

ОСОБЕННОСТИ РАЗВИТИЯ НАПРАВЛЕНИЙ ИНФОРМАТИКИ

В статье раскрывается особенность двух направлений развития информатики и отношение к ним теории К.Шеннона. Показано различие между информатикой и геоинформатикой. Дается ретроспективный анализ подготовки работы К.Э.Шеннона по математической теории связи.

Ключевые слова: информатика, обработка данных, семантика информации, теория информации.

*S. A. Kudzh
V. Ia. Tsvetkov*

FEATURES OF THE DEVELOPMENT OF COMPUTER SCIENCE

The article describes the development of the two areas of computer science. The article analyzes the relationship between the theory of K.Shannon and computer science. Article shows the difference between computer science and geoinformatics. The article reveals the features of the preparation work of K.Shannon on the mathematical theory of communication

Key words: computer science, data processing, data semantics, information theory.

Развитие современных информационных технологий и наук об информации влечет не только к обновлению терминологической основы, но и к анализу научной классификации этих дисциплин. В работе Ю. Ю. Черного [1] обстоятельно представлена история вопроса развития информатики в СССР и современной России. Кратко остановимся на следующих моментах. Проблема двойного трактования понятия информатика нашла отражение в развитии информатики. Двойственность развития обусловлена историческими причинами.

К концу 1940-х гг. в связи с частым употреблением словосочетания «научная информация», ученые ВИНИТИ предложили «теорию научной информации». Эта позиция нашла отражение выражение в работах А. И. Михайлова, В. А. Полушкина, А. И. Черного, Р. С. Гиляревского и других (1962-1965 гг) [1].

Теория информации нужна была философам для познания сущности мира, лингвистам и филологам для понимания смысла, историкам для анализа развития общества. Смысл и знание – вот в чем нуждались гу-

манитарии. Впоследствии эта область была обозначена как «Информатика 1» [1]. В англоязычной научной литературе она имеет аналог – information sciences.

Работа К.Э. Шеннона [2] при ее написании никакого отношения не имела к теории информации, ее целевым назначением было создание общей теории связи [3].

Термин «информатика» появился в 1960-х гг. во Франции для названия области, занимающейся автоматизированной переработкой информации как слияние французских слов information и automatique (F. Dreyfus, 1962). В разных странах информатика (ср. нем. Informatik, англ. Information technology, фр. Informatique, англ. computer science – компьютерная наука – в США, англ. computing science – вычислительная наука – в Великобритании) интерпретировалась как наука о способах получения, накопления, хранения, преобразования, передачи, защиты и использования информации.

При этом термин «информация» не связывался с какой-либо предметной областью, а был обобщением информации как объекта

обработки на компьютере. По существу это была не первичная информация, а ее переработка в виде моделей, пригодных для обработки на компьютере, что тут же создало проблему «семантического разрыва» [4].

В СССР и России довольно долго этот термин использовался как синоним термину «программирование». В России курсы по изучению информатики длительное время (до 2000) включали изучение программирования в первую очередь и применение информационных систем – во вторую. Но никакой теории информации или семантики информации в них не было. Эту информатику и называют «Информатикой 2». В англоязычной научной литературе она имеет аналог – computer sciences.

Объектом изучения (областью исследований) «Информатики 2» или ее доминантой являются методы обработки информации, компьютерные модели, алгоритмы анализа и вычислений – безотносительно к области использования этой информации. Следует подчеркнуть, что «информатика 2» занимается именно обработкой информации.

С этих позиций «информатика 2» – посредник между математикой и логикой, с одной стороны, и прикладными науками, с другой. Она возникла и развивается как наука об обработке информации, безотносительно к ее смысловому содержанию. Напомним определение А.П. Ершова «Информатика – фундаментальная естественная наука, изучающая процессы передачи и обработки информации» [5]. То есть налицо тесные отношения с теорией связи и отсутствие таких отношений с теорией информации.

