отечественных специалистов, изучающих философские и методологические проблемы информационных наук.

Ю.А. Шрейдер опубликовал несколько работ по информационной тематике. Наиболее известны [9, 10]. В книге «Об одной модели семантической информации», не выдвигая никакой самостоятельной теории или даже гипотезы по поводу «семантической информации», он ввел семантическую меру количества информации. Следует отметить затруднительность выдвижения гипотезы и тем более теории семантической информации, поскольку крайне сложно дать определение семантической информации. Признаков у семантической информации, кроме известных для кибернетически-функциональной информации, найти не удастся, за исключением наличия смысла. Но само определение кибернетически-функциональной информации как содержание отношений, возникающих при адаптации субъекта к внешней среде, указывает на актуальность такой информации. Иначе говоря, указывает на ее осмысленность, направленность на полезную для субъекта адаптацию.

В работе [4, с. 79] рассматривается проблема оценки количества информации. Д.И. Менделеев, имея в виду естественные науки, полагал, что наука начинается там, где начинается измерение величин. В настоящее время известны комбинаторный, вероятностный, прагматический, структурный (также его называют объемным или алфавитным) и семантический (по тезаурусной мере) подходы к оценке количества информации. Семантический предложил Шрейдер [9]. Комбинаторный и вероятностный позволяют вычислить количество данных, содержащихся в равновероятных сообщениях. Прагматический подход к измерению количества информации предложил Харкевич в 1960 г. [10]. За меру ценности принимается количество информации, необходимое для достижения поставленной цели, т. е. для максимизации вероятности достижения цели. В отличие от комбинаторного и вероятностного подходов, которым в DIKW-диаграмме отвечают данные, прагматический относится уже к более высокому уровню интерпретации данных, т. е. к уровню информации.

Как было отмечено в [4, с. 82], с точки зрения семиотики комбинаторный и вероятностный подходы отвечают синтаксическому уровню, а прагматический подход Харкевича — прагматическому уровню.

Семантический подход Шрейдера связывает количество информации с тем приращением тезауруса, которое происходит у получателя информации. Шрейдер опирался на идею Н. Винера о том, что для осознанного использования информации получатель должен обладать определенным запасом знаний, выражаемым в тезаурусе. Семантический подход с точки зрения семиотики отвечает уже семантическому уровню, а в соответствии с DIKW-диаграммой — уровню интерпретации информации.

В работе [11], выполненной Шрейдером с соавторами, рассматриваются вопросы изучения информационного взаимодействия с точки зрения субъект-объектных и субъект-субъектных отношений. В 1960-е годы проблематика информационного взаимодействия представлялась неактуальной и малопонятной. По сути, статья опередила свое время. Введенные авторами термины майевтики и фасцинации представлялись очень далекими от практики. Лишь спустя 20 лет оказалось, что оба термина прямо связаны с развиваемой Соколовым теорией информационных когнитивных потребностей и позволяют объяснить парадокс с «пустой SMS-кой» [4, с. 172], когда адресат информации получает больше информации, чем содержится в самом сообщении (майевтика). Фасцинация при параллельном действии индивидуальных когнитивных потребностей вызывает удовлетворение не только коммуникативных ИКП, но и эстетических и этических. Таким образом, спустя многие годы работу [11] можно считать актуальной и практически значимой.

Заключение. Автор настоящей статьи считает особенно важным то, что современная терминосистема «данные — информация — знания», опирающаяся на DIKW-диаграмму, позволяет рассматривать отношения информации с данными и знаниями с привлечением достижений психологии и когнитологии. В этом заключается главное отличие от принятого ранее понимания информации, когда термин «данные» не отделялся от информации. В DIKW-диаграмме данные могут быть синтаксически неверными и не иметь значения для получателя. А информация как результат интерпретации содержит только синтаксически верные сообщения. Поэтому потенциально любая информация — семантическая. Оценка количества семантического содержания информации, предложенная Шрейдером, сводится к увеличению тезауруса у получателя, достигаемого при получении информации. Здесь следует заметить, что приращение тезауруса не сводится лишь к прибавлению нового слова (термина). Часто приращение может ограничиваться освоением получателем нового значения уже известного ему слова (термина). Это утверждение можно проиллюстрировать на примере самого термина «информация». Его новое метафорическое значение является, например в биоинформатике, новым содержанием, которое отличается от содержания исходного понятия и имеет другие признаки.

