

В.С. ВОРОНЦОВ

ИНФОРМАТИКА

НОВЫЙ СИСТЕМАТИЗИРОВАННЫЙ ТОЛКОВЫЙ **С**ЛОВАРЬ



Воройский Ф.С.

**Информатика. Новый
систематизированный
толковый
словарь-справочник
(Введение в
современные
информационные и
телекоммуникационные
технологии в
терминах и
фактах)**



МОСКВА
ФИЗМАТЛИТ ®

УДК 802.0 - 323.2-82:002+002.5(038)

ББК 32.97я2

В 75

Воройский Ф. С. Информатика. Новый систематизированный толковый словарь-справочник (Введение в современные информационные и телекоммуникационные технологии в терминах и фактах). — 3-е изд., перераб. и доп. — М.: ФИЗМАТЛИТ, 2003. — 760 с. — ISBN 5-9221-0426-8.

Толковый словарь-справочник, содержит 16 тыс. русско- и англоязычных терминов, расположенных в логической последовательности по тематическим направлениям: «Основы информационной технологии», «Автоматизация информационных процессов и автоматизированные системы», «Средства технического обеспечения автоматизированных систем», «Средства программного обеспечения автоматизированных систем», «Мультимедиа, гипермедиа, виртуальная реальность и машинное зрение», «Сетевые технологии обработки и передачи данных».

Словарь-справочник ориентирован на широкий круг читателей — преимущественно не специалистов в области вычислительной техники и программирования, использующих в своей деятельности современные программно-технические средства, информационные и телекоммуникационные технологии. Он может использоваться в учебных процессах профессорско-преподавательским составом и студентами вузов, а также квалифицированными техническими специалистами при разработке и эксплуатации различного рода автоматизированных информационных, библиотечно-информационных и других систем по смежным для них областям знаний.

Второе издание — 2001 г.

ОГЛАВЛЕНИЕ

О словаре-справочнике и его авторе	7
Предисловие к третьему изданию	9
I. Основы информационной технологии	13
1.1. Данные, информация, знания, логика	13
1.2. Информационные ресурсы, теория информации, информатика	20
1.3. Документы, документация, носители информации	23
1.4. Средства информационного обеспечения автоматизированных систем и их структура	28
1.4.1. Информационные элементы и их виды	28
1.4.2. Запись, файл, массив, ключ	31
1.4.3. База данных, банк данных	35
1.4.4. Структуры, модели данных и связанные с ними термины	39
1.4.5. Термины, связанные с видами баз данных	53
1.4.6. Формат, поле данных и связанные с ними термины	57
1.5. Информационная технология	67
1.5.1. Общие термины	67
1.5.2. Обработка данных и связанные с нею термины	70
1.5.3. Ввод данных и связанные с ним термины	78
1.5.4. Информационный поиск (общие термины)	84
1.5.5. Информационно-поисковые языки, словарные средства	89
1.5.6. Индексирование данных, поисковый образ документа	95
1.6. Безопасность информационной технологии	103
1.6.1. Общие термины	103
1.6.2. Кодирование/декодирование данных	109
1.6.3. Криптология и связанные с нею понятия	112
II. Автоматизация информационных процессов и автоматизированные системы	118
2.1. Общие термины	118
2.2. Автоматизированные системы	125

2.2.1. Видовой состав вычислительных и автоматизированных систем	125
2.2.2. Функционально-ориентированные автоматизированные системы	133
2.2.3. Персонал и пользователи автоматизированных систем	149
2.2.4. Лингвистическое и информационное обеспечение автоматизированных систем	158
2.2.5. Проектирование автоматизированных систем	163
III. Техническое обеспечение автоматизированных систем.	169
3.1. ЭВМ, их виды и общая классификация	169
3.2. Архитектура, конфигурация, платформа ЭВМ	176
3.3. Персональные ЭВМ	180
3.4. Системный блок и элементы его конструкции	189
3.4.1. Процессоры, их виды и связанные с ними термины . .	190
3.4.2. Память ЭВМ — общие термины	200
3.4.3. Функциональные устройства памяти ЭВМ	210
3.4.4. Адаптеры, интерфейсы и связанные с ними термины .	221
3.4.5. Платы, порты, шины, гнезда	231
3.5. Периферийные (внешние) устройства ЭВМ	242
3.5.1. Внешняя память ЭВМ, накопители и связанные с ними термины	242
3.5.2. Компакт-диски и связанные с ними термины	265
3.5.3. Устройства ввода данных, манипуляторы	273
3.5.4. Устройства вывода данных	283
3.5.5. Модемы, шифраторы, источники бесперебойного питания	301
3.6. Портативные ПК И PC-карты	304
3.6.1. Виды портативных ПК	304
3.6.2. PC-карты	310
3.7. Микроэлектронная база ЭВМ	316
IV. Программное обеспечение автоматизированных систем	325
4.1. Алгоритмы, программы, программирование	325
4.1.1. Общие термины	325
4.1.2. Языки программирования	330

4.1.3. Связанные с программированием термины	343
4.2. Общее программное обеспечение	353
4.2.1. Операционные системы	354
4.2.2. Сервисные средства общего программного обеспечения	364
4.3. Прикладное программное обеспечение автоматизированных систем	365
4.3.1. Общие термины	365
4.3.2. Прикладные программы	368
4.3.3. Компьютерные вирусы и антивирусы	373
4.4. Термины, связанные с работой программных средств . . .	377
4.4.1. Некоторые общие термины	377
4.4.2. Архивация, сжатие-восстановление записей данных .	381
4.4.3. Доступ, адрес и связанные с ними термины	391
V. Мультимедиа, гипермедиа, виртуальная реальность, машинное зрение	404
5.1. Системы мультимедиа и связанные с ними термины . . .	404
5.2. Средства обеспечения музыкального и речевого сопровождения	409
5.2.1. Общие понятия и термины	409
5.2.2. Звуковые файлы и их форматы	414
5.3. Машинная (компьютерная) графика	423
5.3.1. Общие термины	423
5.3.2. Графические файлы и форматы	427
5.3.3. Технология машинной графики	438
5.4. Компьютерное видео, цифровое телевидение и анимация	444
5.4.1. Общие термины	444
5.4.2. Элементы технологии видео	449
5.4.3. Технология анимации	453
5.5. Виртуальная реальность, параллельный мир	458
5.6. Компьютерное зрение	461
VI. Сетевые технологии обработки и передачи данных . . .	464
6.1. Общие термины	464
6.2. Локальные вычислительные сети	469

6.3. Распределенные и глобальные вычислительные сети	480
6.3.1. Общие термины	480
6.3.2. Интранет	489
6.3.3. Интернет	497
6.3.4. Web-технология — «Всемирная паутина»	512
6.3.5. Субъекты юридических отношений в Интернете	523
6.3.6. Сотовая связь и компьютерная телефония	527
6.3.7. Интегрированные службы цифровых сетей — ISDN . .	531
6.3.8. Технологии передачи данных по каналам Интернет . .	534
6.3.9. Телекоммуникационное оборудование зданий	544
6.3.10. Разработки технических средств и их комплексов, основанных на использовании сетевых технологий .	549
6.3.11. Основные стандарты	557
6.4. Средства и технология защиты вычислительных сетей . .	572
VII. Компьютерный сленг	579
VIII. Пиктограммы, используемые в электронной почте . .	608
IX. Сокращения слов и выражений, используемые в Ин- тернете	609
Список литературы	613
Алфавитный указатель русскоязычных терминов	655
Алфавитный указатель англоязычных терминов	705

О СЛОВАРЕ-СПРАВОЧНИКЕ И ЕГО АВТОРЕ

Совсем немного в отечественной и мировой издательской практике найдется литературных источников, относящихся к каким-либо областям науки и техники, которые вызывали бы одинаковый интерес как непрофессионалов, так и самых продвинутых специалистов. Два предыдущих издания словаря-справочника, выпущенные в 1998 и 2001 гг. и быстро реализованные без широкой и шумной рекламной компании, убедительно доказывают, что он относится именно к такому классу литературы. Сказанное подтверждает и значительное число предложений, поступивших Ф.С. Воройскому в последние годы, о размещении его словаря-справочника на Веб-сайтах различных организаций, в том числе и в коллекции «Избранные справочники» широко известного в нашей стране и за рубежом справочно-энциклопедического проекта «Рубрикон» (<http://www.rubricon.ru>).

По отзывам ряда читателей, в основе успеха словаря-справочника помимо актуальности темы для России, быстро набирающей темпы автоматизации информационных процессов, лежат такие его особенности, как форма представления материала (можно читать по разделам и отдельным терминам); емкое и профессиональное изложение материала; наличие справочных данных о важнейших объектах и явлениях, связанных с информатикой и современной информационной технологией; полнота и оперативность отображения новых достижений науки и техники в весьма широкой тематической области и, наконец, стиль изложения словарных статей, доступный разным категориям читателей.

Надеемся, что и третье издание словаря-справочника, существенно дополненное и переработанное, вызовет положительный отклик читательской аудитории и представляем его автора.

