

Федеральное государственное автономное образовательное учреждение
высшего образования
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

На правах рукописи



Кижнер Инна Александровна

**ЦИФРОВЫЕ ТЕХНОЛОГИИ КУЛЬТУРНОГО НАСЛЕДОВАНИЯ
В СОВРЕМЕННОМ ОБЩЕСТВЕ**

Специальность: 24.00.01 – Теория и история культуры (культурология)

Диссертация

на соискание ученой степени кандидата культурологии

Научный руководитель
доктор философских наук
профессор Н.П. Кошцева

Красноярск – 2020

ОГЛАВЛЕНИЕ

| | |
|--|-----|
| Введение..... | 3 |
| 1 Концептуальные и методологические подходы к исследованию роли цифровых технологий в культурном наследовании..... | 22 |
| 1.1 Цифровой поворот в культурном наследии..... | 22 |
| 1.2 Коллекции цифровых изображений и агрегаторы визуального контента в области культурного наследия..... | 61 |
| 2 Методы организации и представления цифровых коллекций в области культурного наследия..... | 81 |
| 2.1 Представление российских цифровых коллекций в области культурного наследия..... | 81 |
| 2.2 Методы организации агрегаторов оцифрованного контента в области культурного наследия (на примере цифровых коллекций Музейного фонда РФ и платформы Google Arts and Culture)..... | 100 |
| Заключение..... | 154 |
| Список литературы..... | 158 |

ВВЕДЕНИЕ

Актуальность темы исследования обусловлена необходимостью понимания места современных цифровых технологий передачи культурного наследия в процессах исследования теории и истории культуры. Цифровые коллекции и собрания цифровых изображений в области культурного наследия выступают сегодня в качестве первичных источников, которые становятся явлением культуры, демонстрирующим способы познания мира и лакуны в понимании культуры. Тенденциозность агрегаторов или цифровых коллекций, возможность или ее отсутствие включить в агрегатор изображения, представляющие ряд культур и сообществ, избыточная репрезентативность одних культур и недостаточная репрезентативность других, способы классификации изображений в коллекции, способы представления культурного наследия и создания «воображаемых сообществ» влияют на представление о культуре, ее развитии, месте разных культур в общей картине мира. К тому же отсутствие представлений о культурах и культурных сообществах в случае их недостаточной представленности в агрегаторах влияет на формирование культурного капитала или степени информированности о чужих культурах и представлений о культурной идентичности других сообществ. Формирование представлений об иной культуре, основанное на недостаточном количестве данных, искаженных данных или на данных, которые представляют культуру сквозь призму западного восприятия, ведет к тому, что культурный капитал не расширяется, а социальная мобильность и взаимное обогащение культур замедляются. Более того, в этом случае анализ культуры, основанный на больших данных, будет представлять искаженную картину, демонстрируя тенденции и закономерности, характерные для культур, представленных в цифровых коллекциях и агрегаторах, и скрывая особенности тех культур, данные о которых представлены в недостаточном количестве. Рекомендательные системы, основанные на алгоритмах машинного обучения, способны дополнительно усилить те эффекты, при которых предрассудки и тенденциозность восприятия культуры доцифровой

эпохи будут тиражироваться и распространяться с помощью цифровых технологий.

Агрегаторы цифровых изображений в области культурного наследия, такие как издательская платформа Google Arts and Culture, наращивают количество изображений в своих базах данных, но мы не знаем, какие культуры, географические регионы и периоды времени представляют эти изображения. Мы также не знаем, какими принципами руководствуются редакторы агрегаторов и цифровых коллекций при отборе изображений для публикации. Цифровые коллекции не сообщают о принципах, которыми они руководствуются при публикации изображений. Несмотря на то, что целью цифровых коллекций, в частности агрегатора Google Arts and Culture, обычно бывает распространение знаний о культурах мира, степень репрезентативности и сбалансированности этих собраний цифровых изображений остается неизвестной.

Следствием избыточной репрезентативности ряда культур является смещение акцента в сторону постколониального представления культурных объектов. Предубеждения доцифрового периода, сложившиеся в эпоху колониальных культур и постколониальных попыток репрезентации «воображаемых сообществ», мигрируют в цифровые коллекции и агрегаторы и формируют новое явление цифрового культурного колониализма. В этих коллекциях доминируют изображения, представленные культурами, которые доминируют в технологическом отношении, преобладают изображения из учреждений культуры центральных городов и столиц. Так или иначе тенденциозность и отсутствие репрезентативности проникают в новые цифровые инфраструктуры по разным техническим, социальным, политическим и культурным причинам и могут повлиять на будущие платформы культурного наследия и их пользователей. Недавние научные исследования, посвященные цифровым коллекциям, рассматривают разные аспекты организации и распространения данных в области культурного наследия, в том числе социальные, политические и культурологические аспекты этого явления.

Степень научной разработанности темы исследования. Проблема культурного наследия является одной из наиболее характерных для современной

культурологии. Она ставится и решается в работах таких исследователей, как Д. Лоуэнталь¹ Р. Ламли², П. Верго³, Р. Хьюисон⁴, П. Нора⁵, Д. Харви⁶ Р. Харрисон⁷, Т. Беннет⁸, Л. Смит⁹, Э. Гринхилл¹⁰. Р. Харрисон обращается к проблеме культурного наследия как к междисциплинарному явлению, которое имеет одновременно социальный и материальный характер, демонстрируя социальные, экономические и политические проблемы современности. Л. Смит предлагает использовать культурологические и социальные подходы в процессе критического анализа культурного наследия, обращая внимание на знаковую природу наследия и важность ценностей и смыслов, которые связаны с материальными объектами. Исследователь показывает, что работа с культурным наследием означает «конст-

¹ Lowenthal D. *The Past is a Foreign Country*. Cambridge: CUP, 1985.