Применение «информатики 2» в любой области основано на адаптации «информатики 2» к методам, задачам и данным этой предметной области. В результате такой специализации появляется специфическая информатика в данной предметной области.

Например, экономическая информатика (бизнес-информатика); информатика в биологии (биоинформатика); информатика в медицине, информатика в геодезии, информатика в геологии и т. д. В этот перечень мы не включили геоинформатику не случайно. Она имеет другое происхождение.

«Специализированные информатики» возникают на основе дифференциации «информатики 2» применительно к проблемной области.

Таким образом, развитие «информатики 2» происходит путем совершенствования методов обработки с одной стороны и ее дифференциации и адаптации к разным предметным областям. Данные, применяемые в «информатике 2», – это компьютер-

ные модели, отражающие взаимодействия в предметную область исследования. То есть никаких специальных условий на эти данные не налагается.

Следует также разграничивать понятия «вычислительная математика» и «информатика 2». Вычислительная математика занимается теорией обработки данных на компьютере, а информатика использует эту теорию.

Важно отметить, что в реальной практике вектор специализированной информатики существенно меняется. Главным для этих специализированных информатик является не обработка информации как таковая (доминанта «информатики 2»), а решение проблем и задач в своей области с использованием информатики.

В отличие от информатики, которая имеет один источник происхождения, геоинформатика [7] возникла на основе интеграции наук о Земле (гео) и методов автоматизированного проектирования.

Объективная потребность в интеграции наук о Земле: геодезии (номенклатура ВАК 25.00.32), фотограмметрии (25.00.34), картографии (25.00.33), дистанционного зондирования Земли, землепользования – назревала давно, независимо от информатики. Преподавание этих дисциплин отличалось большим взаимопроникновением. На практике геодезисты часто работали в фотограмметрии, фотограмметристы работали на геодезических работах или занимались составлением карт и т. д. Поэтому геоинформатика (25.00.35) относится к классу «Науки о Земле» (класс 25) она явилась ответом на необходимость интеграции этих наук в единую систему. В качестве смежной науки геоинформатики – информатика не значит.

«Информатика 2» (Номенклатура ВАК 05.13.01, 05.13.05, 05.13.06, 05.13.11, 05.13.11, 05.13.13, 05.13.15, 05.13.17, 05.13.18, 05.13.19, 05.13.20) – «размазана» по ряду специальностей и относится принципиально к другому классу наук (технические науки – класс 05) согласно классификации ВАК. В качестве смежной науки для информатики – геоинформатика не значит. Точно также в паспорте специальности 25.00.35 информатика не значит в качестве смежной науки для геоинформатики.

Следует отметить, что наряду с геоинформатикой существовали, например, информатика в геодезии (до недавнего времени такие специалисты готовились в вузах России), информатика в геологии, информатика в фотограмметрии, информатика в картографии. Но в отличие от геоинформатики – это дифференцированные науки, которые

создавались на основе дифференциации информатики, в то время, как геоинформатика (неудачное название) создавалась на основе интеграции.

Следует отметить, что зарубежные названия геоинформатики имеют такие значения: «geo science», «geoinformatica» и «geo computer science». В геоинформатике имеются свои данные – «геоданные». Они имеют специфическую структуру и всегда включают три группы «место», «время», «тема».

Синонимом геоинформатики является геоматика, которая с «информатикой 2» тесно не связана. Геоинформатика имеет более тесную связь с теорией искусственного интеллекта, чем информатика 2. В теории искусственного интеллекта (ИИ) с 70-х годов ведутся работы по моделированию и представлению «пространственного знания» [8] (на апрель 2013 на эту работу имелось 732 ссылки). С появлением геоинформатики работы по получению и изучению пространственных знаний объединились с работами в области ИИ. В настоящее время ведутся работы по изучению «пространственно-временного знания» [9].

Информатика изучением пространственного знания не занималась и не занимается. В ней такого понятия нет и не предвидится.

Развитием геоинформатики является геомаркетинг, геостатистика. Объектами исследования геоинформатики являются пространственные отношения, геоданные, геореференция [10, 11]. Все это лежит за рамками «информатики 2».