Воройский Феликс Семенович, кандидат технических наук, профессор, капитан 2 ранга — инженер в отставке. В 1953 г. окончил Высшее военно-морское училище связи им. Попова (г. Петродворец Ленинградской обл.) по специальности «Офицер радиотехнической службы корабля». С 1953 по 1958 гг. служил начальником радиотехнической службы эскадренного миноносца на Черноморском флоте.

В 1961 г. с отличием окончил Военно-морскую академию (г. Ленинград) по специальности «Военный инженер по радиоэлектронике» и с 1961 по 1974 гг. служил в научно-исследовательском центре Министерства обороны, где в 1965 г. защитил диссертацию.

Информационной деятельностью занимается с 1976 г.:

— 1976–1978 гг. — внештатный референт редакции «Машиностроение» ВИНТИ;

— 1978–1997 гг. — старший научный сотрудник, зав. сектором и зав. лабораторией комплектования и эксплуатации машинных массивов Интегрированной системы информационного обеспечения (ИСИО) ВИМИ (в то время — головной информационный орган оборонных отраслей промышленности СССР);

— 1988–1994 гг. — заведующий кафедрой информационного и библиотечного обслуживания, директор Научно-информационного и вычислительного центра ИПКИР.

В начале 1993 г. разработал и приступил к реализации Программы автоматизации массовых (публичных) библиотек Москвы. Под его непосредственным руководством Практически с нулевого уровня была заложена основа автоматизации Центральной городской публичной библиотеки Москвы им. Н. А. Некрасова и 16-ти центральных районных библиотек.

С октября 1994 г. работает в ГПНТБ России; заведует сектором анализа, исследований и экспертиз по проблемам автоматизации библиотечно-информационных процессов.

Ф. С. Воройский — профессор кафедры информационных технологий и электронных библиотек Московского государственного университета культуры и искусств.

Начиная с 1985 г. и по настоящее время принимает участие в разработке ряда крупных проектов в области автоматизации информационных и библиотечных процессов; является ответственным исполнителем проектов создания Корпоративной сети публичных библиотек Москвы, Российского центра корпоративной каталогизации и Единой библиотеки г. Обнинска — единой университетской библиотеки и др.; член Совета и председатель НТС «Информационные ресурсы и обслуживание» Ассоциации российских библиотечных консорциумов (АРБИКОН); эксперт РФФИ; член постоянных оргкомитетов Международных конференций «Крым» и «Либком»; автор более 200 научных работ, в том числе 5-ти монографий, а также 15-ти авторских свидетельств на изобретения и промышленные образцы изделий.

ПРЕДИСЛОВИЕ К ТРЕТЬЕМУ ИЗДАНИЮ

Вычислительная техника и телекоммуникации глубоко вошли во все сферы человеческой деятельности и стали их неотъемлемой частью. Поэтому овладение немалым объемом знаний в области современных информационных технологий и инструментальных средств их обеспечения столь же необходимо каждому, как и элементарная грамотность. В первую очередь это касается терминологии. К сожалению, здесь возникает много проблем, которые связаны не только с тем, что информатика является относительно новой наукой. Не менее существенно и то, что, будучи существенно более динамичной, чем любая другая область знаний, она вводит как в профессиональный, так и бытовой язык большое количество новых понятий и терминов, многие из которых не отражены в отечественных стандартах и трактуются далеко не однозначно.

По-видимому, определенную помощь в преодолении упомянутых проблем могут оказать существующие учебные пособия по информатике, специальная литература по различным программным и техническим средствам, а также терминологические словари. Однако пособия читать довольно трудно, поскольку они ориентированы не на терминологические проблемы. Словари не дают возможности получить систематизированные знания, так как служат лишь для толкования отдельных терминов, понимание которых может быть затруднено, поскольку требует определенных исходных знаний, включающих представление и о других взаимосвязанных понятиях и терминах, извлечь которые из существующих словарей — задача далеко не простая, более того, часто она просто невыполнима.

Так возникла идея создания тематически организованного словаря, который можно было бы читать «поряд» или по разделам, а для обращения к отдельным терминам — использовать алфавитные указатели.

В процессе работы над первыми двумя изданиями словаря появилась необходимость расширения многих словарных статей. Кратких определений терминов (дефиниций) оказалось недостаточно, что часто было связано с неодинаковым их толкованием рядом уважаемых авторов публикаций и специалистов, практикующим в разных областях информатики. В дальнейшем представилось интересным дополнить словарные статьи наиболее важными справочными сведениями о характере объектов описания, в том числе технологий, программных и технических средств, стандартов, ведущих организаций и фирм, ученых и разработчиков. В свою очередь это привело к необходимости введения в словарные статьи ссылок на использованную литературу.

В результате появился «Систематизированный толковый словарь-справочник», отличный по жанру от традиционных толковых словарей. По отзывам ряда читателей преимущественно из числа профессорско-преподавательского состава различных учебных заведений, отдельные разделы словаря они используют в качестве основы учебных курсов по дисциплинам, связанным с информатикой, современными информационными технологиями и инструментальными средствами их обеспечения. Автор также использует материалы данного словаря-справочника в курсах лекций, которые он читает в Московском государственном университете культуры и искусств (МГУКИ) и в системе повышения квалификации библиотечных и информационных работников. Таким образом, словарь-справочник стал также «Введением в современные информационные и телекоммуникационные технологии в терминах и фактах».

Как уже отмечалось, изменения в области современных информационных технологий и технических средств их обеспечения происходят стремительно. Многие из того, что всего год назад относилось к «последнему слову» в данной области, успело «устареть» и заменено или отошло на второй план.

Третье издание отличается от выпущенного в 2001 г. предыдущего издания следующим:

1) существенно дополнены и переработаны более 5 тыс. статей, относящихся к техническим и программным средствам, стандартам и другим объектам описания, в соответствии с изменениями, происшедшими в последние два года;

2) количество терминов и соответствующих им словарных статей увеличено почти на 1,5 тыс.; общее их число — более 16 тыс.;

3) количество ссылок на первоисточники с 644 (во втором издании) увеличено до 839, при этом добавлены или обновлены более 280 ссылок на издания 2001–2003 гг.;

4) наиболее существенно дополнены новыми терминами разделы: 2.2. «Автоматизированные системы»; 6.3. «Распределенные и глобальные вычислительные сети» и 6.4. «Средства и технологии защиты вычислительных сетей»;

5) изменены состав и структура ряда разделов, в частности, выделены в самостоятельные разделы: 1.5. «Информационная технология»; 1.6. «Безопасность информационной технологии»; 5.6. «Компьютерное зрение»; 6.3.5. «Субъекты юридических отношений в Интернете»; 6.3.6. «Сотовая связь и компьютерная телефония»; 6.3.8. «Технологии передачи данных по каналам Интернета»; 6.3.9. «Телекоммуникационное оборудование зданий» и 6.3.10. «Разработки технических средств и комплексов, основанных на использовании сетевых технологий».

В тексте статей словаря-справочника часто встречаются слова **объект**, **тип объекта**, **понятие**, **термин** и **дефиниция**; определим их значение:

— **объект** [object] — любой материальный предмет или явление, с которым связана познавательная, информационная или любая другая практическая деятельность человек;

— **тип объекта** [object type] — обобщенное имя какого-либо множества объектов, характеризующееся определенным набором признаков (в том числе свойств и характеристик), например «автомобиль», «гроза», «компьютерные игры» и т. д.;

— **понятие** [concept, notion] — форма мысли, в которой отражаются общие и существенные признаки предмета или явления; целостная совокупность суждений о каком-либо объекте, отображающая его сущность и являющаяся результатом познания **объекта**;

— **термин** [term] — слово или словосочетание, являющееся обозначением строго определенного **понятия**;

— **дефиниция** [definition, concept definition] — краткое научное определение какого-либо **понятия**; выделяет существенные черты определенного **объекта** таким образом, что он отличается от любых других объектов, и выражается законченным предложением, построенным по правилам логики.

Принципы построения словаря-справочника:

1. **Словарные статьи** вместо общепринятого алфавитного расположения представлены в тематических разделах в логической последовательности: от простого понятия к более сложному или от общего к частному.

2. **Пояснительная часть словарных статей** жестко не формализована, при необходимости расширена и содержит связанные с основным понятием термины¹⁾, а также варианты толкования отдельных терминов, существующие в различных нормативных и других документальных источниках.

3. **Варианты терминов — синонимы**, обозначающие одно и то же понятие, отделены друг от друга запятой, точкой с запятой или косой чертой.

4. **Составные части сложных терминов**, которые используются как синонимы или носят факультативный характер (т. е. могут опускаться), заключены в круглые скобки.

¹⁾ Термины, выделенные шрифтом, как правило, могут быть найдены по алфавитным указателям.

5. **Русскоязычные термины** сопровождаются англоязычными эквивалентами, заключенными в квадратные скобки. В случае отсутствия устоявшегося англоязычного эквивалента термина, он не указывается.

6. **Переводы англоязычных терминов**, не имеющих устоявшихся русскоязычных эквивалентов, выделены кавычками и представлены с прописной буквы.

7. **Англоязычные термины**, на которые русскоязычный эквивалент отсутствует и вводить его признано нецелесообразным, помечаются знаком * (звездочка).

