

Открытые данные: анализ тенденций

Качан Д. А., заместитель директора по научной работе

E-mail: kachan@giac.by

Учреждение «Главный информационно-аналитический центр
Министерства образования Республики Беларусь», ул. Захарова,
д. 59, 220088, г. Минск, Республика Беларусь

Аннотация. В работе проведен анализ использования открытых данных, рассмотрены проблемы, связанные с открытыми данными, и пути их решения. Проведена оценка тенденций развития публикации государственных данных в открытых форматах. Представлена оценка объемов рынка открытых данных в Республике Беларусь. Проведен анализ применимости мирового опыта, рассмотрены актуальные вопросы обеспечения должного уровня защиты информации при публикации наборов данных, предложены эффективные способы их использования, отмечена важность обеспечения достаточного нормативно-правового регулирования. Рассмотрен экономический и социально-культурный эффект от распространения государственных данных в открытом формате. В заключении сформулированы наиболее актуальные вопросы, касающиеся открытых данных, которые будут определять эффективность развития информационных технологий в ближайшем будущем.

Ключевые слова: открытые данные, информационная безопасность, эффект мозаики, нормативно-правовая база открытых данных

Для цитирования: Качан Д. А. Открытые данные: анализ тенденций / Д. А. Качан // Цифровая трансформация. – 2018. – № 1 (2). – С. 72–78.

© Цифровая трансформация, 2018

Open Data: Analysis of Trends

Kachan D. A., Deputy Director for Research

E-mail: kachan@giac.by

Establishment “The Main Information and Analytical Center of the
Ministry of Education of the Republic of Belarus”, 59 Zakharova Str.,
220088 Minsk, Republic of Belarus

Abstract. The paper analyzes the use of open data, discusses problems associated with open data, and solutions for working with them, and assesses trends in the publication of public data in open formats. The estimation of volumes of the open data market in the Republic of Belarus is presented. The analysis of the applicability of the world experience on data discovery for the Republic of Belarus was conducted, the topical issues of ensuring the proper level of information protection when publishing data sets were discussed, effective ways of using open data were suggested, and the importance of ensuring sufficient regulatory regulation was noted. The economic and socio-cultural effect of the dissemination of state data in an open format is considered. In conclusion, the most topical issues of open data are formulated, which will determine the effectiveness of the development of information technology in the near future.

Key words: open data, information security, mosaic effect, the legal basis of open data

For citation: Kachan D. A. Open Data: Analysis of Trends. *Cifrovaja transformacija* [Digital transformation], 2018, 1 (2), pp. 72–78 (in Russian).

© Digital Transformation, 2018

Введение. Текущий этап развития информационно-коммуникационных технологий характеризуют прежде всего безграничные возможности сбора, хранения и обработки информации. В связи с этим происходит интенсивное внедрение информацион-

ных систем во все сферы жизни современного общества. Стало возможным формирование больших массивов данных, доступных массовому потребителю, которые могут использоваться без ограничений. Другим бурно растущим видом деятельности специ-

алистов IT-индустрии является попытка встраивать данные в свободно разрабатываемые приложения. Основной задачей работы является анализ возможностей и ограничений использования технологии открытых данных, в том числе в вопросах обеспечения информационной безопасности.

Основная часть. История возникновения. В 2013 году страны-участники саммита G8, «Большой восьмерки», приняли документ «Хартия открытых данных», который предусматривает публичное раскрытие информации государственных органов в сети Интернет [1].

Тем самым был положен старт движению мировой цивилизации к открытости государственных данных. Поставить свою подпись под Хартией были приглашены и страны, не входящие в G8, что придало инициативе глобальный характер.

Главные принципы, которые прописаны в Хартии — это открытость данных по умолчанию, своевременная их публикация в машиночитаемом виде, прозрачность и обязательство

обеспечивать условия, в которых разработчики будут создавать приложения на основе открытых данных.

В соответствии с приложением к [1] определены 14 приоритетных направлений раскрытия данных, представленных в таблице 1.

Идеология открытых государственных данных поддерживается крупнейшими государствами мира и международными организациями — Open Government Partnership (OGP), International Budget Partnership (IBP), World Wide Web Consortium (W3C), The World Bank Group (Всемирный банк), Организация экономического сотрудничества и развития (ОЭСР), Open Knowledge Foundation (OKF) и т. д.