² Lumley R. *The Museum Time-machine: Putting Cultures on Display*. London: Routledge, 1988.

³ Vergo P. *The New Museology*. London: Reaktion Books, 1989.

⁴ Hewison R. *The Heritage Industry: Britain in a Climate of Decline*. London: Methuen, 1987.

⁵ Nora, P. 1989. "Between Memory and History: Les Lieux De Mémoire." *Representations* 26: 7–25; Nora, P. 2002. "Reasons for the Current Upsurge in Memory." *Transit* 22 (1): 4–8.

⁶ Harvey, D. 1990. *The Condition of Postmodernity: An Enquiry into the Origins of Cultural Change*. Cambridge: Blackwell; Harvey David, C. 2001. "Heritage Pasts and Heritage Presents: Temporality, Meaning and the Scope of Heritage Studies." *International Journal of Heritage Studies* 7 (4): 319–338.

⁷ Harrison, Rodney 2013 *Heritage: Critical Approaches*. Routledge, Abingdon and New York; Harrison, R., ed. 2010. *Understanding the Politics of Heritage*. Manchester: Manchester University Press, with Open University; Bennett, T, Cameron, F, Dias, N, Dibley, B, Jacknis, I, Harrison R and McCarthy, C 2017 *Collecting, Ordering, Governing: Anthropology, Museums and Liberal Government*. Durham: Duke University Press.

⁸ Bennett, T. *The Birth of the Museum: History, theory, politics*, London: Routledge, 1985.

⁹ Smith, L. *The Uses of Heritage*. London: Routledge, 2006; Smith, Laurajane and Natsuko Akagawa (eds) (2009) *Intangible Heritage*. London: Routledge; Gentry, K., & Smith, L. (2019). Critical heritage studies and the legacies of the late-twentieth century heritage canon. *International Journal of Heritage Studies*, 1–21.

¹⁰ E. Hooper-Greenhill. *Museums and the Shaping of Knowledge*, Routledge, London, 1992.

руирование смыслов с материальными последствиями»¹¹. Французский исследователь П. Нора утверждает, что культурное наследие (память) и исторические исследования представляют совершенно разные направления человеческого познания, при этом память динамична и направлена на конструирование настоящего, в то время как исторические исследования используют интеллектуальные и рациональные подходы к документам, пользуются средствами критического анализа, и их выводы могут быть универсальны в любое время и при любых социальных/политических изменениях. Таким образом, ряд ученых, работающих в области критического анализа культурного наследия, обращаются к эффектам, которые возникают при отображении настоящего и будущего в фиксации прошлого.

Цифровые технологии в контексте проблемы культурного наследия – сравнительно новая тема для теории и истории культуры. Однако здесь можно выделить таких фундаментальных ученых как М. Кастельс¹², Н.К. Хейлс¹³, М. Дьюз¹⁴, Л. Манович¹⁵, Н. Коулдри¹⁶.

Концептуальные и методологические подходы к развитию цифровых коллекций и агрегаторов предметов из коллекций культурного наследия представлены в работах, опубликованных в последние пятьдесят лет. Особенно обширно эта тема освещается в литературе за последние двадцать лет, когда сетевые и информационные технологии сделали возможным широкое распространение изображений из цифровых коллекций и размещение изображений высокого разрешения в крупных агрегаторах.

¹¹ Smith, L. *The Uses of Heritage*. London: Routledge, 2006.

¹² Castells, M. *The Rise of the Network Society, the Information Age: Economy, Society and Culture*. Oxford: Oxford University Press, 1996.

¹³ Hayles, N.K. *How We Think: Digital Media and Technogenesis*. Chicago, IL: University of Chicago Press, 2012.

¹⁴ Deuze, Mark (2011) 'Media life', *Media Culture and Society* 33(1): 137–148.

¹⁵ Manovich, L. (2001). *The Language of New Media*. Cambridge, MA: MIT Press.

¹⁶ Couldry N and Hepp A (2017) *The Mediated Construction of Reality*. Cambridge: Polity Press.

Концептуальные подходы к созданию и развитию цифровых коллекций в области культурного наследия близки подходам, которые разрабатывались в двадцатом веке в области структурализма. Необходимость перехода от идиосинкратических описаний произведений культуры и искусства к формализованным стандартам представления культурного наследия в целях предмашинной подготовки и связанного с этим редуцирования представления о сложном произведении отметил в очерке, посвященном истории англо – американских подходов к оцифровке коллекций, Р. Парри¹⁷. Проблеме эпистемических ограничений цифровых коллекций уделил особое внимание В. МакКарти¹⁸ в книге, посвященной машинному моделированию связей между предметами культурного наследия при работе с документацией, фиксирующей эти связи в цифровых коллекциях. Ранние работы, связанные с индексацией и документацией предметов культурного наследия для машинной обработки, касались каталогизации музейных коллекций¹⁹, увеличения технических возможностей цифровых библиотек в области культурного наследия, а вслед за этим увеличения их объемов и масштабов взаимодействия с пользователями²⁰. Цифровые коллекции интересовали культурологов и философов как новый способ организации знания о мире и взаимодействия с миром²¹. В первые десятилетия нашего века цифровые коллекции привлекли

¹⁷ R. Parry. *Recoding the Museum: Digital Heritage and the Technologies of Change*. Routledge, 2007.