8. **Сокращенные термины** представлены, как правило, через запятую вместе с их развернутой формой.

9. **Сокращенные англоязычные термины** и их развернутая форма разделяются круглыми скобками.

10. **Составные термины** могут содержать англо- и русскоязычные части, если каждая из них в отдельности и их комбинации в составном термине являются широкоупотребительными, например **Web-сервер**¹⁾, индекс **iCOMP**, **USB-шина**, **накопитель CD-RW** и т. п.

11. В словарных статьях используются два близких по своему значению сокращения: **ПК** (*рус.*) — по отношению ко всем персональным компьютерам и **РС** (*англ.*) — только по отношению к IBM-совместимым ПК.

12. **Поиск терминов** (в том числе аббревиатур) можно провести по алфавитным русско- и англоязычному указателям, помещенным в конце книги. В алфавитные указатели включены также фамилии разработчиков, наименования фирм, крупных международных и национальных организаций, а также наиболее распространенных программных и технических средств.

13. **Выделенные в пояснительной части статей полужирным шрифтом термины** обращают внимание читателей на возможность их поиска по алфавитным указателям.

14. В **специальные разделы выделены** компьютерный сленг (VII), пиктограммы, используемые в электронной почте (VIII) и сокращения слов и выражений, используемые в Интернете (IX).

Ф. С. Воройский

¹⁾ В настоящее время в различных публикациях все чаще встречается русскоязычное написание — **Веб**. Однако, учитывая профессиональную специфику словаря-справочника, мы преимущественно используем в нем оригинальное англоязычное — международное написание — **Web**.

I. ОСНОВЫ ИНФОРМАЦИОННОЙ ТЕХНОЛОГИИ

1.1. ДАННЫЕ, ИНФОРМАЦИЯ, ЗНАНИЯ, ЛОГИКА

ДАННЫЕ [data, information]

Сведения, факты, показатели, выраженные как в числовой, так и любой другой форме. Указанный в скобках англоязычный эквивалент термина показывает, что он весьма часто рассматривается как синоним термина **информация**. Так, словосочетание **data system** и производные от него могут соответствовать одному из совсем не однозначных, хотя и связанных между собой понятий — **система данных** и **информационная система**. Аналогичное явление прослеживается при русскоязычном использовании этого термина. Например, в одном из изданий «Словаря терминов по информатике» данные трактуются как «Информация, представляющая собой сведения . . . », а в ГОСТ 15971-84 [2] — как «Информация, представленная на материальных **носителях**».

С целью смыслового разделения понятий «информация» и «данные» Ассоциация стандартов Франции (АФНОР) дает следующее определение: «Данные — факт, понятие или инструкции, представленные в условной форме, удобной для пересылки, интерпретации и обработки человеком или автоматизированными средствами».

Согласно другому важному для понимания этого термина определению: «Данные — некоторый факт, то на чем основан вывод или любая **интеллектуальная система**» [1]. Компонентами данных являются цифры и символы естественного языка или их кодированное представление в виде строки двоичных **битов**.

БИТ [bit — от англ. binary digit]

Простое двоичное число (цифра или символ), принимающее значения 1 или 0 и служащее для записи и хранения данных в ЭВМ.

Бит является минимальной двоичной единицей измерения **энтропии и количества информации** в ЭВМ, соответствующей одному **двоичному разряду**. Энтропия сообщения, выраженная в битах, определяется средним числом символов, необходимых для записи этого сообщения. Определенное количество бит составляет размер других единиц — **двоичных слов**, в том числе — **байта, килобайта, мегабайта** (см. далее) и т. д.

Байт [byte] — двоичное слово, способное записывать и хранить в памяти ЭВМ один буквенно-цифровой или другой символ данных. Каждый символ записывается в виде набора двоичных цифр (битов) при помощи определенного кода, например **ASCII**. Количество бит в байте определяет его разрядность, которая может составлять 8, 16, 32 и т. д. Соответственно байт называют 8-разрядным, 16-разрядным и т. д. Один 8-разрядный байт может определять 256 разных значений, например десятичных чисел от 0 до 256. Увеличение разрядности ведет к соответствующему увеличению числа возможных вариантов комбинаций, кодируемых одним байтом. Например, 16-разрядным — до 65536 или 2^{16} , 32-разрядным — до 2^{32} и т. д.

Килобайт, Кбайт [kilobyte] — единица измерения емкости памяти или длины записи, равная 1024 байтам. Часто под килобайтом понимается также величина, равная 10^3 байт.

Мегабайт, Мбайт [megabyte] — единица измерения емкости памяти или длины записи, равная 1024 килобайтам. Часто под мегабайтом понимается также величина, равная 10^3 килобайт или 10^6 байт.

Гигабайт, Гбайт [gigabyte] — единица измерения емкости памяти или длины записи, равная 1024 мегабайтам. Часто под гигабайтом понимается также величина, равная 10^3 мегабайт, 10^6 килобайт или 10^9 байт.

ЗНАК [charter]

Один символ, который может быть представлен и воспринят ЭВМ. К знакам относятся буквы, цифры, пробелы, знаки препинания, специальные символы (например математические, кодовые и т. п.).

Информация [Information] от лат. *Informatio* — *разъяснение, осведомление*]

Данный термин и отражаемое им понятие являются сегодня одними из самых распространенных. Сказанное относится к их использованию, как на бытовом, так и на профессиональном уровнях.

Существует множество различных определений этого понятия, например:

- «Информация — содержание какого-либо сообщения, сведения о чем-либо, рассматриваемые в аспекте их передачи в пространстве и времени . . . » [2];

- «Информация — сведения, подлежащие передаче»;

- «Информация — это значение, вкладываемое человеком в данные на основании известных соглашений, используемых для их представления»;

• «Информация — содержание, значение данных, которое видят в них люди. Обычно данные состоят из фактов, которые ставятся информацией в определенном контексте и понятны людям» [35].

Следует упомянуть здесь также классическое определение **К. Шеннона**, в соответствии с которым информация — это то, что сокращает степень неопределенности (у Шеннона — **энтропии**) у ее адресата о каком-либо объекте (в том числе явлении, передаваемом сигнале и т. п.). Другими словами, по Шеннону информация это то, что увеличивает степень знания ее адресатом интересующих его объектов окружающего мира. В указанном контексте количество информации можно даже рассчитать, например, по увеличению вероятности успешного решения поставленной задачи.

С чем же связано различное представление понятия **информация** разными его пользователями, включая профессионалов?

Во-первых, с его сложной и неоднозначной сущностью, которая к тому же имеет тенденцию достаточно быстро изменяться в ходе научно-технического прогресса. К примеру, в 1992 г. в журнале «Научно-техническая информация» в статье одного уважаемого автора достаточно убедительно доказывался тезис о том, что информация отнюдь не всегда повышает вероятность успешного решения некоторых прикладных задач.

Во-вторых, с тем, что цитируемые и другие определения этого понятия вычлениют только те его признаки, которые служат достижению конкретных целей или соответствуют контексту документов, в которых они опубликованы. Так, наука кибернетика расширенно толкуя понятие **информация**, вывело его за пределы человеческой речи и других форм коммуникаций между людьми, связав его с целенаправленными системами любой природы — биологической (например, биотоки в организмах, связи в гинетических механизмах), технической (например, сигналы в электрических сетях) и социальной (движение человеческих знаний в общественных системах).

Мы остановимся только на тех признаках понятия **информация**, которые необходимы большинству наших читателей, которые, как мы полагаем, тем или иным образом преимущественно связаны с библиотечной и информационной сферами деятельности.

Кратко эти признаки информации можно сформулировать следующим образом: «Информация это сведения или данные, объективно отражающие различные стороны и элементы окружающего мира и деятельности человека на определенном этапе развития общества, представляющие для него какой-либо интерес и материализованные в форме, удобной для использования, передачи, хранения и/или обработки (преобразования) человеком или автоматизированными средствами».

Если на бытовом уровне смешение понятий **данные** и **информация** вполне допустимо, то для профессионалов это может привести и приводит к серьезным последствиям.

Чтобы стать информацией, данные должны правильно отражать объекты описания, в противном случае мы будем иметь дело с **дезинформацией** (ее англоязычные эквиваленты: **false information, misleading information**). Сама по себе «правильность отражения действительности» в соответствии с теорией познания всегда носит условный характер, поскольку связана с уровнем развития знаний на данном этапе развития общества или отдельных социальных групп и индивидуумов. Так, состав и точность данных, которыми владеют или которые необходимы различным организациям и лицам об одном и том же объекте, будут существенно различаться в зависимости от образовательного, возрастного, социального статусов субъекта информирования, а также целей их использования (например, для «повышения общей эрудиции» или для решения научных, технических, производственных, коммерческих и др. задач). В указанном плане данные, которые для одного субъекта будут представляться вполне точными, для другого — могут оказаться грубой «дезой».