Альтернативный взгляд. Поддержка международными организациями государств-участников зачастую политизирована, и участие в этих организациях предполагает проведение конкретных государственных реформ [2], результатом которых станет в том числе и формирование наборов открытых данных.

Таблица 1. Приоритетные направления раскрытия данных

Категория данных	Примеры наборов открытых данных
Компании	Реестр компаний/предприятий
Преступность и правосудие	Статистика преступности, безопасность
Наблюдение за планетой	Метеорологические данные, сведения о погоде, сельском хозяйстве, лесоводстве, рыболовстве и охоте
Образование	Список учреждений образования; результативность, навыки работы с ИКТ
Энергетика и окружающая среда	Уровни загрязнения, энергопотребление
Финансовые и контрактные вопросы	Заклученные сделки, подписанные контракты, поданные заявки на участие в тендере, будущие тендеры, местные бюджеты, национальный бюджет (планируемый и расходный)
Геопространственные данные	Картография, почтовые коды
Международное развитие	Медицинская помощь, продовольственная безопасность, добывающая промышленность, землепользование
Подотчетность правительственного аппарата и демократия	Контактная информация, результаты выборов, нормативно-правовые акты, заработная плата, награды
Здравоохранение	Данные о назначаемых препаратах, данные о результатах
Наука и исследования	Данные о геномах, исследовательская и образовательная деятельность, результаты экспериментов
Статистика	Национальная статистика, перепись, инфраструктура, уровень благосостояния, профессиональные навыки
Социальная мобильность и благосостояние	Жилищное обеспечение, медицинское страхование и пособие по безработице
Транспорт и инфраструктура	Расписание общественного транспорта, транспортная инфраструктура, доступ к сетям связи

Результаты сотрудничества отражены в виде отчетов и наборов открытых данных, размещаемых на официальных порталах международных организаций [3].

По правилам сбора данных, установленных «The World Bank Group» (Всемирный банк), «частные подрядчики, осуществляющие сбор первичных данных, предпочтительнее любых правительственных учреждений или организаций, связанных с правительством» [3].

Венгрия вынуждена была отказаться от членства в Open Government Partnership после поступления в адрес OGP письма [4], подписанного всего четырьмя субъектами. Официальной причиной прекращения сотрудничества послужило то, что двадцатистраничный официальный отчет с данными не был принят во внимание либо принят со значительными искажениями данных [5]. Кроме того, OGP покинули Турция и Азербайджан.

На основании анализа данных, публикуемых международными организациями можно отметить:

- Corruption Perceptions Index (индекс восприятия коррупции), представляющий собой машиночитаемую таблицу открытых данных, сформированную из 13 источников данных, два из которых — открытые данные Всемирного банка;

- данные, предоставляемые Всемирным банком и публикуемые на официальных веб-ресурсах, основаны на опросах фокус-групп, а не на официальной статистике [5;6];

- при проведении опросов фокус-групп наблюдается изменение количества показателей оценки для разных стран: для индустриальных лидеров Европы это как правило 63 показателя, для остальных — около 400 показателей [7–12];

- публикация обновленных наборов данных большинства стран не осуществляется Всемирным банком с 2014 года, однако Corruption Perceptions Index продолжает обновляться с указанием наборов данных при его формировании и использовании, в частности, данных Всемирного банка при формировании показателей (косвенно указывает на то, что данные имеются в наличии) [5].

Подобный подход с реализацией международными организациями своего альтернативного взгляда на формирование открытых данных фактически ставит под сомнение важность принимаемых мер по выработке единого универсального подхода по развитию открытых правительств стран-участников и стран, стремящихся присоединиться к Хартии открытых данных.

Экономический эффект. Экономический эффект создания наборов открытых данных досто-

верно оценить не представляется возможным. Это прежде всего связано с тем, что определение наборов открытых данных, доступных бизнес-сообществу в соответствии с [6], предложено формировать с учетом его интересов. Лишь на этапе формирования технико-экономического обоснования проекта, использующего в своей основе открытые данные, можно провести предварительную оценку экономического эффекта.

При описании экономического эффекта многие авторы работ, посвященных открытым данным, ссылаются на исследования McKinsey Global Institute 2013 года [13], специалисты которого приводят сумму равную трем трлн долларов мирового экономического эффекта. В работе отсутствует указание методики оценки и способы формирования значения. В разделе «Key findings» указано, что Европа получит экономический эффект равный 900 млрд долларов.