Семантическая информатика [12, 13] как теоретическое направление сложилась еще до появления современных работ, рассматривающих новую информационную терминосистему, и не получила заметного развития. В то же время проблематика, связанная с новой информационной терминосистемой и DIKW-диаграммой, вызывает интерес и активно развивается применительно к когнитивной информатике, например, в работах студентов кафедры «Информационная аналитика и политические технологии» МГТУ им. Н.Э. Баумана, которые выступали с докладами, посвященными информационным моделям профессионального сознания специалистов по информационной аналитике [14–16]. Наряду с когнитивной информатикой представляется перспективным изучение информационных моделей сознания на основе расширения DIKW-диаграммы для некоторых направлений искусственного интеллекта.

ЛИТЕРАТУРА

- [1] Выготский Л.С. *Мышление и речь*. Москва, Лабиринт, 1996, 414 с.
- [2] Витгенштейн Л. *Философские работы*. Москва, Гнозис, 1994, 235 с.
- [3] Мазур М. *Качественная теория информации*. Москва, Мир, 1974, 238 с.
- [4] Седякин В.П. *Проблемы методологии и философии информатики*. Москва, Книгиздат, 2019, 278 с.

- [5] Зеркаль О.В. Семантическая информация и подходы к ее оценке. *Философия науки*, 2014, № 1, с. 53–66.
- [6] Хлебников Г.В. О философии информации Лучано Флориди. *Теория и практика общественно-научной информации*, 2013, вып. 21, с. 15–59.
- [7] Арский Ю.М., Гиляревский Р.С., Туров И.С., Черный А.И. *Инфосфера: информационные структуры, системы и процессы в науке и обществе*. Москва, ВИНТИ, 1996, с. 394.
- [8] Дубровский Д.И. О специфике философской проблематики и основных категориальных структурах философского знания. *Вопросы философии*, 1984, № 11, с. 27–34.
- [9] Шрейдер Ю.А. Об одной модели семантической информации. *Проблемы кибернетики*, 1965, вып. 13, с. 41–47.
- [10] Харкевич А.А. О ценности информации. *Проблемы кибернетики*, 1960, вып. 4, с. 54.
- [11] Кузнецов Н.А., Мухешвили Н.Л., Шрейдер Ю.А. Информационное взаимодействие как объект научного исследования. *Вопросы философии*, 1963, № 3, с. 108.
- [12] Шемакин Ю.И. *Систематика*. Москва, Изд-во РАГС, 2006, 251 с.
- [13] Шемакин Ю.И., Романов А.И. *Компьютерная семантика*. Москва, Школа Китайгородской, 1995, 342 с.
- [14] Чжен А.О. Когнитивно-информационная модель творчества в области информационной аналитики. В сб. тезисов докладов Всероссийской студенческой конференции «Студенческая научная весна», посвященной 190-летию МГТУ им. Н.Э. Баумана. Москва, Изд-во МГТУ им. Н.Э. Баумана, 2020, с. 262.
- [15] Фоменко Д.Е. Разработка когнитивно-информационной модели профессионального сознания информационного аналитика. В Сб. тезисов докладов Всероссийской студенческой конференции «Студенческая научная весна», посвященной 190-летию МГТУ им. Н.Э. Баумана. Москва, Изд-во МГТУ им. Н.Э. Баумана, 2020, с. 263.
- [16] Юников С.Р. Расширение DIKW-модели. В Сб. тезисов докладов Всероссийской студенческой конференции «Студенческая научная весна», посвященной 190-летию МГТУ им. Н.Э. Баумана. Москва, Изд-во МГТУ им. Н.Э. Баумана, 2020, с. 265.

Статья поступила в редакцию 02.03.2021

Ссылку на эту статью просим оформлять следующим образом:

Седакин В.П. Семантическая информация и когнитивная информатика. *Гуманитарный вестник*, 2021, вып. 1. <http://dx.doi.org/10.18698/2306-8477-2021-1-703>

Седакин Владимир Павлович — д-р филос. наук, доцент, профессор МГТУ им. Н.Э. Баумана. Область научных интересов: философия науки и техники и инженерной деятельности, политология как научная и учебная дисциплина, политическая этика и этика в преподавании. e-mail: sedyakin@bmstu.ru

Semantic information and cognitive informatics

© V.P. Sedyakin

Bauman Moscow State Technical University, Moscow, 105005, Russia

The paper focuses on issues of information interpretation, existence of "non-semantic" information, as well as the relationship between the terms "meaning", "goal" and "target information", which are methodologically important problems for information and analytical activities.