Чтобы стать информацией, данные должны представлять для субъекта информирования «определенный интерес» и «новизну». Последнее означает, что они должны быть для него связаны с необходимостью решения каких-либо практических или других задач и сокращать «степень неопределенности» об объекте «интереса» (вспомним определение К. Шеннона!). В указанном плане информация помимо того, что она адресно «прибавляет знания» об интересующем объекте, должна доставляться своевременно. К примеру, сообщение о том, что «Сегодня в городе проливной дождь, гололед (или другие напасти)», полученное нами после того, как мы успели выйти из дома и промокнуть или забуксовать на дороге — информацией не является, как не является информацией и сообщение о погоде, скажем, в Чили или деревне Гадюкино, если там не живут наши близкие родственники и мы не собираемся туда ехать. Для нас эти данные являются тем, что в информатике принято называть **информационным шумом**. Однако для кого-то в зоне действия средств массовой информации (кто еще не вышел на улицу или живет в упомянутых местах) эти же сведения — информация.

«Информационным шумом» являются также сообщения и данные, не представляющие для субъекта информирования предмета новизны (другими словами — этими данными он уже владеет). Вспомним в этой связи о часто повторяющейся и вызывающей только раздражение рекламе даже нужных нам товаров и услуг!

Приведенные примеры, носящие откровенно бытовой характер, могут быть заменены другими из любой сферы деятельности человека. Смысл их от этого не изменится: данные, переданные «не по назначению», не своевременно или не представляющие новизну, являются информационным шумом.

Понятие информационный шум может быть также распространено на данные, «не удобные» (см. определение) для «использования, передачи, хранения и/или обработки», поскольку и в этом случае они приводят к бесполезным, а возможно и вредным затратам материальных, временных и других ресурсов.

Подводя итог сказанному, можно существенно сократить определение понятия информации: ***«Информация — это данные, необходимые или полезные тому, кому они передаются».***

И в заключение — небольшое отступление, адресованное нетерпеливому читателю, раздраженному кажущейся ему пространностью этой словарной статьи: «Скажите, с чем связаны бытовавшие и, увы, существующие еще способы оценки эффективности работы библиотек, информационных органов, а также отдельных их служб и работников по количеству выданных пользователям (или читателям) документов, справок, библиографических описаний и т. п. без учета соответствия содержащихся в них данных признакам, делающим их информацией?».

Термин информация в сочетании с различными прилагательными широко употребляется в следующих случаях.

1. *Сведения или данные, полученные в процессе какого-либо вида деятельности*, отражают результаты этого вида деятельности или имеют отношение к ней и предназначены для **справочно-информационного обслуживания** и/или **информационного обеспечения** (см. далее) заинтересованных пользователей, например, **научно-техническая информация** [scientific-technical information], **юридическая информация** [juridical information], **патентная информация** [patent information] и т. п.

2. *Сведения или данные имеют определенное назначение*, напр.:

- **справочная информация** [reference information] — сведения или данные для выдачи справок о чем-либо;

- **сигнальная информация** [alert information, current awareness information] — информация, предназначенная для быстрого предварительного оповещения.

3. *Словарная статья, определяет характер, принадлежность, форму или вид данных*, используемых в информационном процессе, например:

- **априорная информация** [aprior information] и **апостериорная информация** [aposterior information] — соответственно —

данные, имевшиеся до проведения какого-либо опыта или другого действия, и сведения, полученные после его выполнения;

- **коммерческая информация** [**commercial information**] — данные, сведения и содержащие их документы, являющиеся объектом продажи их собственником;

- **личная информация** [**private information**] — сведения (данные) о гражданах и организациях, затрагивающие их интересы и запрещенные для распространения без их согласия;

- **библиографическая информация** [**bibliographic information**] — библиографические данные, описания и их перечни;

- **графическая информация** [**graphical (pictorial, image, pattern) information**] — сведения или данные, представленные в виде схем, эскизов, изображений, графиков, диаграмм, символов;

- **ретроспективная информация** [**retrospective information**] — сведения, содержащиеся в накопленных за два и более лет массивах данных или полученные в результате поиска в этих массивах (так называемого **ретроспективного поиска**).

4. *Словарная статья характеризует средства закрепления, отображения и/или передачи данных, например:*

- **документальная информация** [**documentary information**] — сведения, закрепленные на каком-либо материальном **носителе**; содержание документа или текста;

- **устная информация** [**oral information**] — содержание устного сообщения и т. п.

ЗНАНИЯ [**knowledge**]

Совокупность хранимых в **базах знаний** или памяти человека (специалиста-эксперта) фактов (данных, сведений) о некоторой предметной области, их взаимосвязей и правил, которые могут быть использованы для получения новых фактов и/или решения каких-либо задач, связанных с этой предметной областью. Знания отражают множество возможных ситуаций, характеризующихся состоянием и конкретной реализацией объектов определенного типа, способы перехода от одного описания объекта к другому. Для знаний характерна внутренняя интерпретируемость, структурированность, связанность и активность. Условно можно записать: «знания = факты + убеждения + правила» [265].

Различают следующие виды знаний [722]:

- **базовые (фундаментальные) знания** [**deep knowledge**] — систематизированные знания, основанные на модели, описывающей все значимые аспекты некоторой предметной области, которые описывают ее сущности, их свойства и различные связи между ними;

- **компилятивные знания** [**compiled knowledge**] — знания, полученные на основе уже известных знаний, путем их структурирования и/или систематизации в форме, необходимой для их использования в новых целях;
- **неполные знания** [**incomplete knowledge**] — знания, из которых не могут быть получены некоторые значимые факты в данной предметной области;
- **нечеткие знания** [**fuzzy knowledge**] — знания, основанные на **нечеткой логике** (см. далее);
- **поверхностные знания** [**surface knowledge**] — неполные, фрагментарные знания, часто связанные с недостаточной изученностью предметной области. Тем не менее, такие знания могут быть успешно использованы в случаях, не затрагивающих ситуаций, которые ими не охвачены;
- **предметные знания** [**domain knowledge**] — знания, относящиеся к конкретной узкой предметной области;
- **процедурные знания** [**procedural knowledge**] — знания, воплощенные в компьютерных программах для решения тех или иных задач.

ЛОГИКА [logic]

Наука о законах и формах мышления, методах познания и условиях определения истинности знаний и суждений.

Нечеткая логика [**fuzzy logic**] — *в математике и вычислительных системах*: форма представления **знаний** или данных, связанных с описанием различных объектов понятиями вида **тяжелый, громкий, горячий** и т. п., имеющими неточные или неопределенные значения. Более строгое определение значения указанных понятий возможно только с привлечением ряда дополнительных сведений или данных, входящих в **нечеткие множества** и составляющих перечни дополнительных данных или условий. Центральным понятием нечеткой логики является понятие **вероятность члена множества**, определяющее степень правомерности отнесения данного члена к указанному множеству. Например, для члена нечеткого множества, относящего предметы к тяжелым, вес 20 кг может иметь такое значение с вероятностью 90 % для миниатюрных женщин и, скажем, только для 20 % мужчин.

Нечеткая логика широко используется в различного рода **экспертных системах** для автоматизированного принятия решений, близких к человеческим, на основе адекватного реагирования на сигналы, поступающие от связанных с ними датчиков, а также команды с пульта управления.

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

(от лат. *intellectus* — разум, рассудок)

1. Искусственная система, имитирующая решение человеком сложных задач, связанных с различными видами его деятельности.

2. Научное направление, связанное с созданием на базе средств вычислительной техники средств обработки больших объемов данных и выработки на основе **моделирования** органов человека и/или заданных им **алгоритмов** решений определенных практических задач. Примерами использования искусственного интеллекта являются «**экспертные системы**», «**интеллектуальные системы**» и «**компьютерное зрение**» [4].

1.2. ИНФОРМАЦИОННЫЕ РЕСУРСЫ, ТЕОРИЯ ИНФОРМАЦИИ, ИНФОРМАТИКА

ИНФОРМАЦИОННЫЕ РЕСУРСЫ, ИР [information resources]

В общем случае под ИР понимается вся совокупность сведений, получаемых и накапливаемых в процессе развития науки и практической деятельности людей, для их многоцелевого использования в общественном производстве и управлении. ИР отображают естественные процессы и явления, зафиксированные в результате научных исследований и разработок или других видов целенаправленной деятельности в различного рода документах (например отчетах о НИР, патентах, проектно-конструкторской документации, массивах данных и т. п.), понятиях и суждениях, а также более сложных моделях действительности [4].

Данный термин начал широко использоваться в конце 1970-х — начале 1980-х гг. в результате осознания растущей зависимости промышленно развитых стран, отдельных организаций и фирм от источников информации (технической, политической, военной и т. д.), а также от уровня развития и использования средств передачи и переработки информации. С ним связаны термины: **национальные информационные ресурсы** (в том числе — **государственные и негосударственные информационные ресурсы**), **информационные ресурсы территориально-административных образований, фирм (организаций), их подразделений и т. п.**

В современном обществе ИР относятся к материальным и наиболее важным видам ресурсов, определяющих экономическую, политическую и/или военную мощь их владельца. В подтверждение этого тезиса можно привести ставший классическим пример с Японией: страна, практически

лишенная природных ресурсов и обладающая весьма скромными людскими ресурсами, является крупнейшим в мире производителем и экспортером не только изделий микроэлектроники, но и такой материалоемкой продукции, как автомобили и супертанкеры.