¹⁸ McCarty, Willard. *Humanities Computing*. London: Palgrave Macmillan, 2005.

¹⁹ Chenhall, R. G. *Museum cataloguing in the computer age*, Nashville, TN: American Association for State and Local History, 1975; R. B. Light, D. A. Roberts and J. D. Stewart (eds) *Museum Documentation Systems: developments and applications*, London, 1986; Parry, R. (ed.). *Museums in a Digital Age*, Leicester Readers in Museum Studies, Routledge, 2007.

²⁰ Lynch, C. A. (2002). Digital collections, digital libraries and the digitization of cultural heritage information. *First Monday*, 7(5).

²¹ Manovich, L. *The Language of New Media*. Cambridge, MA: MIT Press, 2001; Deuze, Mark (2006) 'Participation, Remediation, Bricolage: Considering Principal Components of a Digital Culture', *The Information Society* 22(2): 63–75; Cameron, F. (2003). *Digital Futures I: Museum col-*

внимание как способ осуществлять экспериментальные работы в области оцифровки коллекций²². Многочисленные работы, рассматривающие использование цифровых коллекций, касаются как использования изображений общей публикой, так и профессиональными пользователями²³. Важность открытых лицензий для повторного использования изображений в творческой деятельности и для научных исследований подчеркивает ряд исследователей, которые поддерживают подход, связанный с открытостью коллекций и распространением изображений в случае, когда произведения культуры и искусства перешли в общественное достояние²⁴.

lections, digital technologies, and the cultural construction of knowledge. *Curator*, 46, 325-340; Mak, Bonnie. 2014. "Archaeology of a Digitization." *Journal of the Association for Information Science and Technology* 65.8 (2014): 1515–26; Bode K. (2020). Why you can't model away bias, *Modern Language Quarterly* 81.1.

²² D. Koller, B. Frischer, and G. Humphreys, "Research challenges for digital archives of 3D cultural heritage models", in *ACM Journal on Computing and Cultural Heritage*, Vol. 2 Iss. 3, Dec. 2009; Chabries D. M., Booras S. W., Bearman G. H. (2003). Imaging the past: recent applications of multi-spectral imaging technology to deciphering manuscripts. *Antiquity*, 77(296): 359–72; Terras M. (2006). *Image to Interpretation: Intelligent Systems to Aid Historians in the Reading of the Vindolanda Texts*. Oxford Studies in Ancient Documents, Oxford University Press, Oxford; Tanner S., Bearman G. (2009). *Digitising the Dead Sea Scrolls: Archiving 2009*. Arlington, VA: The Society for Imaging Science and Technology, pp. 119–23.

²³ Walsh, D., Hall, M.M., Clough, P., Foster, J.: Characterising online museum users: a study of the national museums liverpool museum website. *International Journal on Digital Libraries* (Jul 2018); Elena Villaespesa (2019) *Museum Collections and Online Users: Development of a Segmentation Model for the Metropolitan Museum of Art*, *Visitor Studies*, 22:2, 233-252; Lopatovska, I., Bierlein, I., Lember, H. & Meyer, E. "Exploring requirements for online art collections." *Proceedings of the American Society for Information Science and Technology*, 50(1), pp.1–4.

²⁴ Kapsalis E. (2020). 21st-Century Diffusion with Smithsonian Open Access. *Smithsonian*, 25 February 2020. URL: <https://www.si.edu/openaccess/updates/21st-century-diffusion>; Kelly, K. (2013) *Images of Works of Art in Museum Collections: The Experience of Open Access*. <https://clir.wordpress.clir.org/wp-content/uploads/sites/6/pub157.pdf>; Terras M. (2015). *Opening Ac-*

Исследования, связанные с агрегированием цифрового контента, уделяют особое внимание стандартизации метаданных и аннотаций с целью совместимости коллекций и соединения данных из разрозненных коллекций в единую систему, объединенную связанными данными²⁵.

Критический анализ цифровых коллекций и агрегаторов визуальных и текстовых данных в области культурного наследия рассматривает цифровые коллекции и агрегаторы как первичные источники данных, которые формируют представление о культуре географических регионов, стран, культурных сообществ и иногда всего мира²⁶. В недавней работе К. Боуд уделяет большое внимание проблеме репрезентативности и сбалансированности цифровых коллекций и агрегаторов, указывая, что тенденциозность и искажение результатов, полученных при анализе больших данных, являются неизбежным следствием тенденциозно-

cess to collections: the making and using of open digitised cultural content. *Online Information Review*, 39(5): 733 – 752.