По оценкам европейских специалистов [14] прямое и косвенное воздействие информации государственного сектора на экономику 28 стран Европейского союза (EU28+) по состоянию на 2017 год превышает 200 млрд евро. Объемы рынка по непосредственному использованию информации государственного сектора в 2017 году для 28 европейских стран (EU28), а также Норвегии, Исландии, Лихтенштейна, Швейцарии превысили 50 млрд евро.

Рост указанных показателей до 2020 года оценивается в 265–286 и 83,5 млрд евро соответственно. Динамика роста на период 2016–2020 приведена на рисунке 1 [14].

По оценкам российских специалистов [15] прямой доход государства от продажи информации госсектора меньше экономического эффекта от непосредственного использования информации государственного сектора бизнес-сообществом в 10–20 раз.

Среди факторов, определяющих необходимость обеспечения улучшенного доступа к открытым данным, европейские специалисты выделяют: оценку воздействия на окружающую среду, позволяющую снизить затраты на 20% в год;

открытый доступ к результатам НИОКР, который позволит получать дополнительно около 6 млрд евро в год;

полноценный доступ к государственной информации, позволяющий каждому из жителей стран Евросоюза сэкономить хотя бы два часа в год и обеспечивающий экономию в масштабах Евросоюза не менее 1,4 млрд евро в год;

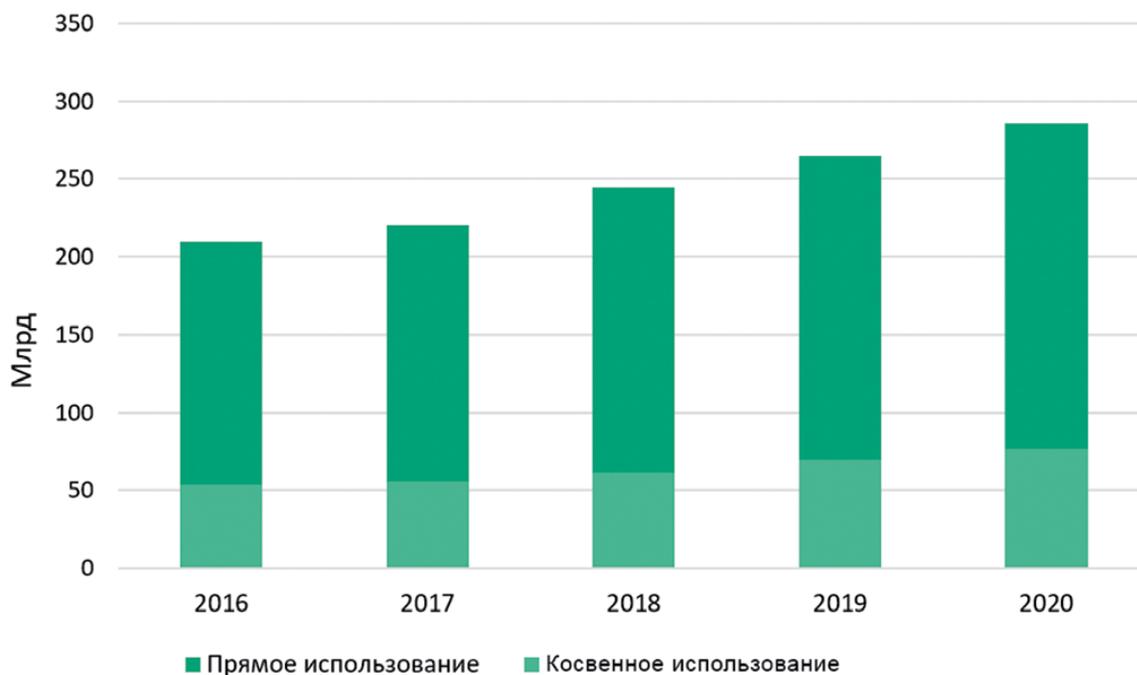


Рисунок 1. Объемы рынка открытых данных по 28 странам Европейского союза в 2016–2020 гг.

ряд детальных национальных исследований, показывающих высокий потенциал рынков и новых приложений информации государственного сектора [15].

Объем рынка открытых данных. Объем рынка использования наборов открытых данных органов государственного управления Республики Беларусь можно оценить по аналогии с рынком EU28+, полагая, что объемы этих рынков соотносятся так же, как количество пользователей сети Интернет. Величина для Республики Беларусь составит 300–500 млн евро в год.