Отличием ИР от других материальных видов ресурсов (например полезных ископаемых) является их воспроизводимость. Как и другие виды ресурсов, ИР являются объектами импорта–экспорта, а также конкуренции, политической и экономической экспансии. Следует отметить, что границы понятия ИР в настоящее время четко не установлены. Так, некоторые ученые включают в его толкование также степень профессиональной подготовки общества или его части, способность воспроизводить и использовать ИР. Другие ограничивают ИР только совокупностью зафиксированных в документах и данных сведений, «представляющих ценность для учреждения (предприятия)» или, добавим, другого владельца ИР [5]. Заметим, что в последнем случае в понятие ИР не включены средства передачи и переработки информации.

Не будем спорить ни с теми, ни с другими авторами, хотя признаемся, что мы считаем предпочтительней более широкий подход к определению этого очень важного и интересного понятия. Для тех читателей, которые интересуются данной проблемой, рекомендуем монографию Г. Р. Громова [6].

Виртуальные (информационные) ресурсы [**virtual (information) resources**] — информационные ресурсы других организаций, предприятий, фирм и т. п., доступные пользователям в режиме **теледоступа** по каналам глобальной связи, например Интернета.

ТЕОРИЯ ИНФОРМАЦИИ [**information theory**]

Раздел **кибернетики**, изучающий общие стороны процессов передачи, хранения, извлечения и классификации **информации** различной природы (в том числе биологической, технической, социальной и др.) в независимости от ее семантического (смыслового) содержания. Общим средством анализа, описания и количественной оценки исследуемых процессов теории информации является ее математический аппарат, представляющий собой основу разрабатываемых и используемых теорий и методов.

Важнейшей частью теории информации является **теория передачи информации** [**theory of communication**], основоположником которой является американский математик **К. Шеннон**. Основные понятия этого раздела теории: **энтропия** (количественная мера неопределенности ситуации) произошло от греческого *entropia* — поворот, превращение и **количество информации**, измеряемое величиной изменения энтропии в условиях, связанных с получе-

нием информации. С использованием этих понятий выражается пропускная способность канала связи между источником информации и ее адресатом, равная максимально допустимой скорости передачи информации со сколь угодно малой вероятностью ошибки [4].

Составной частью теории информации является также **теория кодирования** (или — **теория оптимального кодирования**), рассматривающая вероятностные аспекты проблем кодирования и декодирования информации. Большой вклад в разработку теории информации внесли отечественные ученые **А. Н. Колмогоров, А. Я. Хинчин, Р. Л. Добрушин, В. А. Котельников, А. А. Харкевич** и др. Возросшая необходимость не только количественного, но и содержательного анализа информационных процессов породила появление новой науки — **информатики**.

ИНФОРМАТИКА [informatics, information science]

Упомянутый выше «Словарь по кибернетике» [4] содержит следующее определение: «Информатика — наука, изучающая информационные процессы и системы в социальной среде, — их роль, методы построения, механизм воздействия на человеческую практику, усиление этого воздействия с помощью вычислительной техники. Возникла как дополнение и конкретизация теории информации из потребностей автоматизации социально-коммуникативных процессов, и начала формироваться в 1970-е гг. как научная база использования электронных вычислительных машин в управлении, науке, проектировании, образовании, сфере услуг и т. д.». Как всякая относительно новая и быстро развивающаяся отрасль знания, не только связанная с социальной сферой, но и широко используемая в ней, информатика получила в последние годы множество толкований и не все они однозначны.

Наибольшие противоречия связаны с той частью данного понятия, которая определяет его семантические границы распространения. В качестве примера приведем другое определение: «Информатика — отрасль знания, изучающая закономерности сбора, преобразования, хранения, поиска и распространения документальной информации и определяющая оптимальную организацию информационной работы на базе современных технических средств» [5].

Видимые отличия цитируемых определений заключается, в частности, в том, что второе ограничивает понятие «информатика» технологическими процессами, входящими в функции информационных органов, а также документальной информацией. Следует отметить, что несмотря на давность этого определения (1971 г.), оно используется и в настоящее время в среде работников инфор-

мационных органов и служб, в недрах которых изначально и было порождено.

Еще один подход связан с организациями, подведомственными Комитету по информатизации России, который основное внимание акцентирует на инструментальных (программных и технических) средствах «информатики» и «информатизации»: «**Информатика [informatics, computer science]** — ... группа дисциплин, занимающихся различными аспектами применения и разработки ЭВМ: прикладная математика, **программирование, программное обеспечение, искусственный интеллект, архитектура ЭВМ, вычислительные сети**» [265].

Анализируя сказанное, мы склонны предпочесть вариант «кибернетиков» как более объективный и полный.

С целью более глубокого понимания нашими читателями указанного термина продолжим выборочное цитирование соответствующей статьи «Словаря по кибернетике»: «... Важнейшими категориями информатики являются понятия информационных сред (социальных подсистем, в которых осуществляются информационные процессы и куда внедряются ЭВМ как усилители человеческого интеллекта), полного информационного цикла (включающего зарождение информации, ее переработку, передачу, использование для снижения энтропии рассматриваемой социальной системы), полезной работы (отдачи) ЭВМ. *Отдача ЭВМ, коэффициент полезного действия (подчеркнуто мною — Ф. В.) зависят от уровня функционирования социальной среды, в которой они задействованы, ее упорядоченности, системности, условий для творческой деятельности людей, сложности и важности задач, решаемых с помощью машин.* Информатика не заменяет собой кибернетику, теорию информации, электронику, системотехнику, а взаимодействует с ними, имея ряд общих проблем. Интегральный характер информатики заключается также в ее взаимодействии с такими дисциплинами, как теория познания, семиотика, лингвистика, документалистика, библиотековедение».

1.3. ДОКУМЕНТЫ, ДОКУМЕНТАЦИЯ, НОСИТЕЛИ ИНФОРМАЦИИ

НОСИТЕЛЬ [media]

Обобщающее наименование материала, на который можно записывать данные. Носители подразделяются на **человекочитаемые (твердые)** и **машиночитаемые** (см. далее).

• **Машиночитаемый носитель** [**machine-readable media**] — носитель, пригодный для непосредственной записи и считывания данных программно-техническими средствами (ЭВМ). Термин обычно применяется к устройствам **внешней памяти ЭВМ**, например **магнитным и оптическим** дискам, **дискетам** и т. п. Однако он может использоваться и по отношению к определенной части **твердых носителей**, если они допускают использование специальных считывающих устройств, например **сканеров**. Подробнее о видах носителей см. раздел 3.4.2. «Память ЭВМ — общие термины», а об их характеристиках — раздел 3.5.1. «Внешняя память ЭВМ, накопители и связанные с ними термины».

• **Человекочитаемый (твердый) носитель** [**human readable media, hard media**] — носитель, пригодный или используемый для записи данных непосредственно считываемых человеком, например — бумага.

• **Микроформа, микроноситель** [**microform**] — общее наименование носителей, на которых тексты или графические изображения представлены в уменьшенном фотографическом способом виде. Микроформы широко используются в библиотеках для хранения в человекочитаемом виде (чтение их производится с использованием специальных читальных аппаратов) больших объемов копий документов и данных. Создание микроформ производится при помощи специальной (микрофильмирующей) фотоаппаратуры, а также компьютерной записи (см. также «**Computer Out on Microfilm/microfiche**» — «**COM**»). Основными типами микроносителей являются:

1. **Микрофильм** [**microfilm**] — пленка обычного фотоаппарата содержащая уменьшенные копии текстов и графических изображений;

2. **Микрофиша** [**microfiche**] — стандартный (105 × 148 мм) прямоугольный лист фотопленки, на котором располагается 420 кадров уменьшенных изображений страниц текста или графики.

ДОКУМЕНТ

[**document** от лат. *documentum* — свидетельство, доказательство]

Материальный носитель информации, зафиксированной вне памяти человека. В соответствии с ГОСТ 16487-70 [8] «документ» является средством «закрепления различным способом на специальном материале сведений о фактах, событиях явлениях объективной действительности и мыслительной деятельности человека». Документы могут содержать текстовую, цифровую, графическую и аудиоинформацию. Они могут подвергаться процессам записи (преобразования), хранения, поиска, передачи, получения, сбора и

чтения. В зависимости от характера обработки содержащихся в них данных документы принято разделять на **первичные** и **вторичные**.

- **Первичные документы** [**primary document, source**] — документы, содержащие исходную запись сведений, полученных в процессе исследований, разработок, наблюдений, анализа или других видов человеческой деятельности независимо от ее характера или тематики. Деление документов на первичные и вторичные в значительной степени условно, поскольку один и тот же документ может содержать сведения, относящиеся и к первой, и ко второй группе. Так, считающиеся безусловно первичными видами документов — отчеты по НИР, проектные документы, монографии и т. д., могут содержать наряду с оригинальными сведения или данные ранее включенные в другие первичные, или даже вторичные документы.

- **Вторичные документы** [**secondary document**] — документы, полученные в результате аналитико-синтетической и логической переработки сведений или данных, содержащихся в первичных документах. Примерами вторичных документов являются справочные и энциклопедические издания, рефераты и реферативные издания, библиографические издания, указатели и списки, обзоры (за исключением аналитических обзоров) и т. п.