²⁵ Freire, N.; Voorburg, R.; Cornelissen, R.; de Valk, S.; Meijers, E.; Isaac, A. Aggregation of Linked Data in the Cultural Heritage Domain: A Case Study in the Europeana Network. *Information* 2019, 10, 252; Julia Marden, Carolyn Li-Madeo, Noreen Whysel, and Jeffrey Edelstein. 2013. Linked Open Data for cultural heritage: Evolution of an information technology. In *Proceedings of the 31st ACM International Conference on Design of Communication*. ACM, New York, NY, 107–112; Daquino, M.; Mambelli, F.; Peroni, S.; Tomasi, F.; Vitali, F. Enhancing semantic expressivity in the cultural heritage domain: Exposing the Zeri Photo Archive as Linked Open Data. *J. Comput. Cult. Herit.* 2017, 10, 21–42; de Boer, V., Wielemaker, J., van Gent, J., Hildebrand, M., Isaac, A., van Ossenbruggen, J., Schreiber, G.: Supporting linked data production for cultural heritage institutes: The amsterdam museum case study. In: *The Semantic Web: Research and Applications*, pp. 733–747. Springer (2012).

²⁶ Bode, Katherine (2018). *A World of Fiction*. en. Ann Arbor, MI: University of Michigan Press; Bode, K. (2020) Why you can't model away bias. *Modern Language Quarterly* (2020) 81 (1): 95–124; Mak, Bonnie. 2014. "Archaeology of a Digitization." *Journal of the Association for Information Science and Technology* 65.8 (2014): 1515–26; Fraistat, N., Jones, S.E., Stahmer, C. (1998). "The Canon, The Web, and the Digitization of Romanticism." *Romanticism On the Net*, 10. URL: <http://id.erudit.org/iderudit/005801ar>; Hyvonen, E.: Using the Semantic Web in Digital Humanities: Shift from data publishing to data-analysis and serendipitous knowledge discovery. *Semantic Web – Interoperability, Usability, Applicability* 11(1), 187–193 (2020).

сти, возникшей при отборе предметов культурного наследия для физических коллекций, а также при ремедиации данных, переводе их в цифровую форму и отборе для последующей публикации и распространения²⁷. Таким образом, канон произведений искусства и культуры, сформировавшийся в печатную эпоху, не расширяется за счет произведений, доступных благодаря оцифровке и ремедиации, а фиксируется новой цифровой инфраструктурой²⁸.

Метод анализа количественных параметров, характеризующих состав коллекции в применении к репрезентативности и сбалансированности цифровых коллекций, восходит к первым цифровым коллекциям лингвистических данных, которые являются источниками исследований в корпусной лингвистике²⁹. К. Боуд применяет метод оценки количественных параметров к коллекциям литературных данных как для определения характеристик литературных процессов, так и для выявления ограничений, связанных с коллекцией текстовых документов как первичного источника литературоведческого исследования³⁰. Метод количественной оценки масштабов, типов и географического распределения цифровых коллекций применяется в проекте ENUMERATE Европейской Комиссии. Задачей проекта является определение размеров оцифрованного контента и масштабов

²⁷ Bode, K. (2020) Why you can't model away bias. *Modern Language Quarterly* (2020) 81 (1): 95–124.

²⁸ Fraistat, N., Jones, S.E., Stahmer, C. (1998). "The Canon, The Web, and the Digitization of Romanticism." *Romanticism On the Net*, 10. URL: <http://id.erudit.org/iderudit/005801ar>; Mak, Bonnie. 2014. "Archaeology of a Digitization." *Journal of the Association for Information Science and Technology* 65.8 (2014): 1515–26

²⁹ Biber D. (1993). Representativeness in Corpus Design. *Literary and Linguistic Computing*, 8 (4): 243-257. Clear J. (1992). Corpus sampling. In Leitner G. (ed.) *New Directions in English Language Corpora*. Berlin: Mouton-de-Gruyter, pp. 21 - 31; Biber D., Conrad S., Reppen R. (1998). *Corpus Linguistics: Investigating Language Structure and Use*. New York: Cambridge University Press.

³⁰ Bode, Katherine. *A World of Fiction*. en. Ann Arbor, MI: University of Michigan Press, 2018;

публикационной активности в области культурного наследия в Европе³¹. Кроме этого, в данном диссертационном исследовании использованы стандартные методы поиска информации, которые дают возможность переходить от одного источника к другому для обнаружения широкого культурологического контекста исследуемых явлений³².

Проблема исследования. Хотя большие коллекции изображений доступны через различные каналы, включая цифровые библиотеки и агрегаторы данных о культурном наследии³³, мы не знаем, в какой мере оцифрованный контент представляет различные культуры, жанры и географические регионы. Это означает, что если в больших агрегаторах культурного контента будет присутствовать большее количество объектов, изображений и историй, связанных с определенной идеей или представлением отдельных знаний, то к этим идеям и концепциям будет легче получить доступ, распространить их и изучить, и они станут основой нового цифрового канона³⁴. «Мифологические, художественные и политические

³¹ Nauta J. G., van den Heuvel W. (2015). Survey report on digitization in European Cultural Institutions 2015. Europeana Pro. URL: <http://pro.europeana.eu/enumerate/statistics/results>

³² Ellis, D. 1989. 'A Behavioural Model for Information Retrieval System Design', *Journal of Information Science*, 15(4–5), 237–247.