В качестве основы теории информации длительное время применялась «математическая теория связи». По этому поводу необходимо отметить следующее. Фундаментальная работа К.Э Шеннона «A Mathematical Theory of Communication» 1948 [2] (математическая теория связи), которую теперь многие отождествляют с теорией информации, не была направлена на создание теории информации.

Перед 1948 годом связь была строго технической дисциплиной, с небольшой научной теорией для ее поддержки. К тому времени использовались следующие системы: телеграф (1830), телефон (1870), беспроводный телеграф (1890), радио (1900), телевидение (1930), телетайп (1930), системы частотной модуляции (1930), импульсно кодовая модуляция (ИКМ) (1930), вокодер (1930) и др. Однако, не имелось единой теории, объединяющей разные подходы в области связи. Именно поэтому Шеннон поставил перед собой цель – создание научной теории связи, которая могла бы стать по возможности универсальной для объединения

разных направлений в теории связи (того времени). Проблемы связи были много важнее, чем проблемы теории информации. Поэтому Шеннон не собирался писать теорию информации, это не входило в его планы.

Математическую теорию связи [2] Шеннона, написанную в 1948 г., в современных условиях нельзя считать адекватной теорией информации. Она рассматривает только передачу дискретных сообщений и измеряет только информационную емкость символов без учета смыслового содержания сообщений и информации.

В зарубежной литературе, в отличие от Советской и Российской [12] для информации, количество которой измеряет мера Шеннона, ввели специальный термин «Self-information». В соответствии с интерпретацией этого термина «Self-information» обозначает такое количество (физической) информации, которое как частичное знание относительно некоторого явления, добавляется к какому-то полному знанию об этом явлении.

В российской литературе [13] есть неполный перевод этого термина как «Собственно информация». Внимательное изучение первоисточника дает основание перевести данный термин как «Собственная информация (сообщения, символа)», синонимом которому является термин «Информационная емкость (сообщения, символа)» [14].

В современных информационных технологиях и операционных системах устойчиво употребляется такой термин как «контейнер». Информационная емкость такого контейнера может быть разной, он может быть пустым и наполненным. Именно емкость такого контейнера и определяет мера К.Э. Шеннона. Это не информация в широком смысле слова, а только какая-то специфическая ее часть.

Следует отметить, что одновременно с работой К. Э. Шеннона вышла и работа Н. Винера [15], в которой он также рассматривает эту проблему. Винер называет теорию К.Э. Шеннона «статистической теорией количества информации»

«В этой теории за единицу количества информации принимается количество информации, передаваемое при одном выборе между двумя равновероятными альтернативами. Такая мысль возникла одновременно у Р. А. Фишера, у доктора Шеннона из Bell Labs и у автора настоящей книги» [15].

«При этом Фишер исходил из статистической теории, а Шеннон из проблемы кодирования информации» [15].

Следует отметить, что основатель математической статистики Р. А. Фишер ввел свое

понятие информации, которое используется до настоящего времени и называется «информация по Фишеру». В математической статистике информацией Фишера называется дисперсия функции вклада выборки.

Основываясь на работах Найквиста и Хартли, К. Э. Шеннон ясно различал передаваемое сообщение и семантику сообщения. Он четко сформулировал положение о том, что: «семантические аспекты информации не релевантны техническим проблемам связи» [2]

По существу Шеннон рассматривал связь и передачу неких информационных объектов, имеющих определенную информационную емкость. Модель информации, которая лежит в основе теории Шеннона, основана на раскрытии неопределенности. Это статистический, а не семантический подход. Результатом развития теории информации по К. Э. Шеннону явились различные вариации на тему энтропии и две замечательные формулы.

