

ISSN 1816-0301 (Print)
ISSN 2617-6963 (Online)

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ
INFORMATION TECHNOLOGY

УДК 20.15.05
<https://doi.org/10.37661/1816-0301-2020-17-3-87-94>

Поступила в редакцию 07.02.2019
Received 07.02.2019

Принята к публикации 20.05.2020
Accepted 20.05.2020

О количественных характеристиках информационных ресурсов

Л. И. Точицкий

*Объединенный институт проблем информатики
Национальной академии наук Беларуси, Минск, Беларусь
E-mail: Leonard.Tochicki@tut.by*

Аннотация. Предлагаются количественные характеристики информационных ресурсов автоматизированных информационных систем (АИС). Рассматриваются сущности, представляемые в АИС отображением информационного объекта, и их составные части (информационная единица, информационный элемент). Обосновывается состав информационного пространства такой системы, которое представляет собой всю обрабатываемую в ней информацию, вспомогательные информационные средства информационных ресурсов и сведения о системе и ее пользователях. Применение количественных значений к характеристикам информационных ресурсов создаваемых систем, учитывающих смысловые свойства отображаемых в АИС информационных объектов, позволит повысить качество проектных решений по АИС и их эксплуатационные показатели на стадиях создания и эксплуатации в течение жизненного цикла.

Ключевые слова: информационный ресурс, информационный объект, отображение информационного объекта, информационная единица, информационный элемент, информационное пространство системы

Для цитирования. Точицкий Л. И. О количественных характеристиках информационных ресурсов / Л. И. Точицкий // Информатика. – 2020. – Т. 17, № 3. – С. 87–94. <https://doi.org/10.37661/1816-0301-2020-17-3-87-94>

On quantitative characteristics of information resources

Leonard I. Tochicki

*The United Institute of Informatics Problems of the National Academy
of Sciences of Belarus, Minsk, Belarus
E-mail: Leonard.Tochicki@tut.by*

Abstract. The quantitative characteristics of information resources for automated information systems (AIS) are proposed in the paper. The entities represented in the AIS by the information object mappings and their components (information unit and information element) are considered. The composition of the system information space is substantiated, which represents the whole amount of information being processed in the system, auxiliary information resources and all data of the system and its users. Application of the quantitative values to AIS information resources characteristics, taking into account the sense properties of information objects being mapped in AIS, will make it possible to raise the design decision making quality of AIS and its performance within the lifecycle.

Keywords: information resource, information object, representation of information object, information unit, information element, information space of the system

For citation. Tochicki L. I. On quantitative characteristics of information resources. *Informatics*, 2020, vol. 17, no. 3, pp. 87–94 (in Russian). <https://doi.org/10.37661/1816-0301-2020-17-3-87-94>

Введение. Оценка эффективности функционирования АИС сложна, так как количество характеристик (параметров), определяющих качество ее функционирования, весьма велико. Показатели эффективности функционирования АИС в значительной степени зависят от характеристик ее информационных ресурсов (ИР). В настоящее время используется так называемая объемная система характеристик ИР АИС – объем памяти, занятой и зарезервированной для ИР на серверах АИС.

В статье рассматриваются основные положения о количественных характеристиках ИР, применение которых позволит повысить обоснованность проектных решений по их конфигурации в течение жизненного цикла АИС.

Парадоксальным свойством ИР, которые представляют собой организованную совокупность документированной информации, включающей базы данных, или совокупность взаимосвязанной информации в информационных системах [1], является присущее им увеличение объемов на протяжении всего времени использования в отличие от природных, материальных, трудовых, финансовых и других видов ресурсов, объемы которых уменьшаются по мере их потребления.

В течение жизненного цикла информации некоторая ее часть теряет смысл и становится неактуальной. Однако есть определенная часть информации, которая содержит документированные результаты научно-технической, общественно-политической, исторической и культурной деятельности. Поэтому жизненный цикл этой информации не имеет конечного срока хранения и, соответственно, ИР, где размещаются такие сведения, должны функционировать постоянно.

Общеизвестные виды размещаемой на ИР информации:

- текстовая (буквенная; числовая, отображающая количественную меру сущностей) и символьная специальная (символы математические, технические и т. п.);
- графическая (картины (рисунки), графики, чертежи, схемы и графические изображения);
- звуковая (речевая, музыкальная, сигнальная);
- видеoinформация (мультимедийная);
- комбинированная (объединяющая несколько видов информации).