Алгоритм определения наборов данных для бизнес-сообщества. Алгоритм действий при определении наборов открытых данных может включать следующие шаги:

- совместная консультационная работа ответственных за подготовку открытых данных с представителями бизнес-сообщества;
- определение бизнес-сообществом тех наборов данных, которые на их взгляд способны дать максимальный экономический эффект;
- оценка органами государственного управления возможности подготовки наборов открытых данных и согласование наборов;
- подготовка бизнес-сообществом технико-экономического обоснования и бизнес-модели для окончательного решения в пользу открытия данных.

Такой алгоритм связан прежде всего с вероятными значительными трудовыми затратами органов

госуправления на подготовку данных, которые в итоге могут быть не востребованы.

Социально-культурный эффект. Социально-культурный эффект открытия данных заключается в повышении прозрачности и подотчетности деятельности органов государственной власти, развитии гражданского общества через возможность контроля и мониторинга деятельности государственных органов, укреплению доверия граждан к органам власти.

На основании наборов открытых данных осуществляется создание социально-значимых сервисов, повышающих комфорт граждан, предоставляющих дополнительные возможности для улучшения их жизни и деятельности [16].

Формирование наборов данных возможно на основании результатов опроса потенциальных пользователей и экспертов, а также специалистов в области обработки статистических данных.

Информационная безопасность и защита персональных данных. Вопросы нормативно-правового обеспечения процессов открытия наборов данных не рассматриваются в данной работе, однако стоит упомянуть, что страны-участники данной инициативы имеют свои собственные законодательные акты, определяющие регламенты проведения процедур.

Ввиду отсутствия подобных актов законодательства в Республике Беларусь необходимо руководствоваться Законом Республики Беларусь от 19 июля 2010 г. №170-З «О государственных секретах».

В разрезе оценки данных на предмет возможности предоставления публичного доступа необходимо привести базовую классификацию данных:

1. Инфраструктурные данные — данные о состоянии мира, например, описывающие землю, транспортные сети, структуры правительства, метеорологические измерения и так далее. Данные, за незначительным исключением, не являются конфиденциальными (в некоторых государствах проблемы безопасности могут ограничивать предоставление инфраструктурных данных, например, данные о географической границе и потоке воды в реках Северной Индии ввиду напряженных отношений с определенными государствами [17]).

2. Данные государственных служб — данные о деятельности правительства, начиная от мест предоставления государственных услуг и их бюджетов и заканчивая публичными реестрами, а также детальной статистикой работы учреждений образования, здравоохранения и объектов социально-культурного назначения.

3. Данные ограниченного распространения — персонифицированные данные о физических лицах, данные, отнесенные к государственным секретам и иная информация, распространение которой ограничено.

В качестве открытых данных рассматриваются практически все наборы данных из группы 1 и группы 2.

История развития концепции открытых данных имеет один показательный прецедент, произошедший в США: после трагических событий в учебном заведении города Сэнди Хук, штат Коннектикут, где в результате использования злоумышленником огнестрельного оружия погибли 20 учащихся и 6 взрослых [19], одна из газет опубликовала адреса людей, обладающих огнестрельным оружием. В электронной версии статьи присутствовало изображение картографического сервиса Google с нанесенными отметками [20]. Публикация вызвала серьезный резонанс и общественное обсуждение в США относительно пределов допустимого вмешательства в частную жизнь граждан.

В контексте рассмотренной выше ситуации, можно было предоставить обезличенные

данные — подготовить необходимую для публикации инфографику, характеризующую цветовой градацией те или иные характеристики рассматриваемой территории. Обезличенные персонифицированные данные наряду с инфраструктурными и данными государственных служб являются источником ценной информации как для бизнес-сообщества, так и для формирования активной гражданской позиции общества и осуществления мониторинга деятельности госорганов.

Эффект мозаики. Так называемый эффект мозаики возникает, когда информация отдельного набора данных не создает угрозу нарушения информационной безопасности, но в сочетании с другими наборами открытых данных может представлять такой риск [20].

В связи с этим при определении наборов открытых данных необходимо проводить комплекс мер на предмет поиска комбинации данных, способной нести угрозу информационной безопасности.

Заключение. Можно сделать следующие выводы:

1. Формирование наборов открытых данных способно лишь косвенно повлиять на членство в международных организациях.

2. Сбор открытых данных международными организациями осуществляется в интересах промышленно развитых стран.