Юридический документ — документ, оформленный в соответствии с действующим юридическим законодательством и имеющий правовое значение.

В зависимости от характера связи документов с технологическими процессами в автоматизированных системах и вида физического носителя информации различают:

- **машинно-ориентированный документ**, предназначенный для обработки части содержащейся в нем информации средствами вычислительной техники [9]. Примерами машинно-ориентированных документов могут служить заполненные специальные формы бланков библиографической записи, различного рода анкет и т. п., предназначенные для последующего считывания в ЭВМ записанных в них данных с использованием клавиатурных операций;

- **машиночитаемый документ** [**machine-readable document**], документ, пригодный для автоматического считывания содержащейся в нем информации [9]. Средства автоматического считывания — **сканеры** предъявляют определенные требования к характеру оформления соответствующих текстовых, графических и других видов записей, включая их **формат**, виды шрифтов, наличие специальных служебных знаков, контраст и т. п.;

- **документ на машиночитаемом носителе, электронный документ**, созданный средствами вычислительной техники, записанный на **машиночитаемый носитель**: магнитную ленту (МЛ), маг-

нитный диск (МД), дискету, оптический диск и т. п. и оформленный в установленном порядке¹⁾. В зависимости от того, на каком носителе записан машиночитаемый (электронный) документ или документы, принято указывать его вид, например, «документ(ы) на магнитном (оптическом) диске», «документ(ы) на магнитной ленте»;

- **документ-машинограмма, распечатка [hard copy document]** — документ на бумажном носителе, созданный средствами вычислительной техники и оформленный в установленном порядке;

- **документ на экране дисплея**, документ, созданный средствами вычислительной техники, отображенный на экране дисплея (монитора) и оформленный в установленном порядке;

- **электронный документ; электронный текст [electronic document]** — совокупность данных в памяти вычислительной системы, предназначенная для восприятия человеком с помощью соответствующих программных и аппаратных средств;

- **title — электронный документ**, который идентифицируется и передается как единое целое;

- **служебный документ [internal document]**, содержание которого отражает способ или результаты решения какой-либо функциональной задачи автоматизированной системы;

- **входной документ [input document]**, составленный по определенной форме и содержащий данные, предназначенные для ввода в ЭВМ. Входные документы могут быть условно разделены на две категории: **информационные документы и запросы**;

- **информационный документ [information document]**, основное назначение которого — пополнение массивов или баз данных (БД) ЭВМ;

- **выходной документ [output document]**, являющийся носителем результатов обработки данных ЭВМ и/или формируемый автоматизированной системой и выданный системными средствами вывода.

- **электронное издание, ЭИ** — самостоятельный законченный продукт, содержащий информацию, представленную в машиночитаемой (электронной) форме, все копии (экземпляры) которого соответствуют оригиналу, и предназначенный для длительного хранения, широкого распространения и многократного использования неопределенным кругом пользователей [687].

¹⁾ *Примечание:* Фраза «оформленный в установленном порядке» в цитируемом выше ГОСТе [9] относится к форме документа, а не к его юридическому статусу. Однако при выполнении определенных условий «оформления» эти дефиниции можно распространить и на юридические документы.

ДОКУМЕНТАЦИЯ

[documentation, collection of documents, file of documents]

Совокупность документов, объединенных по определенным признакам (например, по назначению, содержанию и т. п.) и оформленных по единым правилам.

Применительно к программным продуктам и средствам вычислительной техники термином «документация» обозначают:

1. Руководства по использованию.

2. Совокупность текстов, описывающих строение и применение соответствующих средств или изделий.

Документация предназначена для облегчения использования программных и технических средств и включает руководства, справочники, учебники, краткие справочники, обучающие программы, а также средства диалоговой документации и подсказки [9].

СИСТЕМА ДОКУМЕНТАЦИИ [documentation system]

Совокупность документов, состав, содержание, структура и правила оформления которых определены Государственными стандартами.

Очень часто понятие «система документации» подменяется другим — **вид документации**, который строго говоря не является его синонимом. Так, различают следующие виды документации: научно-техническая (научная и техническая), директивная, юридическая, нормативная (нормативно-техническая), технологическая, проектно-конструкторская (проектная и конструкторская), эксплуатационная, технико-экономическая и др. Данная классификация достаточно условна в том смысле, что состав и правила оформления большинства используемых в реальной практике видов документации, нормативно не определены (см. ГОСТ 6.10-2-83 [9]).

ДОКУМЕНТАЛИСТИКА

Направление в кибернетике, занимающееся изучением и оптимизацией документальных систем независимо от их назначения (в первую очередь — документов).

1.4. СРЕДСТВА ИНФОРМАЦИОННОГО ОБЕСПЕЧЕНИЯ АВТОМАТИЗИРОВАННЫХ СИСТЕМ И ИХ СТРУКТУРА

1.4.1. Информационные элементы и их виды

ИНФОРМАЦИОННЫЙ ЭЛЕМЕНТ, ИЭ [information element (item)]

Единица информации, подлежащей обработке, хранению и передаче пользователям системы или предназначенная для обеспечения ее работы.

ИЭ является обобщенным наименованием структурной единицы информации, не зависящей от ее назначения, состава данных, характера или материала носителя и т. п. Например, ИЭ может быть названо содержание книги, статьи, библиографического описания или его части. С учетом сказанного ИЭ могут быть делимыми на части или не делимыми.

Виды информационных элементов

Составной ИЭ — информационный элемент, который может быть разделен на части средствами системы (без участия человека в выполнении логических или интеллектуальных операций, связанных с анализом и разделением содержащихся в этом информационном элементе данных), причем каждая из его частей также является информационным элементом системы.

Исходя из данного определения, составной ИЭ характеризуется видом и организацией содержащихся в нем данных. Так, неструктурированные вторичные документы — библиографическое описание, реферат и аннотация — не могут считаться составными ИЭ, если **автоматизированная система** не позволяет на физическом уровне своими средствами выделить содержащиеся в них различные виды данных (например, в библиографическом описании — сведения об ответственности, фамилии авторов, продолжение названия и т. п.) или аспекты описания.

В общем случае ИЭ, которые не могут быть разделены на части автоматизированными средствами, принято называть **данными**.

В указанном контексте уместно упомянуть еще одно существующее определение понятия «данные»: это «информация, представленная в виде пригодном для обработки автоматическими средствами при возможном участии человека» [265].

Элемент данных, ЭД [data item] — неделимый информационный элемент, ЭД являются минимальной структурной единицей информации, поскольку части ЭД теряют признаки информацион-

ного элемента. Примерами ЭД могут служить в библиографическом описании: фамилия автора, год издания, название издания и т. п. Деление на части **лексических единиц**, образующих соответствующие понятия, не приводит к образованию новых ИЭ.

Из сказанного следует, что вид элемента данных определяется как характером содержащихся в нем сведений, так и особенностью его организации или записи (см. далее статьи «**Логическая запись**» и «**Физическая запись**»).

В зависимости от характера содержащихся в ИЭ данных они могут подразделяться на следующие виды:

- **документографический (документальный) ИЭ** — информационный элемент, содержащий сведения о документах (например, библиографическое описание, реферат, аннотация, поисковый образ документа);

- **полнотекстовый ИЭ** — информационный элемент, содержащий полные тексты документов или их частей;

- **фактографический ИЭ** — информационный элемент, содержащий описание отдельных фактов или некоторой ограниченной совокупности фактов, объединенных по каким-либо признакам (например, сведения о курсе валют, расписания движения транспорта и т. п.);

- **объектографический ИЭ** — информационный элемент, содержащий структурированное и логически связанное описание различных объектов (см. начало словаря) науки, техники или управления. Объектографические информационные элементы являются более сложным вариантом **фактографических ИЭ**, поэтому многие специалисты и организации, работающие в области информатики, используют в рассматриваемых случаях термины — фактографические ИЭ или его синонимы.

Развитие вычислительной техники может привести к появлению терминов, расширяющих приведенный ряд, например **графический ИЭ, аудио ИЭ** и т. п., хотя, следует признать, что в литературе мы с такими терминами или их синонимами пока не встречались.

ТИП ДАННЫХ [data type]

Множество допустимых значений данных, объединенных общим содержанием и именем (например «Библиографическое описание», «Автор», «Год издания», «Стоимость» и т. п.), а также совокупностью допустимых операций, которые можно выполнять над этими данными, включая способ их хранения в памяти ЭВМ. Понятие «тип данных» делает манипулирование данными с использованием средств вычислительной техники абстрактным процессом

и скрывает лежащее в основе обращения с ними представление их в виде двоичного кода [35].

Виды типов данных:

- **аналоговые данные** [analog data] — данные, принимающие произвольные значения из заданного диапазона, и представляемые в виде непрерывно изменяющихся физических величин, например напряжения, длительности сигнала и т. п.;

- **дискретные (цифровые) данные** [digital data], представленные в дискретном коде в определенной, например, двоичной системе счисления;

- **аналого-цифровые данные** [analog-digital data] — аналоговые данные, преобразуемые для обработки в цифровой код;

- **двоичные данные** [binary data], представленные в двоичном коде;

- **десятичные данные** [decimal data], представленные в десятичном коде;

- **алфавитно-цифровые (текстовые) данные** [alphanumeric data], значения которых составлены из любых знаков алфавита;

- **числовые (арифметические) данные** [arithmetic data], над которыми можно выполнять арифметические операции.