Применение количественных значений к характеристикам ИР. Характеристики ИР определяются назначением АИС и алгоритмами выполняемых на ее ресурсах процессов. На стадии эксплуатации назначение АИС может изменяться (могут расширяться виды работ и т. п.), соответственно, будут меняться и требования к ИР.

Для описания объектов ИР предлагается использовать следующие понятия.

Информационный объект (ИО) – сущность внешней среды, рассматриваемая и используемая в информационной системе. Внешней средой системы (ИР) являются объекты, не принадлежащие ей, но оказывающие на нее влияние и воспринимающие от нее воздействия.

В ГОСТ 33707-2016 дается и другое определение: «Информационный объект: Совокупность данных и программного кода, обладающая свойствами (атрибутами) и методами, позволяющими определенным образом обрабатывать данные. Самостоятельная единица применения и хранения в интеллектуальной информационной системе (ИИС)» (URL: <http://docs.cntd.ru/document/1200139532>). Это определение является частным, так как не всегда в ИО имеются в наличии методы, позволяющие обусловленным образом обрабатывать данные.

Содержанием ИР являются отображения ИО ($O_{ИО}$). В ИР хранятся первичные и реплицированные (вторичные) $O_{ИО}$.

Количество первичных $O_{ИО}$ ($O_{ИО\text{ перв}}$), размещаемых на ИР АИС и занесенных в реестр ИР, определяется выражением

$$| O_{ИО\text{ перв. АИС}} | = N, \quad (1)$$

где $| O_{ИО\text{ перв. АИС}} |$ – количество элементов конечного множества $O_{ИО\text{ перв. АИС}}$; N – количество $O_{ИО\text{ перв}}$, представленных в ИР.

Количество вторичных $O_{ИО}$ ($O_{ИО\text{ вт}}$), которые являются также составной частью ИР АИС, определяется стратегией управления конфигурацией ИР, точками ввода запросов информации пользователями и ее вывода и местами размещения используемых приложений.

Информационная единица (ИЕ) – запись, непосредственно представляющая описание ИО или его части и содержащая логически законченную информацию.

$O_{ИО}$ могут быть простыми, содержащими одну ИЕ, и составными, содержащими две и более ИЕ. Состав ИР АИС с помощью термина «информационная единица» $O_{ИО}$ предлагается определять как сумму первичных и реплицированных ИЕ. $O_{ИО}$ представлено в ИР совокупностью ИЕ, которая устанавливает различные аспекты отображаемого ИО:

$$\sum_1^n ИЕ_{O_{ИО}} \subseteq O_{ИО}, \quad (2)$$

где n – количество $ИЕ_{O_{ИО}}$ в $O_{ИО}$, \subseteq – функция принадлежности подмножества (объединения элементов в одно целое).

Состав ИР ($C_{ИР\ АИС}$), образованный первичными ИЕ, представим выражением

$$C_{ИР\ АИС} = \sum_1^n ИЕ_{O_{ИО1}} + \sum_1^n ИЕ_{O_{ИОi}} + \sum_1^n ИЕ_{O_{ИОn}}, \quad (3)$$

где $ИЕ_{O_{ИО}}$ – ИЕ соответствующего $O_{ИО}$, n – количество ИЕ в соответствующем $O_{ИО}$.

ИЕ в ИР предлагается определять следующими свойствами: идентификатором, параметрами связанности с другими ИЕ, семантическими характеристиками, правилами доступа и изменения.

Идентификатор ИЕ – уникальная последовательность символов (знаков), присваиваемая ИЕ, которая формируется по обусловленным правилам. Он является одной из информационно-поисковых характеристик ИЕ и однозначно определяет ее в системе. Параметры связанности ИЕ с другими единицами устанавливают виды ее связей (представление составной части ИО, общей темы, общей первопричины формирования и т. д.). Семантическая характеристика описывает смысловое содержание ИЕ. Правила доступа и изменения обозначают права пользователя на доступ к ИЕ, ее изменения, а также порядок работы с ИЕ. В свою очередь, ИЕ может содержать один или более информационных элементов.

Информационный элемент (ИЭ) – единица информации, подлежащей обработке, хранению и предоставлению пользователю ИР [2]:

$$ИЕ_{O_{ИО}} = \sum_1^n ИЭ_{ИЕ}, \quad (4)$$

где $ИЭ_{ИЕ}$ – информационный элемент ИЕ; n – количество ИЭ, составляющих $ИЕ_{O_{ИО}}$.