3. Публикация наборов открытых данных международными организациями является в большей степени политическим инструментом, нежели механизмом развития мирового сообщества.

4. Объем рынка использования наборов открытых данных Республики Беларусь составляет около 300–500 млн евро ежегодно.

5. Публикация данных осуществляется на основании нормативно-правовых актов и разработанных с этой целью регламентов.

6. Публиковаться могут обезличенные персонифицированные данные, инфраструктурные данные и данные государственных служб после их детального рассмотрения и анализа в совокупности.

7. Для проведения анализа по оценке требуется разработка методических указаний.

Список литературы

1. G8 Open Data Charter and Technical Annex [Electronic resource]. – Mode of access: <https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex#technical-annex>. – Date of access: 19.04.2018.

2. The road to budget transparency: learning from country experience [Electronic resource]. – Mode of access: <https://www.internationalbudget.org/publications/road-to-budget-transparency-six-country-synthesis/>. – Date of access: 19.04.2018.
3. World Bank Open Data [Electronic resource]. – Mode of access: <https://data.worldbank.org/>. – Date of access: 19.04.2018.
4. OGP Hungary response policy letter [Electronic resource]. – Mode of access: <http://www.opengovpartnership.org/sites/default/files/attachments/OGP%20Hungary%20response%20policy%20letter.pdf>. – Date of access: 19.04.2018.
5. Media Briefing: Government of Hungary Withdraws from the Open Government Partnership [Electronic resource]. – Mode of access: <https://www.opengovpartnership.org/about/news-and-events/media-briefing-government-of-hungary-withdraws-open-government-partnership>. – Date of access: 19.04.2018.
6. Corruption perceptions index 2017 [Electronic resource]. – Mode of access: https://www.transparency.org/news/feature/corruption_perceptions_index_2017#resources. – Date of access: 19.04.2018.
7. France — Global Financial Inclusion (Global Findex). Database 2014 [Electronic resource]. – Mode of access: http://microdata.worldbank.org/index.php/catalog/2419/data_dictionary. – Date of access: 19.04.2018.
8. Germany — Global Financial Inclusion (Global Findex) Database 2014 [Electronic resource]. – Mode of access: http://microdata.worldbank.org/index.php/catalog/2409/data_dictionary. – Date of access: 19.04.2018.
9. Belarus - Enterprise Survey 2013 [Electronic resource]. – Mode of access: http://microdata.worldbank.org/index.php/catalog/1766/data_dictionary. – Date of access: 19.04.2018.
10. Russian Federation — World Bank Group Country Survey 2014 [Electronic resource]. – Mode of access: http://microdata.worldbank.org/index.php/catalog/2184/data_dictionary. – Date of access: 19.04.2018.
11. Kazakhstan — Enterprise Survey 2013 [Electronic resource]. – Mode of access: http://microdata.worldbank.org/index.php/catalog/1977/data_dictionary. – Date of access: 19.04.2018.
12. China — Enterprise Survey 2012 [Electronic resource]. – Mode of access: http://microdata.worldbank.org/index.php/catalog/1559/data_dictionary. – Date of access: 19.04.2018.
13. Open data: Unlocking innovation and performance with liquid information [Electronic resource]. – Mode of access: <https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/open-data-unlocking-innovation-and-performance-with-liquid-information>. – Date of access: 19.04.2018.
14. Market value Open Data to reach 286 billion by 2020 [Electronic resource]. – Mode of access: <https://www.consultancy.uk/news/3019/market-value-open-data-to-reach-286-billion-by-2020>. – Date of access: 19.04.2018.
15. Аналитический отчет по результатам проведения сравнительного анализа российского и зарубежного опыта в сфере обеспечения доступа к государственным открытым данным в целях создания социально-значимых интернет-сервисов. – НИУ ВШЭ, 2012.
16. Городские данные // Открытые данные Беларуси [Электронный ресурс]. – Режим доступа: <https://opendata.by/group/1430?page=1>. – Дата доступа: 19.04.2018.
17. Khadka N. S. China and India water 'dispute' after border stand-off [Electronic resource]. – Mode of access: <http://www.bbc.com/news/world-asia-41303082>. – Date of access: 19.04.2018.
18. Sandy Hook Elementary School shooting [Electronic resource]. – Mode of access: https://en.wikipedia.org/wiki/Sandy_Hook_Elementary_School_shooting. – Date of access: 19.04.2018.
19. Ingram M. Newspapers and guns: If data is available, should it always be published? [Electronic resource]. – Mode of access: <https://gigaom.com/2012/12/27/newspapers-and-guns-if-data-is-available-should-it-always-be-published/>. – Date of access: 19.04.2018.
20. M-13-13 — Memorandum for the Heads of Executive Departments and Agencies [Electronic resource]. – Mode of access: https://project-open-data.cio.gov/policy-memo/#i_definitions. – Date of access: 19.04.2018.