Первая формула – оценки коэффициента эмерджентности системы. В этом случае формула Хартли принимает вид

$$I = \text{Log}_2 W^\varphi \quad (1)$$

где I – количество информации; W – число «чистых» состояний системы; φ – коэффициент (степень) эмерджентности системы; φ рассматривают как степень системности и оценки качества системы, в частности, выделяют следующие альтернативы: $\varphi < 1$ деструктивная система; $\varphi = 1$ аддитивная система; $\varphi > 1$ синергетическая система.

Вторая формула (формула Харкевича) – оценки ценности информации, которая выражается через приращение вероятности достижения цели. Если до получения информации вероятность достижения цели была p_0 , а после получения информации – p_1 , то величина ценности информации определяется по формуле Харкевича [16]:

$$I_0 = \log_2(p_1/p_0). \quad (2)$$

Очевидно, что она может быть и отрицательной, если $p_1 < p_0$

Но дальше развития этой формулы не произошло. В результате можно говорить о застое, в котором оказалась «информатика 2». Одним из итогов развития информатики можно считать «информациологию» И.И. Юзвишина [17], написанную на основе субстанционального понимания информации.

Как необходимость выхода из этого состояния явились предложения по созданию «Интегральной информатики» или «Информатики 3» [18] в России и работы по созданию новой «Семантической теории информации» профессора Оксфордского университета Лучиано Флориди [19, 20].

Не удовлетворенный теорией К.Э. Шеннона, Л. Флориди взял за основу подход Н. Виннера и попытался написать теорию (общей) семантической информации. В ИНИОН РАН в 2011г. начал свою работу семинар «Методологические проблемы наук об информации». Начало было положено обсуждением [21] работы Л. Флориди «Открытые проблемы в философии информации», опубликованной в 2004г., в которой сформулированы 18 проблем. Одна из них гласит – «Возможна ли большая объединенная теория информации?»

Сложнее с «информатикой 3». Объектом ее названы свойства информации, а предметом – изучение законов и методов преобразования и распределения информации в природе и обществе для эффективной информатизации. Однако естественная интеграционная основа в этой науке не определена. Она по прежнему развивается изолировано от других наук. Характерным примером является то, что в трудах института проблем информатики (ИПИ) публикуются одни сотрудники этой организации, боясь допустить взгляд со стороны или иную точку зрения.

В работе [18] автор (идеолог интегральной информатики) признает неудачным название термина «информатика» и предлагает другой «компьютинг». Но за рубежом существует достаточное количество журналов с таким названием, на русский они переводятся как «вычисление» или «вычислительные методы».

Выводы. В то время как за рубежом интенсивно развивают теории семантической теории информации на основе подхода Н. Винера и Л.Флориди, в России нет согласия между сторонниками разных информатик. Сторонники «информатики 2» находятся в состоянии поиска. Их подход - главное это обработка и передача сообщений. Одной из их основных концепций является упорное использование энтропии К.Э. Шеннона как универсального средства в теории информации. Сторонники «информатики 1» придерживаются иной точки зрения. Их подход в том, что семантика является основой информации и информационных сообщений, безотносительно к форме представления. Этот подход широко применяется в лингвистике, в теории искусственного интеллекта, в теории информационного поиска и др.