ИЭ является неделимой структурной частью ИЕ. Существует и другое определение данного термина: «Информационный элемент – наименьшая, логическая неделимая часть данных, допускающая независимое обращение в процессе обработки данных» [3].

На рис. 1 показано, что количество характеристик и уровней детализации $O_{ИО}$ может быть различным, а $ИЕ_{ИО}$ и $ИЭ_{ИЕ}$ одного уровня могут быть присущи нескольким $ИЕ_{ИО}$ и $ИЭ_{ИЕ}$ вышележащих уровней.

Таким образом, в базах данных достаточно размещать один экземпляр конкретных $ИЕ_{ИО}$ и $ИЭ_{ИЕ}$ с указанием их связей с характеристиками не только одного конкретного $O_{ИО}$, но и всех других. Это позволит более рационально использовать ресурсы технических средств хранения и уменьшить время формирования предоставляемой пользователю информации. Весь ИР АИС предлагается представлять в виде информационного пространства системы.

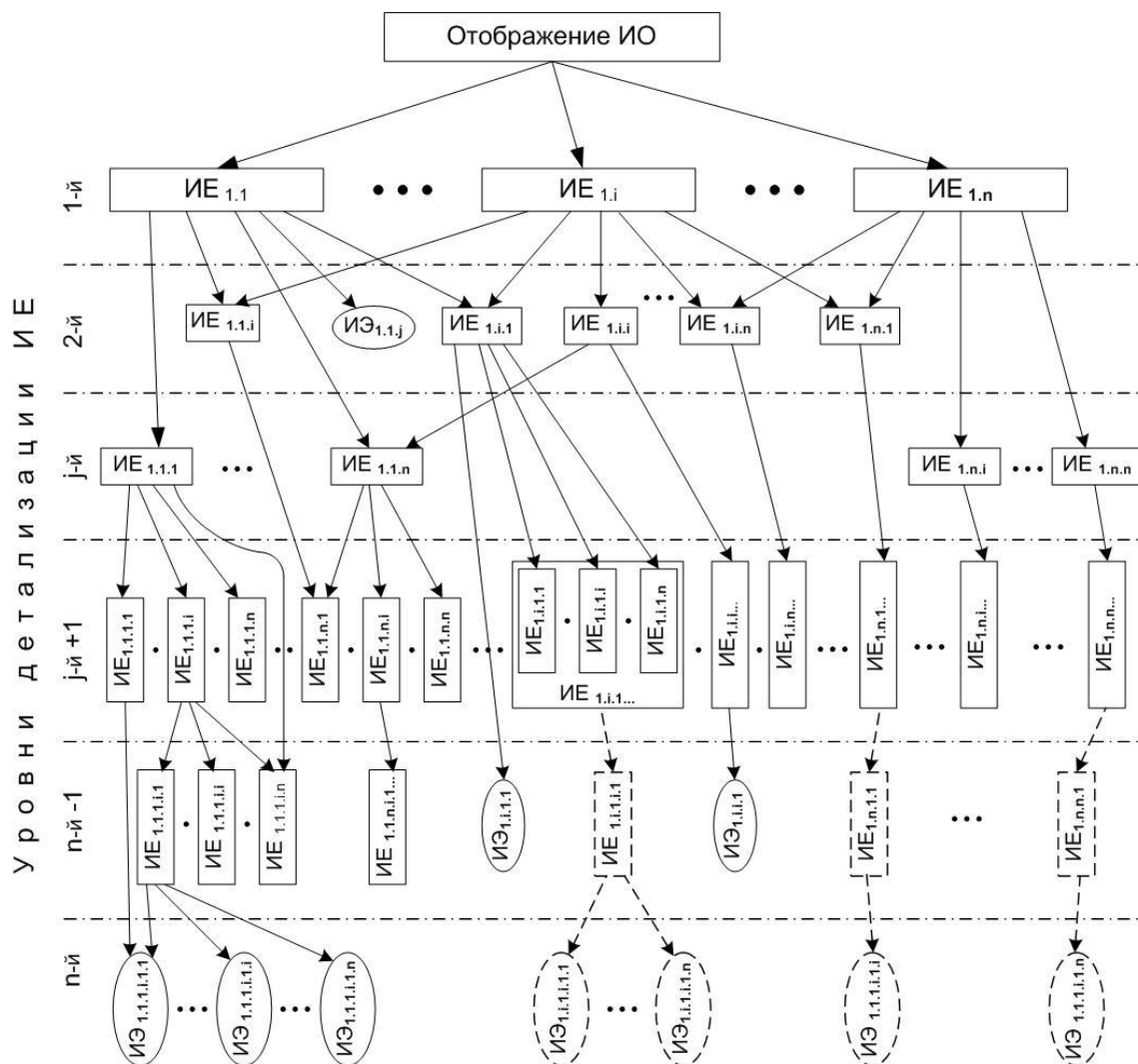


Рис. 1. Пример структуры ОИО

Информационное пространство системы (пространство ИР) представляет собой всю обрабатываемую в системе информацию, вспомогательные информационные средства системы (словари, справочники, классификаторы) и сведения о системе и ее пользователях (рис. 2).

Учитывая состав функций, которые выполняются в АИС и используют ОИО, пространство ИР можно представить в виде двух частей:

основных средств – ОИО, помещаемых, хранимых и предоставляемых пользователю (процессу);