³³ Sood, A. (2016). Every piece of art you've ever wanted to see - up, close and searchable. TED lecture. URL: https://www.ted.com/talks/amit_sood_every_piece_of_art_you_ve_ever_wanted_to_see_up_close_and_searchable; Petras, V., Hill, T., Stiller, J. and Gäde, M. (2017), Europeana – a search engine for digitized cultural heritage material, *Datenbank Spektrum*, Vol. 17 No. 1, pp. 41-46; Petras, Vivien, Juliane Stiller. "A Decade of Evaluating Europeana – Constructs, Contexts, Methods & Criteria." In *Research and Advanced Technology for Digital Libraries. TPDFL 2017*. Ed. by Jaap Kamps, Giannis Tsakonas, Yannis Manolopoulos et al. 233–245. Berlin: Springer, 2017.

³⁴ A.E. Earhart. 2012. Can information be unfettered?: Race and the new digital humanities canon. In M. K. Gold, editor, *Debates in the Digital Humanities*, pages 309–318. University of Minnesota Press, Minneapolis, London; Feldman, M.H. (2017). Rethinking the Canon of Near Eastern Art in the Internet Age, *Journal of Ancient Near Eastern History*. 3: 57-79; Zaagsma, G. (2019). Digital History and the Politics of Digitization. Presented at Digital Humanities 2019 Conference, 9-12 July, 2019.

представления о странах»³⁵ или представления о взаимодействии стран будут укрепляться и расширяться на уровне инфраструктур больших данных в области культурного наследия, которые будут влиять на исследования в гуманитарных науках и культурологии в ближайшем будущем. Цифровой культурный колониализм и доминирующие эпистемологии³⁶, которые возникнут в результате этих инфраструктур, могут привести к статистически необъективным «представлениям о знаниях и их политической субъективности»³⁷. Это, в свою очередь, повлияет на подходы машинного обучения, созданные в цифровой информационной среде, а также на развитие культурологических исследований и представления пользователей о культурном наследии.

«Культурный колониализм» цифровых коллекций и тенденциозность, возникшую в результате агрегации или унаследованную от музеев, библиотек и архивов, можно измерить с помощью оценки репрезентативности оцифрованного контента или определяя, насколько пропорционально каждая страна, регион или период времени представлены с помощью оцифрованных изображений. Таким образом, можно отразить социальные и культурологические эффекты, которые возникают вместе с преимуществами и достоинствами цифровых коллекций.

Проблема, которую ставит диссертационное исследование, заключается в возможности и необходимости четкого понимания тенденциозности и неравноправия представления данных, свойственных цифровым платформам, работающим с культурным наследием. Можем ли мы оценить тенденциозность и неравноправие, представленные платформами, публикующими оцифрованный контент в области культурного наследия? Какова возможная методологическая

³⁵ Knell, Simon. (2010). Preface. In Knell, Simon J, Peter Aronsson & Arne Bugge Amundsen (eds), *National Museums: New Studies from Around the World*, London: Routledge, p. xix.

³⁶ Santos, B. De Sousa. *The End of the Cognitive Empire. The Coming of Age of Epistemologies of the South*. Durham: Duke University Press, 2018; Milan, S., & Treré, E. (2019). Big Data from the south(s): Beyond data universalism. *Television and New Media*, 20(4), 319–335.

³⁷ Knell, Simon. (2010). Introduction. In Knell, Simon J, Peter Aronsson & Arne Bugge Amundsen (eds), *National Museums: New Studies from Around the World*, London: Routledge, p. 4.

основа этой оценки? Можно ли применить существующие методы количественного анализа параметров, сопутствующих изображениям оцифрованных произведений культурного наследия для оценки тенденциозности и неравноправия?

Объектом диссертационного исследования являются цифровые технологии, с помощью которых создаются цифровые коллекции, представляющие визуальный контент в области культурного наследия.

Предметом диссертационного исследования выступает культурное наследие и современные механизмы его сохранения, воспроизводства, трансляции.

Цель исследования – выявить складывающиеся культурные закономерности в современных процессах формирования цифровых коллекций и агрегаторов в области культурного наследия.

Достижение поставленной цели связано с реализацией следующих **задач**:

– провести анализ концептуальных и методических подходов к развитию современных цифровых технологий в области культурного наследия, определивших цифровой поворот для всех современных культурных процессов;

– выявить основные способы и формы создания коллекций цифровых изображений и функционирования агрегаторов визуального контента в области культурного наследия;

– сделать исследование представленности российских цифровых коллекций, формирующих современное культурное наследие;

– провести эмпирическое количественное исследование цифровых коллекций Музейного фонда РФ и цифрового агрегатора Google Arts and Culture.

Теоретико-методологическими основаниями для проведения исследования выступают представления о цифровой культуре³⁸, визуальной культуре³⁹

³⁸ Deuze, Mark (2006) 'Participation, Remediation, Bricolage: Considering Principal Components of a Digital Culture', *The Information Society* 22(2): 63–75; Manovich, L. (2001). *The Language of New Media*. Cambridge, MA: MIT Press

³⁹ Schwartz JM (2004) Negotiating the visual turn: new perspectives on images and archives. *Am Arch* 67:107–122; Evans, J., & Hall, S. (Eds.). (1999). *Visual culture: The reader*. London: Sage and

и возможности конструирования культурных концептов благодаря цифровым коллекциям и платформам⁴⁰. Возможность реконструировать мир, представлять его через сочетания новых культурных концептов и с разных точек зрения отмечалась в многочисленных работах конца двадцатого и первых десятилетий двадцать первого веков. Влиятельные работы Б. Андерсона⁴¹, Э. Саида⁴², Л. Мановича⁴³ и М. Дьюза⁴⁴ подчеркивали тенденцию конструирования точек зрения при создании представлений о сообществах и географических регионах и возможности новых медиа в процессе использования изображений, документов и данных в новых контекстах. Расширяя наше понимание культуры, платформы, которые предлагают пользователю цифровой контент, влияют на наше восприятие культуры, поскольку они являются источником опосредованных социальных конструкций⁴⁵. Агрегированный цифровой контент является мощным инструментом создания социальных структур, в которых мы живем⁴⁶. Онлайн-платформы «предлагают определенные нормы и ценности, вписанные в их архитектуру»⁴⁷.