ЛИТЕРАТУРА

1. Черный Ю.Ю. Полисемия в науке: когда она вредна? (на примере информатики) // Открытое образование.- М., 2010, N 6, с.97-106
2. С.Е. Shannon, (1948), "A Mathematical Theory of Communication", Bell System Technical Journal, vol. 27, pp. 379–423 & 623–656, July & October, 1948
3. Cover T. M., Thomas J. A. Elements of Information Theory. – New York: J. Wiley , Sons, 1991
4. V. Y. Tsvetkov. Information Interaction as a Mechanism of Semantic Gap Elimination // European Researcher, 2013, Vol.(45), № 4-1, p.782- 786.
5. Ершов А.П. Информатика: предмет и понятие. // Кибернетика. Становление информатики. М.: Наука, 1986. С-28-31.
6. Буйневич М.В., Кудж С.А. Выбор рационального варианта информационно-технического взаимодействия в распределенных системах // Вестник ИНЖЭКОНа. Серия: Технические науки. 2011. № 8. С. 37-47
7. Максудова Л.Г., Савиных В.П., Цветков В.Я Интеграция наук об окружающем мире в геоинформатике // Исследование Земли из космоса. - 2000.- №1. - С.46-50.
8. Benjamin Kuipers. Modeling Spatial Knowledge (1978) // Cognitive Science - №2. - p. 129-153.
9. Antony Galton. Spatial and temporal knowledge representation // Earth Science Informatics, September, 2009, Volume 2, Issue 3, pp 169-187
10. Hill Linda L. Georeferencing: The Geographic Associations of Information – MIT Press Cambridge, Massachusetts, London, England- 2009, - 272 p
11. Майоров А.А., Цветков В.Я. Геореференция как применение пространственных отношений в геоинформатике // Геодезия и аэрофотосъемка, - 2012.- №3. - С. 87 -89
12. Большая Советская Энциклопедия / 3-е издание в 30 томах.- М.: Сов энциклопедия, 1969 - 1978 гг.
13. Ожиганов А.А., Тарасюк М.В. Передача данных по дискретным каналам/ - СПб.: Санкт-Петербургский государственный институт точной механики и оптики 1999. 102 с.
14. Хелд Г. Технологии передачи данных. 7-изд. - СПб.: Питер, 2003. - 720 с.
15. Winner N. Cybernetics or Control and Communication in the Animal and the Mashine. The Technology Press and John Wiley & Soris Inc. New York — Herman et Cie, Paris, 1948 99 p.
16. Харкевич А.А. Избранные труды в 3-х томах. Том 3. Теория информации. Опознавание образов. М.: Наука, 1973. 524 с
17. Юзвишин И.И. Информациология или закономерности информационных процессов.- Высшая Школа. М. 2001
18. Колин К.К. Становление информатики как фундаментальной науки и комплексной научной проблемы // Системы и средства информатики. Специальный выпуск. Научно-методологические проблемы информатики./ под ред. К.К. Колина. – М.: ИПИ РАН, 2006. С.7-58.
19. Floridi L., Semantic Conceptions of Information First published Wed Oct 5, 2005; substantive revision Fri Jan 28, 2011 <http://plato.stanford.edu/entries/information-semantic>
20. V. Ya. Tsvetkov. Semantic Information Units as L. Floridi's Ideas Development // European Researcher, 2012, Vol. (25), № 7, p.1036- 1041
21. Хлебников . Г.В. Философия Лучано Флориди//Доклад на заседании семинара “Методологические проблемы наук об информации”.- М: ИНИОН, 2011.

REFERENCES

1. С.Е. Shannon, (1948), "A Mathematical Theory of Communication", Bell System Technical Journal, vol. 27, pp. 379–423 & 623–656, July & October, 1948
2. Chernyi Yu.Yu. Polysemy in science, when it is harmful? (for example computer science). *Otkrytoe obrazovanie - Open Education*, 2010, no.6, pp.97-106 (in Russian).
3. Cover T. M., Thomas J. A. Elements of Information Theory. – New York: J. Wiley , Sons, 1991.
4. V. Y. Tsvetkov. Information Interaction as a Mechanism of Semantic Gap Elimination. *European Researcher*, 2013, Vol.(45), no.4-1, p.782-786.
5. Ershov A.P. Computer science: the subject and the concept. *Kibernetika. Stanovlenie informatiki* [Cybernetics. Becoming a science]. Moscow, Nauka, 1986. pp. 28-31.
6. Buinevich M.V., Kudzh S.A. Choosing a rational variant of information technology interaction in distributed systems. *Vestnik INZhEKONa. Serii: Tekhnicheskie nauki - Journal ENGECON. Series: Engineering*, 2011, no.8. pp.37-47 (in Russian)
7. Maksudova L.G., Savinykh V.P., Tsvetkov V.Ia. Integration of Sciences about the world in Geoinformatics. *Issledovanie Zemli iz kosmosa - Study of Earth from space*, 2000, no.1. pp.46-50 (in Russian).
8. Benjamin Kuipers. Modeling Spatial Knowledge (1978) // Cognitive Science - №2. - p. 129-153.
9. Antony Galton. Spatial and temporal knowledge representation // Earth Science Informatics, September, 2009, Volume 2, Issue 3, pp 169-187
10. Hill Linda L. Georeferencing: The Geographic Associations of Information – MIT Press Cambridge, Massachusetts, London, England- 2009, - 272 p