Keywords: DIKW models, data, information, text analyses, semantics, analytics

REFERENCES

- [1] Vygotskiy L.S. *Myshlenie i rech* [Mentality and speech]. Moscow, Labirint Publ., 1996, 414 p.
- [2] Wittgenstein L.J.J. *Philosophische Untersuchungen*, Suhrkamp Verlag AG; Neuauflage, 2010 [In Russ.: Wittgenstein L.J.J. *Filosofskie raboty*. Moscow, Gnozis Publ., 1994, 235 p.].
- [3] Mazur M. *Jakościowa teoria informacji*. WNT, Warszawa 1970. [In Russ.: Mazur M. *Kachestvennaya teoriya informatsii*. Moscow, Mir Publ., 1974, 238 p.].
- [4] Sedyakin V.P. *Problemy metodologii i filosofii informatiki* [Problems of methodology and philosophy of informatics]. Moscow, Knigizdat Publ., 2019, 278 p.
- [5] Zerkal O.V. *Filosofiya nauki — Philosophy of Sciences*, 2014, no. 1, pp. 53–66.
- [6] Khlebnikov G.V. *Teoriya i praktika obshchestvenno-nauchnoy informatsii (Theory and practice of social science information)*, 2013, no. 21, pp. 15–59.
- [7] Arskiy Yu.M., Gilyarevskiy R.S., Turov I.S., Cherny A.I. *Infosfera: informatsionnye struktury, sistemy i protsessy v nauke i obshchestve* [Infosphere: information structures, systems and processes in science and society]. Moscow, VINITI Publ., 1996, 394 p.
- [8] Dubrovskiy D.I. *Voprosy filosofii — Russian Studies in Philosophy*, 1984, no. 11, pp. 27–34.
- [9] Shreider Yu.A. *Problemy kibernetiki (Problems of cybernetics)*, 1965, no. 13, pp. 41–47.
- [10] Kharkevich A.A. *Problemy kibernetiki (Problems of cybernetics)*, 1960, no. 4, p. 54.
- [11] Kuznetsov N.A., Muskheshvili N.L., Shreider Yu.A. *Voprosy filosofii — Russian Studies in Philosophy*, 1963, no. 3, p. 108.
- [12] Shemakin Yu.I. *Sistemantika* [Systemantics]. Moscow, RANEPА Publ., 2006, 251 p.
- [13] Shemakin Yu.I., Romanov A.I. *Kompyuternaya semantika* [Computer semantics]. Moscow, Shkola Kitaygorodskoy Publ., 1995, 342 p.
- [14] Chzhen A.O. *Kognitivno-informatsionnaya model tvorchestva v oblasti informatsionnoy analitiki* [Cognitive information model of creativity in the field of information analytics]. In: *Sb. tezisov dokladov Vserossiyskoy studencheskoy konferentsii «Studencheskaya nauchnaya vesna», posvyaschennoy 190-letiyu MGTU im. N.E. Baumana* [Collection of abstracts of the All-Russian student conference “Student scientific spring” dedicated to the 190th anniversary of Bauman Moscow State Technical University]. Moscow, BMSU Publ., 2020, p. 262.

- [15] Fomenko D.E. Razrabotka kognitivno-informatsionnoy modeli professionalnogo soznaniya informatsionnogo analitika [Development of a cognitive-information model of the professional consciousness of an information analyst]. In: *Sb. tezisov dokladov Vserossiyskoy studencheskoy konferentsii «Studencheskaya nauchnaya vesna», posvyaschennoy 190-letiyu MGTU im. N.E. Baumana* [Collection of abstracts of the All-Russian student conference “Student scientific spring” dedicated to the 190th anniversary of Bauman Moscow State Technical University]. Moscow, BMSTU Publ., 2020, p. 263.
- [16] Yunikov S.R. *Rasshirenie DIKW modeli* [Extending the DIKW model]. In: *Sb. tezisov dokladov Vserossiyskoy studencheskoy konferentsii «Studencheskaya nauchnaya vesna», posvyaschennoy 190-letiyu MGTU im. N.E. Baumana* [Collection of abstracts of the All-Russian student conference “Student scientific spring” dedicated to the 190th anniversary of Bauman Moscow State Technical University]. Moscow, BMSTU Publ., 2020, p. 265.

Sedyakin V.P., Dr. Sc. (Philosoph.), Assoc. Professor, Professor, Department of Information Analytics and Political Technologies, Bauman Moscow State Technical University. Research interests: philosophy of science, technology and engineering, political science as a scientific and academic discipline, political ethics and ethics in teaching.
e-mail: svp134@mail.ru, sedyakin@bmstu.ru