Open University Press; Mirzoeff, N. (1998). What is visual culture? In N. Mirzoeff (Ed.). *The Visual Culture Reader* (pp, 1-13). New York: Routledge; Mirzoeff, N. (1998). What is visual culture? In N. Mirzoeff (Ed.). *The Visual Culture Reader* (pp, 1-13). New York: Routledge.

⁴⁰ Deuze, Mark (2006) 'Participation, Remediation, Bricolage: Considering Principal Components of a Digital Culture', *The Information Society* 22(2): 63–75; Cameron, F. (2003). Digital Futures I: Museum collections, digital technologies, and the cultural construction of knowledge. *Curator*, 46, 325-340.

⁴¹ Anderson B. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. Verso Books, 2006.

⁴² Said E.W. (2000). Invention, memory, and place. *Critical Inquiry*, 26: 175–192.

⁴³ Manovich, L. *The Language of New Media*. Cambridge, MA: MIT Press, 2001.

⁴⁴ Deuze, Mark (2006). Participation, Remediation, Bricolage: Considering Principal Components of a Digital Culture, *The Information Society* 22(2): 63–75.

⁴⁵ Couldry, Nick, and Andreas Hepp. *The Mediated Construction of Reality*. Cambridge: Polity, 2017.

⁴⁶ Van Dijck J, Poell T and Waal M de *The Platform Society*. New York: Oxford University Press, 2018.

⁴⁷ Там же, с. 3.

На отбор и интерпретацию данных цифровых коллекций оказывают влияние социальные представления и конструкты, созданные во времена расцвета империй и колониализма⁴⁸ и усиленные обработкой больших данных⁴⁹. Культурный колониализм онлайн – агрегаторов и тенденциозность, возникшую в результате агрегации или унаследованную от музеев, библиотек и архивов, можно измерить помощью оценки репрезентативности агрегатора или определяя, насколько пропорционально каждая страна или период времени представлены с помощью платформы. Таким образом, возможно отразить социальные и политические эффекты, которые возникают вместе с преимуществами и достоинствами цифровых агрегаторов.

Методологической основой исследования стал подход, который основан на количественных методах оценки репрезентативности и сбалансированности цифровых наборов данных в гуманитарных науках. Понятие репрезентативных и сбалансированных цифровых коллекций применяется к агрегированному визуальному контенту, чтобы продемонстрировать ограничения электронных коллекций данных. При этом в научной литературе цифровые коллекции данных ранее определялись как «тенденциозные с точки зрения представления данных (skewed)»⁵⁰ или «неизбежно нерепрезентативные (inevitably biased)»⁵¹. В научной

⁴⁸ Karp, Ivan, and Steven D. Lavine. (1991). *Exhibiting Cultures: The Poetics and Politics*.

of Museum Display. Washington, DC: Smithsonian Institution Press; Knell, Simon J, Peter Aronsson & Arne Bugge Amundsen (eds), (2010). *National Museums: New Studies from Around the World*, London: Routledge; Boast R (2011) Neocolonial collaboration: museum as contact zone revisited. *Mus Anthropol* 34:56–70.

⁴⁹ Flyverbom M and Murray J (2018) Data structuring – Organizing and curating digital traces into action. *Big Data and Society* 5(2): 1–12.

⁵⁰ Chomsky, Noam. *Syntactic structures*. The Hague: Mouton, 1957.

⁵¹ Clear, J. (1992). Corpus sampling. In G. Leitner (ed.) *New directions in English language corpora*, Mouton-de-Gruyter, Berlin, pp. 21 - 31, см. также Raineri, S. and Debras, C. ‘Corpora and Representativeness: Where to go from now?’ *CogniTextes*. Vol. 2019.

литературе показано, что предвзятость корпуса усугубляется особенностями отбора произведений, которые свойственны тем галереям, музеям, архивам и библиотекам, которые участвуют в составлении корпуса⁵² особенностями, вызванными социальными обстоятельствами⁵³ или особенностями, связанными с самим материалом, представленным в цифровых коллекциях⁵⁴. Действительно, отмечается, что корпус редко бывает репрезентативен. При этом объем корпуса считают самой важной его характеристикой, часто компенсирующей недостаточную репрезентативность и сбалансированность⁵⁵. Однако недавние работы К. Бод в области анализа коллекций литературных данных обращают внимание на необходимость давать оценку тенденциозности, возникающей в связи с недостаточной репрезентативностью и сбалансированностью корпуса⁵⁶. Данное диссертационное исследование применяет методологические подходы, используемые для оценки цифровых данных в текстовых коллекциях к исследованию цифровых коллекций визуальных данных в области культурного наследия.