11. Maiorov A.A., Tsvetkov V.Ia. Georeferense as the use of spatial relationships in Geoinformatics. *Geodeziia i aerofotos"emka – Geodesy and aerial photography*, 2012, no.3. pp.87-89 (in Russian).
12. Thesaurus. Moscow, Soviet Encyclopedia, 1969. 1978 p.
13. Ozhiganov A.A., Tarasiuk M.V. *Peredacha dannykh po diskretnym kanalām* [The transfer of data from discrete channels]. Saint Petersburg, 1999. 102 p.
14. Kheld G. *Tekhnologii peredachi dannykh* [Data Technologies]. Saint Petersburg, Piter, 2003. 720 p.
15. Winner N. *Cybernetics or Control and Communication in the Animal and the Mashine*. The Technology Press and John Wiley & Soris Inc. New York – Herman et Cie, Paris, 1948. 99 p.
16. Kharkevich A.A. *Izbrannye trudy v 3-kh tomakh. Tom 3. Teoriia informatsii. Opoznanie obrazov* [Selected works in three volumes. Volume 3. Information Theory. Recognition of images]. Moscow, Nauka, 1973. 524 p.
17. Iuzvishin I.I. *Informatsiologiia ili zakonomernosti informatsionnykh protsessov* [Informationology or patterns of information processes]. Moscow, Vysshiaia Shkola, 2001.
18. Kolin K.K. *Stanovlenie informatiki kak fundamental'noi nauki i kompleksnoi nauchnoi problemy* [The formation of computer science as basic science and complex scientific problems] / *Sistemy i sredstva informatiki. Spetsial'nyi vypusk. Nauchno-metodologicheskie problemy informatiki* [Systems and Informatics. Special Issue. Scientific and methodological problems of computer science]. Moscow, IPI RAN, 2006, pp.7-58.
19. Floridi L., *Semantic Conceptions of Information* First published Wed Oct 5, 2005; substantive revision Fri Jan 28, 2011. Available at: <http://plato.stanford.edu/entries/information-semantic> (accessed 14 November 2013).
20. V. Ya. Tsvetkov. *Semantic Information Units as L. Floridi's Ideas Development*. *European Researcher*, 2012, Vol.(25), no.7, p.1036-1041.
21. Khlebnikov G.V. *Filosofia Luchano Floridi* [Luciano Floridi Philosophy] / *Doklad na zasedanii seminara "Metodologicheskie problemy nauk ob informatsii"* [Talk at the workshop "Methodological problems in information sciences"]. Moscow, INION, 2011.

Информация об авторах

Кудж Станислав Алексеевич (Россия, Москва) – Профессор, доктор технических наук, ректор Московского государственного технического университета радиотехники, электроники и автоматики. Автор свыше 100 печатных работ. E-mail: mirearec1@yandex.ru

Цветков Виктор Яковлевич (Россия, г. Москва) – Профессор, доктор технических наук, советник ректора. Московский государственный университет геодезии и картографии. E-mail: cvj2@mail.ru

Information about the authors

Kudzh Stanislav Alekseevich (Russia, Moscow) – Professor, doctor of technical sciences, rector of Moscow State Technical University of Radio Engineering, Electronics and Automation. The author of over 100 publications. E-mail: mirearec1@yandex.ru

Tsvetkov Viktor Iakovlevich (Russia, Moscow) – Professor, doctor of technical sciences, advisor to the rector. Moscow State University of Geodesy and Cartography. E-mail: cvj2@mail.ru