инфраструктуры ИР – комплекса взаимосвязанных обслуживающих структур системы, обеспечивающих функционирование ее основной части (основных средств ИР) [4].

Предлагается следующий состав инфраструктуры ИР:

– средства, обеспечивающие доступ к ресурсам, прием, хранение и предоставление ИО пользователю;

– средства управления ИР;

– нормативно-правовые акты, устанавливающие правила и порядок работы с информацией, размещенной на ИР;

– нормативные технические правовые документы: международные, межгосударственные, национальные, отраслевые, организаций, устанавливающие технические требования к документированию информации;

– системы классификации и кодирования размещаемой в ИР информации;

- средства, содержащие сведения о наличии и местонахождении информации в ИР (метаданные): справочник адресов данных об ИО и справочник адресов документов (рубрикатор);
- базовые массивы данных: словарь рубрик документов, справочники баз данных документов, классификаторы объектов деятельности организаций – пользователей ИР, классификаторы документов организаций – пользователей ИР;
- библиотеки шаблонов документов, разрабатываемых организациями – пользователями ИР;
- справочно-информационные материалы: справочники организаций – пользователей ИР, справочники сведений о приложениях, использующих ИР, справочники системных сведений об ИР, словари (языковые, толковые, энциклопедические, специальные).

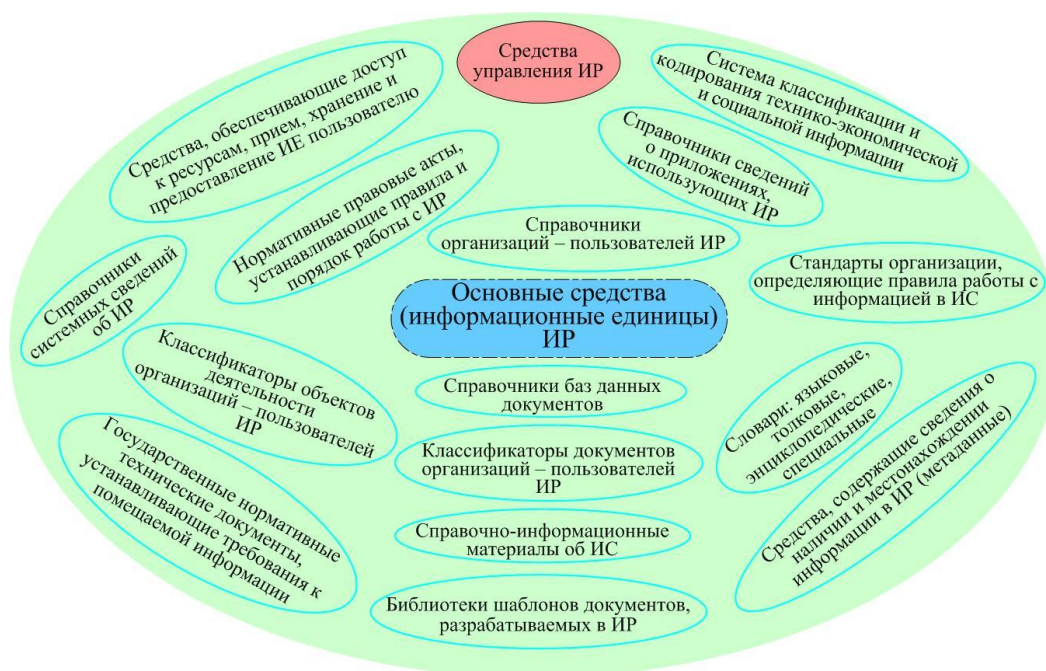


Рис. 2. Пространство ИР

Функциональные части инфраструктуры ИР выполняют:

- ведение библиотек, справочников;
- учет запросов пользователей ИР;
- анализ запросов, условий и требований пользователя;
- мониторинг ИР;
- определение задач по усовершенствованию организации ИР;
- организацию функциональной структуры ИР;
- оперативное управление ИР, куда входят анализ результатов обслуживания пользователей ИР и определение возможности адаптации ИР, процедуры адаптации ИР, а также разработка мероприятий по коррекции планирования использования ИР;
- стратегическое (упреждающее) управление ИР, включающее разработку и реализацию мероприятий по усовершенствованию организации ИР, а также прогнозирование характеристик нагрузок на ИР;
- модернизацию (развитие) ИР.

Объем пространства ИР определяется выражением

$$\Pi_{\text{ИР}} = \sum_{d=1}^m \sum_{l=1}^n \text{ИЕ}_{\text{ИО осн. ср}} + \sum_{g=1}^h \sum_{l=1}^k \text{ИЕ}_{\text{ИО инф}}, \quad (5)$$

где $\Pi_{\text{ИР}}$ – пространство ИР; $\text{ИЕ}_{\text{ИО осн. ср}}$ – ИЕ основных средств обрабатываемой информации ИР;

n – количество ИЕ основных средств, представляющих конкретный ИО; m – количество $O_{ИО}$ основных средств, представленных в ИР; $ИЕ_{ИО\ инф}$ – ИЕ инфраструктуры ИР; k – количество ИЕ, представляющих конкретный $O_{ИО}$ инфраструктуры ИР; h – количество $O_{ИО}$ инфраструктуры ИР.

Предлагается следующий состав основных характеристик ИР.

Общий объем памяти, занимаемой первичными $O_{ИО}$ основных средств и инфраструктуры ИР, размещаемых на ИР АИС, определяется как сумма:

$$Q_{\text{перв. } O_{ИО}} = \sum_{i=1}^n (Q_{\text{перв. } O_{ИО} \text{ осн. ср}}^i + Q_{\text{перв. } O_{ИО} \text{ инф}}^i)_{i\text{-го сервера АИС}}, \quad (6)$$

где $Q_{\text{перв. } O_{ИО}}^i$ – объем памяти основных средств ИР, размещаемых на ИР i -го сервера информационной системы, МБ; $Q_{\text{перв. } O_{ИО} \text{ инф}}^i$ – объем памяти средств инфраструктуры ИР, размещаемых на ИР i -го сервера информационной системы, МБ; n – количество серверов системы, на которых хранятся первичные $O_{ИО}$ основных средств, представленных в ИР.

Объем доступной пользователю информации по проблеме определяется согласно выражению

$$Q = \sum_{i=1}^n I_{i \text{ вид}}, \quad (7)$$

где $I_{i \text{ вид}}$ – объем i -го вида информации, МБ; n – количество видов информации, предоставляемой пользователю.

Полнота предоставляемой информации ИР (П) – это соотношение между имеющейся в ИР информацией по проблеме и той информацией, которая может быть предоставлена пользователю согласно правам доступа. Чем больше информации содержит ИР по конкретной проблеме, тем выше эффективность использования данного ресурса при последующем его предоставлении пользователям, имеющим различный уровень знаний о ней:

$$П = \frac{Q_{\text{инф. дост}}}{Q_{\text{инф. ИР}}} = 1 \quad (8)$$

при

$$Q_{\text{инф. дост}} \rightarrow Q_{\text{инф. ИР}},$$

где $Q_{\text{инф. дост}}$ – объем доступной пользователю информации, МБ; $Q_{\text{инф. ИР}}$ – объем информации ИР, МБ.

Применение формулы (8) позволяет количественно определить значение коэффициента полноты предоставляемой пользователю информации по отношению к информации, находящейся в ИР.