Основные результаты и научная новизна диссертационного исследования. В диссертационном исследовании получены следующие результаты:

<https://journals.openedition.org/cognitextes/1671>; Bode, K. (2020). 'Why You Can't Model Away Bias,' *Modern Language Quarterly* 81.1.

⁵² Huster, Angela C. (2013). *Assessing Systematic Bias in Museum Collections: A Case Study of Spindle Whorls*. *Advances In Archaeological Practice*. 1(2):77-90.

⁵³ Корчагин К. М. «Поэзия XX века в поэтическом подкорпусе Национального корпуса русского языка: Проблема репрезентативности». *Труды Института русского языка им. В. В. Виноградова* 6 (2015б): 235–256

⁵⁴ Caliskan A., Bryson J. J., Narayanan A. (2017). *Semantics derived automatically from language corpora contain human-like biases*. *Science* 356 (6334): 183–86.

⁵⁵ Egbert J. (2019). *Corpus design and representativeness*. In Sardinha T. B. and Pinto V. (eds.) *Multi-dimensional Analysis. Research Methods and Current Issues*. Bloomsbury Academic, pp. 27-42.

⁵⁶ Bode, K. *A World of Fiction: Digital Collections and the Future of Literary History*. Ann Arbor, MI: University of Michigan Press 2018; Bode K. (2020). *Why you can't model away bias*, *Modern Language Quarterly* 81.1.

1. Проведен аналитический обзор концептуальных и методических подходов к развитию современных цифровых технологий в области культурного наследия. Рассмотрены изменения, определившие цифровой поворот для всех современных культурных процессов. Выявлено изменение методологических подходов, связанных с новыми способами фиксации, анализа и трансляции культурного наследия, связанные с развитием цифровых технологий в области культурного наследия;

2. Выявлены основные способы и формы создания коллекций цифровых изображений и функционирования агрегаторов визуального контента в области культурного наследия. Уточнены ограничения коллекций цифровых изображений и агрегаторов визуального контента при представлении культурного наследия всего мира, отдельных географических регионов и культурных сообществ. Показаны последствия неравномерного представления коллекций цифровых изображений и агрегаторов визуального контента для распространения знаний о культурном наследии;

3. Проведено исследование представленности российских цифровых коллекций, формирующих современное культурное наследие. Рассмотрены масштабы оцифровки коллекций Музейного фонда Российской Федерации. Выявлены цели и складывающиеся культурные закономерности представления и трансляции российских коллекций с помощью цифровых технологий.

4. Проведено эмпирическое количественное исследование цифровых коллекций Музейного фонда РФ и цифрового агрегатора Google Arts and Culture. Показано, что ряд географических регионов и периодов времени представлены в значительно большей степени, нежели другие регионы и сообщества. Проведена интерпретация этого явления в контексте научной литературы, рассматривающей представление культурного наследия в рамках колониальных представлений доцифрового периода. Сделаны выводы о развитии цифрового культурного колониализма в новой информационной среде и с помощью новых медиа.

Основные положения, выносимые на защиту:

1. Анализ концептуальных и методологических подходов к развитию современных цифровых технологий в области культурного наследия позволяет сделать вывод о том, что цифровой поворот для всех современных культурных процессов направлен на создание и распространение нового знания о культурном наследии. Цифровые технологии способствуют увеличению культурного капитала, обеспечивают ремедиацию и реконструирование знания, обнаружение новых контекстов и создание творческих ресурсов. Уточнен спектр социальных, политических и культурологических проблем при дальнейшем развитии цифровых технологий в области культурного наследия и развитии складывающихся культурных закономерностей в современных процессах формирования цифровых коллекций и агрегаторов в области культурного наследия.

2. Анализ основных способов и форм создания коллекций цифровых изображений и функционирования агрегаторов визуального контента в области культурного наследия показал, что цифровая коллекция может неравномерно представить пропорциональное соотношение произведений культуры и искусства. Рассмотрены складывающиеся культурные закономерности в формировании цифровых коллекций в области визуальных произведений культурного наследия. Уточнена связь понятий тенденциозности и репрезентативности коллекций цифровых изображений и агрегаторов визуального контента в области культурного наследия с исследованиями этих понятий в корпусной лингвистике и цифровых литературных исследованиях.

3. Проведенное исследование представленности российских цифровых коллекций, формирующих современное культурное наследие, продемонстрировало неравномерные масштабы формирования представлений о культурном наследии. Показано, что оцифровка музейных коллекций в РФ носит учетный, инвентаризационный характер и не направлена на распространение и создание знаний о коллекциях культурного наследия или создание инфраструктуры для культурных и гуманитарных исследований. Выявлены складывающиеся культурные закономерности и ограничения, связанные с распространением цифровых изобра-

жений произведений из открытого достояния. Продемонстрирован разрыв между ожиданиями пользователей и существующей ситуацией ограничений в распространении знаний о культуре.

4. Культурологическое исследование, в котором был проведен эмпирический количественный анализ цифровых коллекций Музейного фонда РФ и цифрового агрегатора Google Arts and Culture позволяет сделать вывод о недостаточной репрезентативности и сбалансированности агрегатора. Можно утверждать, что складывающиеся культурные закономерности в современных процессах формирования цифровых коллекций и агрегаторов в области культурного наследия демонстрируют культурные предрассудки эпохи печатных изданий, что приводит к цифровому культурному колониализму. Показано, что Google Arts and Culture не является научным изданием и, таким образом, не отвечает требованиям, предъявляемым к инфраструктуре, которая поддерживает культурологические исследования и исследования в области других гуманитарных наук.