МНОЖЕСТВО [set]

Совокупность каких-либо **объектов**, представляемых как единое целое. Множество может включать в себя как однородные объекты, объединенные каким-либо общим признаком, так и неоднородные. Объединение последних в конкретное «множество» определяется на основе ассоциативных и других связей между его элементами (см. далее «Теория множеств»).

Некоторые виды множеств:

- **конечное множество** [finite set], содержащее определенное (конечное) число элементов;

- **нечеткое множество** [fuzzy set], принадлежность объекта к которому определяется функцией, принимающей значения в области $[0, 1]$;

- **ограниченное множество** — упорядоченное множество, имеющее нижнюю и верхнюю границы;

- **пустое множество** [empty set], не имеющее ни одного элемента;

- **подмножество** [subset] — множество, являющееся частью другого множества;

- **эквивалентное множество** [equivalent set] — множество, имеющее взаимно однозначное соответствие с другим множеством,

при котором каждому элементу одного из них соответствует один единственный элемент другого и наоборот.

ТЕОРИЯ МНОЖЕСТВ [set theory]

Раздел математики, исследующий общие свойства **множеств** и операций над ними. Понятия и методы теории множеств широко используются в **информатике**, в частности, в теории и практике построения **информационно-поисковых систем**¹⁾.

1.4.2. Запись, файл, массив, ключ

ЗАПИСЬ [record, writing]

1. Единица обмена данными между программой и внешней памятью ЭВМ (**record**).

2. Процесс или результат закрепления (фиксирования) данных на носителе информации (**writing**). Запись может содержать один **информационный элемент**, включая и составной.

Логическая запись [logical record] — совокупность записей взаимосвязанных ИЗ (в том числе элементов данных, данных и составных ИЭ), рассматриваемая в логическом плане как единое целое. Одна логическая запись может состоять из нескольких физических или быть частью одной физической записи.

Физическая запись [physical record] — порция данных, пересылаемых как единое целое между основной и внешней памятью ЭВМ.

На машиночитаемом носителе (например, магнитном или оптическом диске, дискете и т. п.) физическая запись реализуется в форме «**Поля данных**», «**Файла**» (см. далее) и др. Физическая запись может содержать одну логическую запись, ее часть или несколько логических записей (см. также «**Блок данных**»).

ТИП ЗАПИСИ [record type]

1. Обобщенное имя записи типа данных.

2. В базах данных: тип, к которому относится данная запись²⁾.

¹⁾ Подробнее см.: Словарь по кибернетике. 2-е изд. — Киев: Гл. редакция Украинской Советской энциклопедии им. М. П. Бажова, 1989. — С. 358, 422.

²⁾ Подробнее о типах записей в базах данных см.: Когаловский М. Р. Энциклопедия технологий баз данных [722].

3. Характеристика, определяющая возможность записи менять свою длину (по этому признаку различают записи: **фиксированной, переменной и неопределенной длины**). См. также «**Формат записи**» и «**Переменный формат**».

БЛОК ДАННЫХ, БЛОК [data block, block]

1. Несколько последовательных логических записей, объединенных в одну физическую.

2. Единица доступа к внешней памяти ЭВМ.

3. Выделенный фрагмент текстового материала, который можно удалить, переместить или выполнить над ним преобразования.

4. В *вычислительных сетях*: последовательность передаваемых данных, рассматриваемая как единое целое. Обычно имеет **адрес** и **контрольную сумму**, позволяющую обнаруживать ошибки, и другие служебные элементы, требуемые соответствующим **протоколом** канала связи.

ПАКЕТ [packet]

Единица информации в сети передачи данных. Пакет имеет строго определенную структуру. Она предусматривает наличие заголовка, который содержит адреса получателя и отправителя, данные для контроля ошибок, а также самого сообщения или его части, поскольку передаваемое сообщение (например, текстовый файл) может быть разделено и пересылаться в виде последовательности пакетов (см. также «**Фрейм**»).

Формат пакета [packet format] — в *вычислительных сетях*: формат, определяющий размеры, размещение и содержание полей, входящих в состав пакета. Формат пакета регламентируется **протоколом** управления соответствующей сети передачи данных.

ФАЙЛ [file]

1. Идентифицированная последовательность или множество записей однотипных **информационных элементов**.

2. Поименованная целостная совокупность данных на машиночитаемом носителе.

3. Поименованная область **внешней памяти** ЭВМ.

4. В *английском языке*: картотека, архив, комплект, подшивка и т. п. (см. также «**Массив**»).

Файл — основной структурный элемент хранения данных в ЭВМ, обеспечивающий возможность машине и человеку отличать один набор данных от другого при их поиске, изменении, удалении или выполнении с ними других операций.

Содержанием файла может стать одна поименованная логическая запись или несколько логических записей, объединенных общим именем. Например, файл может содержать записи некоторого множества библиографических описаний, каждое из которых по своей сути является логической записью, объединенной с другими каким-либо общим признаком (вид документа, тематика и т. п.). В других случаях содержанием записи в файле могут быть отдельные документы, управляющие работой ПЭВМ, программы или их части. «Идентификация» или «поименование» указанных записей или групп записей выражается присвоением каждому из них «уникального» или отличного от других имени, позволяющего автоматизированным средствам находить их в массивах других записей, в том числе файлов.

На физическом уровне содержимое файла может быть не структурировано и представлять собой единственную физическую запись или структурировано, например **полями данных**. В последнем случае оно будет включать соответствующее множество физических записей.

Некоторые типы файлов:

- **файл данных [data file]** — файл, содержащий **данные** в отличие от «программных» файлов, содержащих записи программ или их частей.

- **файл регистрации [log file]** — файл, в котором хранятся записи о других файлах, например, если какие-либо файлы были архивированы (см. «**Архивация**»), список имен этих файлов может храниться в файле регистрации.

- **файл с произвольным доступом [random-access file]** — файл, в котором любая запись может быть считана, записана или изменена без необходимости считывания других записей. Произвольный доступ должен быть обеспечен **операционной системой** и реализован специальной машинной программой.

- **основной файл [master file]** — файл, являющийся основным источником данных для решения определенного класса задач, для какой-либо цели или назначения. Поддержка этого файла в актуальном состоянии, при котором содержащиеся в нем данные отвечают требованиям новизны и точности, производится с использованием операций **обновления файла**.

- **упорядоченный файл [sequential file]** — файл, в котором записи упорядочены по ключевому полю для ускорения доступа к определенной записи.

- **последовательный файл [serial file]** — файл, в котором записи не упорядочены; поэтому, чтобы прочитать нужную запись, требуется прочитать все предыдущие.

• **скрытый файл [hidden file]** — файл, который не отображается на экране **монитора** при просмотре каталога файлов (директории). Скрытым можно сделать любой файл, с целью затруднения его нахождения, чтения и/или порчи другими лицами, которые могут иметь **доступ** к ЭВМ и др. [27, 265, 369, 722].

МАССИВ [array]

1. Упорядоченная структура множества документов или данных одного типа.

2. Поименованная совокупность однотипных (логически однородных), упорядоченных по индексам записей ИЭ.

3. Упорядоченное множество элементов одного типа. Каждый элемент массива должен иметь имя (идентификатор, индекс), обеспечивающее возможность его нахождения. Элементы массива могут быть одномерными и многомерными.

Применительно ко многим задачам автоматизированной обработки данных термины **массив** и **файл** могут использоваться как синонимы. Однако соответствующие понятия имеют и отличия: так, понятие **массив** не обязательно связано с записями информационных элементов (включая полнотекстовые документы) на машиночитаемых носителях. Понятие же **файл** в русскоязычной практике, как правило, предполагает указанное условие. О видах массивов см. [722].

Информационно-поисковый массив, поисковый массив [information collection, file] — массив документов или данных (соответствующие англ. эквиваленты — **document collection, data collection**), в котором производится **информационный поиск** [27, 265, 369].

АРХИВ [archives]

Организованная совокупность **массивов** данных или программ, длительно хранимых на внешних машиночитаемых носителях (например на **гибких магнитных дисках, магнитных лентах** и/или **CD-ROM**) с целью обеспечения возможности их дальнейшего использования.

Целями создания архивов являются:

1) создание страховочных копий информационных и программных продуктов на случай их утраты или порчи в ходе эксплуатации вычислительных средств;

2) освобождение внешней памяти ЭВМ (например **накопителя на жестком магнитном диске**) от программ и данных, потребность в оперативном использовании которых частично, полностью или временно отпала.

КЛЮЧ [key]

1. **Информационный элемент**, однозначно идентифицирующий **запись** или указывающий ее местоположение. Ключ может также служить средством для идентификации некоторого множества (в том числе массива) записей и располагаться в одном или нескольких из его **полей**.