Достаточность информации ИР (Д) является его количественной характеристикой, которая определяет возможность достижения поставленной цели при наличии информации, доступной пользователю данного ИР, т. е. когда значение $Д$ стремится к единице [5]:

$$Д = \frac{Q_{\text{инф. дост}}}{Q_{\text{инф. треб}}} = 1 \quad (9)$$

при

$$Q_{\text{инф. дост}} \rightarrow Q_{\text{инф. треб}},$$

где $Q_{\text{инф. дост}}$ – объем доступной пользователю информации, МБ; $Q_{\text{инф. треб}}$ – объем требуемой информации, МБ.

Применение формулы (9) позволяет количественно определить значение коэффициента достаточности предоставляемой пользователю информации по отношению к требуемой информации, находящейся в ИР.

Доступность информации ИР характеризует предоставление пользователю возможности работы с информацией, ее получения (в том числе и в течение заданного времени) и дальнейшего использования. При этом предоставляется возможность получения доступа к источникам, содержащим требуемую информацию, в удобной для пользователя форме, что в определенной мере может влиять на восприятие предоставляемой информации и извлечение из нее необходимых пользователю сведений.

Доступность информации, как правило, регламентируется системой ограничений различного уровня: от установленных в конкретной организации – пользователе АИС – правил до законодательных актов, принятых на государственном и международном уровнях.

Использование перечисленных выше характеристик позволит определить значения показателей качества услуг по предоставлению информации, необходимых для прогнозирования и корректировки регламента доступа к информации обслуживания.

Определение «доступность информации – свойство быть доступным и готовым к использованию по запросу авторизованного субъекта» (ГОСТ 33707-2016 Информационные технологии (ИТ). Словарь Information technologies, URL: <http://docs.cntd.ru/document/1200139532>), по мнению автора, является не совсем корректным, так как оно в данном случае характеризует не тип и вид объекта, к которому предоставлен доступ, а то, что информация сама может устанавливать данное свойство.

Рассмотренные показатели характеризуют статические свойства ИР.

Обслуживание пользователей ИР выполняется по установленным правилам в соответствии с принимаемыми запросами. ИР осуществляют прием, отбор и систематизацию вводимых сведений. Свойства ИР на стадии эксплуатации жизненного цикла постоянно изменяются и определяются динамическими характеристиками системы.

Одним из основных показателей потока запросов к ИР является частота их обращения, которая характеризуется следующими свойствами:

- частота запросов информации ($F_{\text{зап. инф}}$) значительно больше, чем частота помещения ее на сервер ($R_{\text{пом. инф}}$) и обновлений ($F_{\text{зап. инф}} > R_{\text{пом. инф}}$);
- частоты запросов информации и ее помещения на сервер примерно одинаковы ($F_{\text{зап. инф}} \approx R_{\text{пом. инф}}$);
- частота запросов информации значительно меньше, чем частота ее помещения на сервер ($F_{\text{зап. инф}} < R_{\text{пом. инф}}$).

В качестве характеристик потока запросов предлагается взять показатели частотности запросов информации к ИР: постоянно запрашиваемую информацию (invariable, $F_{\text{зап. инф inv}}$) и эпизодически запрашиваемую информацию (occasionally, $F_{\text{зап. инф occ}}$). Характеристики потоков запросов и процессов репликации баз данных определяются свойствами ИО: тематикой информации, ее актуальностью и др.

В корпоративных АИС, как правило, выполняется рассылка циркулярных уведомлений об изменениях в базах данных, которая предшествует процессам репликации данных. Процессы репликации баз данных серверов характеризуются такими показателями, как оперативность репликации информации (минимальное время ожидания рассылки данных после окончания их поступления или обновления) и ритмичность помещения информации в базу данных.

Схема рассылки реплицируемых баз данных в зависимости от их содержания может изменяться. Она определяется наличием запросов информации к конкретным серверам.

Заключение. Предлагаемые количественные функциональные характеристики ИР ориентированы на содержательные свойства размещаемой в них информации и позволяют структурировать ее именно по смыслу с неограниченной степенью детализации, необходимой для поиска и систематизации данных. Эти характеристики можно применять только для детализации представления ИО в ИР, но ни в коем случае не для их синтеза.

Количественные значения функциональных характеристик $O_{ио}$, помещенных в ИР, дают возможность более рационально использовать ресурсы памяти технических средств хранения АИС, уменьшить время формирования предоставляемой пользователю информации и повысить ее информативность за счет поиска по смысловому содержанию.