References

1. G8 Open Data Charter and Technical Annex. Available at: <https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex#technical-annex> (accessed: 19.04.2018).
2. The road to budget transparency: learning from country experience. Available at: <https://www.internationalbudget.org/publications/road-to-budget-transparency-six-country-synthesis/> (accessed: 19.04.2018).
3. World Bank Open Data. Available at: <https://data.worldbank.org/> (accessed: 19.04.2018).
4. OGP Hungary response policy letter. Available at: <http://www.opengovpartnership.org/sites/default/files/attachments/OGP%20Hungary%20response%20policy%20letter.pdf> (accessed: 19.04.2018).
5. Media Briefing: Government of Hungary Withdraws from the Open Government Partnership. Available at: <https://www.opengovpartnership.org/about/news-and-events/media-briefing-government-of-hungary-withdraws-open-government-partnership> (accessed: 19.04.2018).
6. Corruption perceptions index 2017. Available at: https://www.transparency.org/news/feature/corruption_perceptions_index_2017#resources (accessed: 19.04.2018).
7. France — Global Financial Inclusion (Global Findex). Database 2014. Available at: http://microdata.worldbank.org/index.php/catalog/2419/data_dictionary (accessed: 19.04.2018).

8. Germany - Global Financial Inclusion (Global Findex) Database 2014. Available at: http://microdata.worldbank.org/index.php/catalog/2409/data_dictionary (accessed: 19.04.2018).
9. Belarus - Enterprise Survey 2013. Available at: http://microdata.worldbank.org/index.php/catalog/1766/data_dictionary (accessed: 19.04.2018).
10. Russian Federation — World Bank Group Country Survey 2014. Available at: http://microdata.worldbank.org/index.php/catalog/2184/data_dictionary (accessed: 19.04.2018).
11. Kazakhstan - Enterprise Survey 2013. Available at: http://microdata.worldbank.org/index.php/catalog/1977/data_dictionary (accessed: 19.04.2018).
12. China — Enterprise Survey 2012. Available at: http://microdata.worldbank.org/index.php/catalog/1559/data_dictionary (accessed: 19.04.2018).
13. Open data: Unlocking innovation and performance with liquid information. Available at: <https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/open-data-unlocking-innovation-and-performance-with-liquid-information> (accessed: 19.04.2018).
14. Market value Open Data to reach 286 billion by 2020. Available at: <https://www.consultancy.uk/news/3019/market-value-open-data-to-reach-286-billion-by-2020> (accessed: 19.04.2018).
15. *Analiticheskij otchet po rezul'tatam provedenija sravnitel'nogo analiza rossijskogo i zarubezhnogo opyta v sfere obespechenija dostupa k gosudarstvennym otkryтым dannym v celjah sozdaniya social'no-znachimyh internet-servisov* [Analytical report on the results of a comparative analysis of Russian and foreign experience in providing access to public data in order to create socially significant Internet services]. Higher School of Economics, 2012.
16. *Gorodskie dannye* [City data]. Available at: <https://opendata.by/group/1430?page=1> (accessed: 19.04.2018).
17. Khadka N. S. China and India water 'dispute' after border stand-off Available at: <http://www.bbc.com/news/world-asia-41303082> (accessed: 19.04.2018).
18. Sandy Hook Elementary School shooting. Available at: https://en.wikipedia.org/wiki/Sandy_Hook_Elementary_School_shooting (accessed: 19.04.2018).
19. Ingram M. Newspapers and guns: If data is available, should it always be published? Available at: <https://gigaom.com/2012/12/27/newspapers-and-guns-if-data-is-available-should-it-always-be-published/> (accessed: 19.04.2018).
20. M-13-13 — Memorandum for the Heads of Executive Departments and Agencies. Available at: https://project-open-data.cio.gov/policy-memo/#i_definitions (accessed: 19.04.2018).

Received: 20.04.2018

Поступила: 20.04.2018