В качестве ключа может служить **элемент данных** (в том числе **индекс**, **адрес**, **код** и др.) или группа элементов данных. Так, в **индексно-последовательном файле** ключ является обязательным элементом **записи** в каждом поле данных. Набор значений некоторой совокупности **атрибутов в реляционной модели данных** служит ключом, который однозначно идентифицирует **кортеж** или группу кортежей конкретного **отношения**.

2. Параметр шифрования, представляющий один из возможных вариантов шифра. Для дешифрирования помимо ключа необходимо знать последовательность операций или правило его использования (см. также «**Алгоритм**»).

3. Значение, используемое для подтверждения полномочий на **доступ** к некоторой информации (в том числе к базам данных, отдельным файлам и т. п.), программным и/или техническим средствам. См. также значение «**Ключ**» в разделе 1.6.2.

1.4.3. База данных, банк данных

БАЗА ДАННЫХ, БД [Data Base, database, DB]

Организованная совокупность блоков информационных элементов, представленных на машиночитаемых носителях, предназначенных и пригодных для оперативного решения пользовательских, служебных и других задач с использованием средств вычислительной техники.

Существует великое множество дефиниций, посвященных данному понятию. Например:

- «Организованная в соответствии с определенными правилами и поддерживаемая в памяти компьютера совокупность данных, характеризующих актуальное состояние некоторой предметной области и используемая для удовлетворения информационных потребностей пользователей» [722];

- «Совокупность данных, организованных по определенным правилам, предусматривающим общие принципы описания, хранения и манипулирования данными, независимая от прикладных программ. Является информационной моделью предметной обла-

сти. Обращение к БД осуществляется с помощью системы управления базами данных» (см. «СУБД»);

- «Совокупность взаимосвязанных данных, используемых несколькими приложениями под управлением системы управления базой данных» [27];

- «Совокупность предназначенных для машинной обработки данных, которая служит для удовлетворения нужд многих пользователей в рамках одной или нескольких организаций» [28];

- «Объективная форма представления и организации совокупности данных (статей и т. д.), систематизированных таким образом, чтобы эти данные могли быть найдены и обработаны ЭВМ» [159].

При всем многообразии определений понятия БД всегда предполагается, что БД отвечает следующим признакам.

1. БД содержит некоторое множество данных необходимых (и, желательно, достаточных) для решения конкретных задач многих пользователей (в том числе как реальных, так и потенциальных) или, что в принципе почти одно и то же, — удовлетворения соответствующих информационных потребностей.

2. Данные или информационные элементы в БД определенным образом структурированы и связаны между собой (т. е. организованы), при этом структура, состав данных и их содержание в БД не зависят от особенностей прикладных программ, используемых для управления БД.

3. Данные (информационные элементы) представлены на машиночитаемых носителях в форме пригодной для оперативного использования их с применением средств вычислительной техники, включая и систем управления базами данных (СУБД).

Совокупность данных или **информационных элементов**, не отвечающих любому из трех перечисленных выше признаков, не может называться базой данных!

Последнее замечание является реакцией на весьма распространенную практику многих работников отечественных библиотек и информационных органов — использовать указанный термин по отношению к **массивам** документов и данных, базами данных не являющимися. Это относится, в частности:

- к отдельным **файлам**, а также совокупности нескольких файлов, содержащих данные, состав и организация которых не обеспечивает решение каких-либо **задач** пользователей или удовлетворение определенных информационных потребностей;

- к любому набору документов и данных на **твердых** или **человекочитаемых носителях**;

- к **архивам** документов и данных на машиночитаемых носителях.

Использование баз данных характеризуется следующими свойствами.

1. **ОПЕРАТИВНОСТЬ**: средства вычислительной техники позволяют осуществлять оперативный **доступ** к информации.

2. **ПОЛНАЯ ДОСТУПНОСТЬ**: вся информация, содержащаяся в БД, доступна для использования.

3. **ГИБКОСТЬ**: имеется возможность легко изменять состав и форму выдачи интересующих пользователя данных, изменения в БД вносятся также достаточно просто.

4. **ЦЕЛОСТНОСТЬ** (данных): минимизируется дублирование данных, предоставляется возможность упорядочения и согласованности данных, а также работ по их обновлению (см. далее «Целостность БД»).

С общими характеристиками БД связан также ряд взаимозависимых понятий.

- **Целостность БД [database integrity]** — состояние БД, при котором все значения данных правильно отражают предметную область (в пределах заданных ограничений по точности и согласованности во времени) и подчиняются правилам взаимной непротиворечивости. Поддержание целостности БД предполагает ее проверку и восстановление или корректировку из любого неправильного состояния, которое может быть обнаружено. Это входит в функции **администратора БД**, который пользуется средствами системы управления БД (см. «СУБД»). Аналогично можно говорить и о **целостности файла**, хотя в типичных случаях файлы подвергаются менее обширным проверкам на целостность.

- **Защищенность БД [database security]** — наличие и характеристика средств (аппаратных, программных, организационных, технологических, юридических и т. п.) обеспечивающих предотвращение или исключение:

- а) доступа к информации лиц, не получивших на то соответствующего разрешения (см. также «**Несанкционированный доступ**»);

- б) умышленного или непредумышленного разрушения или изменения данных.

- **Безопасность БД [database safety]** — свойство БД, которое заключается в том, что содержащиеся в ней данные не причиняют вреда пользователю при правильном их применении для решения любых **функциональных задач** системы, для которой она была создана. Часто понятия безопасность и защита БД рассматриваются как синонимичные, например в [265]. См. также «**Информационная безопасность**».

- **Эффективность БД [database efficiency]**

1. Степень соответствия результатов использования БД затратам на ее создание и поддержание в рабочем состоянии, в случае оценки этого показателя в денежном выражении он носит наименование **экономической эффективности БД**.

2. Обобщающий показатель качества состояния и использования БД по совокупности признаков (в том числе скорость, доступность, гибкость, целостность, защищенность, безопасность и др.).

- **Техническая эффективность БД** — эффективность БД применительно к условиям их использования в конкретных автоматизированных системах.

БАНК ДАННЫХ (БнД) [databank]

Совокупность одной или нескольких **баз данных** и средств управления (**манипулирования**) **данными**. БнД является важнейшей составной частью автоматизированных информационных и других систем, функционирование которых связано с обработкой больших объемов данных (в том числе АСУ, АСУП, АСУТП и др.).

Разные авторы и документальные источники (в том числе нормативные) неоднозначно трактуют данное понятие. В одних случаях оно определяется, как совокупность баз данных и **СУБД** (см. далее), в других — в него включают также технические и организационные средства. В указанном плане уместно привести одно из самых расширенных представлений указанного термина, определяемого в «Толковом словаре по информатике» В. И. Першикова и В. М. Савинкова [265] следующим образом: «**БнД — Автоматизированная информационная система** централизованного хранения и коллективного использования данных. В состав БнД входят одна или несколько баз данных, справочник баз данных, система управления базами данных, а также библиотеки запросов и **прикладных программ**».

Учитывая изложенное, мы не считаем необходимым жестко определять состав средств, характеризующих данное понятие. Заметим только, что является грубой, хотя и весьма распространенной, ошибкой использовать термины **Банк данных** и **база данных** как синонимичные.

Распределенный банк данных [distributed databank] — система территориально разобобщенных **банков данных**, объединенных средствами вычислительной сети и функциональным управлением. Основу такой системы составляют **распределенная база данных** и система управления распределенными базами данных (см. «СУРБД»).

Банк знаний [knowledge bank] — см. «Экспертная система».

1.4.4. Структуры, модели данных и связанные с ними термины

СТРУКТУРА [structure]

Фиксированное упорядоченное множество **объектов** и связей между ними.

С понятием «структура» связаны следующие термины:

- **структура данных [data structure]** — множество элементов данных, объединенных и упорядоченных определенным образом;

- **структура информационной базы [information support structure]** — упорядоченная по определенным правилам совокупность подмножеств записей **информационных элементов**, образующих **информационную базу**, и необходимых для реализации функций автоматизированной системы;

- **структура базы данных [DB structure]** — принцип или порядок организации записей в **базе данных** и связей между ними. Структуру БД принято рассматривать на разных уровнях **абстракции** (представления) и, в частности 1) концептуальном (с позиции администратора предприятия), 2) реализации или внешнем (с позиций конечного пользователя и прикладного программиста), 3) физическом или внутреннем (с позиций системного аналитика и системного программиста). Соответственно этим уровням различают концептуальную, внешнюю и физическую модели и/или схемы организации данных;

- **абстракция [abstraction]** — использование для описания или представления общих свойств объекта без конкретной его реализации (например **типов объектов** — «читатель», «фирма», «автомобиль», но не конкретно названных читателей, фирм, марок автомобилей и т. п.);

- **абстрактная структура данных [abstract logic design]** — структура данных, определенная функционально посредством выполняемых на ней операций. Такая структура не связана с определенными типами объектов.

СПИСОК [list]

1. Структура данных, представляющая собой логически связанную последовательности записей — элементов списка.

2. Перечень различных объектов.

Связанные со «структурой» и «списком» понятия:

- **цепной список, связанный список [chained (linked) list]** — список данных, в котором порядок элементов списка задан посредством указателей, включенных в их запись;