Предлагаемые в статье количественные характеристики целесообразно применять для ИР, в которых размещаются большие объемы структурированной информации, например big data, и при создании ИИС. Они могут быть также использованы для описания ИР в формализованном виде и, соответственно, предоставления возможности аналитической обработки и оценки их эффективности.

Применение рассматриваемых количественных характеристик $O_{ио}$ ИР позволит на стадиях создания и эксплуатации (реализации процессов непрерывного совершенствования) АИС повысить качество проектных решений и эксплуатационные показатели этих АИС.

Список использованных источников

1. Об информации, информатизации и защите информации. Статья 1. Основные термины, применяемые в настоящем Законе, и их определения [Электронный ресурс] : Закон Респ. Беларусь от 10 нояб. 2008 г. № 455-3 // Национальный правовой интернет-портал Республики Беларусь. – Режим доступа: <http://pravo.by/document/?guid=3871&p0=H10800455>. – Дата доступа: 10.01.2019.
2. Информационные элементы и их виды [Электронный ресурс] // Государственная публичная научно-техническая библиотека России. – Режим доступа: <http://www.gpntb.ru/win/book/1/Doc8.HTML>. – Дата доступа: 10.01.2019.
3. Методы анализа и синтеза модульных информационно-управляющих систем / Н. А. Кузнецов [и др.]. – М. : Физматлит, 2002. – 800 с.
4. Инфраструктура : большой энциклопедический политехнический словарь [Электронный ресурс] // АКАДЕМИК. – Режим доступа: <https://dic.academic.ru/dic.nsf/polytechnic/3509/ИНФРАСТРУКТУРА>. – Дата доступа: 10.01.2019.
5. Информация. Источники научной информации. Аналитико-синтетическая переработка информации [Электронный ресурс] // Library.ru – информационно-справочный портал. – Режим доступа: <http://library.mephi.ru/icb2/glav2.html>. – Дата доступа: 10.01.2019.

References

1. Law of the Republic of Belarus November 10, 2008 № 455-3, Ob informacii, informatizacii i zashhite informacii. Stat'ja 1. Osnovnye terminy, primenjaemye v nastojashhem Zakone, i ih opredelenija [About information and informatization and data protection, article 1. main terms used in this law, and their definitions]. Nacional'nyj pravovoj internet-portal Respubliki Belarus' [National Legal Internet Portal of the Republic of Belarus] (in Russian). Available at: <http://pravo.by/document/?guid=3871&p0=H10800455> (accessed 10.01.2019).
2. Informacionnye jelementy i ih vidy [Information elements and their types]. Gosudarstvennaja publichnaja nauchno-tehnicheskaja biblioteka Rossii [State Public Scientific and Technical Library of Russia] (in Russian). Available at: <http://www.gpntb.ru/win/book/1/Doc8.HTML> (accessed 10.01.2019).
3. Kuznetsov N. A., Kulba V. V., Kovalevsky S. S., Kosyachenko S. A. Metody analiza i sinteza modul'nyh informacionno-upravljajushhih sistem. *Methods of Analysis and Synthesis of Modular Information and Control Systems*. Moscow, Fizmatlit, 2002, 800 p. (in Russian).
4. Infrastruktura : bol'shoj jenciklopedicheskij politehnicheskij slovar' [Infrastructure : large encyclopedic polytechnic dictionary]. АКАДЕМИК [ACADEMICIAN] (in Russian). Available at: <https://dic.academic.ru/dic.nsf/polytechnic/3509/ИНФРАСТРУКТУРА> (accessed 10.01.2019).
5. Informacija. Istochniki nauchnoj informacii. Analitiko-sinteticheskaja pererabotka informacii [Information. sources of scientific information. Analytical and synthetic processing of information]. Library.ru – informacionno-spravochnyj portal. [Library.ru – Information and Reference Portal] (in Russian). Available at: <http://library.mephi.ru/icb2/glav2.html> (accessed 10.01.2019).

Информация об авторе

Тоцицкий Леонард Иванович, главный конструктор проекта отдела технологий цифровой трансформации, Объединенный институт проблем информатики Национальной академии наук Беларуси, Минск, Беларусь.
E-mail: Leonard.Tochicki@tut.by

Information about the author

Leonard I. Tochicki, Chief Designer of the Project of the Department of Digital Transformation Technology, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Belarus.
E-mail: Leonard.Tochicki@tut.by