Теоретическая и научно-практическая значимость исследования заключается в расширении понимания места цифровых коллекций и агрегаторов в развитии цифровой культуры, определении того, насколько они соответствуют ожиданиям пользователей и способны сделать вклад как в развитие культурного капитала и творческого потенциала общей публики, так и в развитие культурологических и гуманитарных исследований. Представлены концептуальные и методологические подходы к развитию цифровых коллекций как инфраструктуры, способной создать и распространить новое знание. Показаны ограничения цифровых коллекций и агрегаторов, препятствующие использованию изображений в собственных творческих работах и влияющие на результаты анализа больших данных в гуманитарных науках. Раскрыты проблемы и ограничения, связанные с тенденциозностью и неравномерностью представления данных в цифровых коллекциях. Показана возможность применения понятий репрезентативности и сбалансированности, заимствованных из корпусной лингвистики и исследований цифровых коллекций литературных данных, для оценки цифровых коллекций визуальных произведений культуры и искусства. Теоретические положения данной

исследовательской работы могут использоваться при анализе больших данных в гуманитарных науках, составлении учебных программ при разработке курса, касающегося анализа больших данных в культурологии, работах культурологического характера, связанных с формированием культурного капитала и исследований культурной идентичности.

Результаты настоящего диссертационного исследования могут быть использованы при работе над созданием цифровых коллекций и агрегаторов и составлении методических указаний по использованию цифровых материалов в научной и учебной деятельности. Результаты могут применяться в работах, имеющих отношение к теории новых медиа, культурологии, политических и социологических аспектах представления культур в новой информационной среде, культурной антропологии, философии и эпистемологии.

Диссертационная работа может способствовать формированию дальнейших гипотез и проведению исследований в области развития цифровых коллекций и агрегаторов как структур, отражающих способы познания в обществе, демонстрирующих пробелы и закрытые области в эпистемической культуре. Данное исследование может быть полезно для пробуждения интереса к культурам, недостаточно представленным с помощью новых медиа, и популяризации способов привлечения внимания к скрытым (забытым) способам познания, представленным в культурах, которые не входят в новый цифровой канон.

Апробация результатов исследования. Диссертация обсуждена и рекомендована к защите на заседании кафедры культурологии и искусствоведения Гуманитарного института Сибирского федерального университета. Результаты исследования обсуждались в докладах диссертанта в ходе работы конференции Digital Humanities 2016 в Кракове в 2016 году, Digital Humanities 2018 в Монреале в 2018 году, Digital Humanities 2019 в Утрехте в 2019 году, семинаре «Коллекции как данные» в Кракове в 2016 году, семинаре на кафедре информационных исследований Университета Бар – Илан в Израиле в 2019 году, конференции «Информационные технологии в гуманитарных науках» в Красноярске в 2017 году.

Результаты диссертационного исследования опубликованы в виде статей, три из которых опубликованы в рецензируемых научных журналах, рекомендуемых ВАК РФ, таких как журнал Сибирского федерального университета серия «Гуманитарные науки» и «Сибирский антропологический журнал». Одна статья опубликована и одна статья принята в печать в рецензируемом научном журнале «Digital Scholarship in the Humanities». Журнал индексируется базой данных Scopus.

Структура и объем диссертации. Диссертация содержит введение, две главы, четыре параграфа, заключение, список использованной литературы, включающий 266 наименований.

Благодарности. Диссертант выражает благодарность профессору Мелиссе Террас, Университет Эдинбурга, Великобритания, за консультации и помощь в процессе проведения исследования, а также обсуждения и представления результатов.

1 КОНЦЕПТУАЛЬНЫЕ И МЕТОДОЛОГИЧЕСКИЕ ПОДХОДЫ К ЦИФРОВЫМ ТЕХНОЛОГИЯМ КУЛЬТУРНОГО НАСЛЕДОВАНИЯ В СОВРЕМЕННОМ ОБЩЕСТВЕ

1.1 Цифровой поворот в области культурного наследия

В первом параграфе представлен обзор научных исследований развития технологий в области культурного наследования и, как следствие, рост количества цифрового контента, представляющего культурное наследие. С развитием технологии связаны технические особенности цифровых инфраструктур, позволяющие воспринимать их как структуры, создающие и распространяющие знание, которое может быть вписано в иные контексты и объединено с данными из других источников. Первый параграф также представляет научный обзор феномена повторного использования оцифрованных коллекций и применения открытых лицензий в этой области. Эта часть диссертационного исследования затрагивает проблему тенденциозности цифровых коллекций и расширения цифрового канона в области культурного наследования. Обзор научной литературы, представленный в параграфе, показывает, что цифровые коллекции следует воспринимать как первичные источники, которые могут служить предметом культурологического исследования, так как их развитие обусловлено рядом культурных, социальных и политических факторов. Критический анализ подобного рода источников предполагает разработку рекомендаций по их созданию, демонстрирующих ограничения и показывающих онтологические, эпистемические и культурологические проблемы, которые вызывают ограничения и пробелы в коллекциях и агрегаторах. Таким образом, несмотря на ожидания пользователей, которые стремятся увидеть возможность ремедиации и реконструирования цифровых коллекций в собственных исследованиях, интерфейсах и проектах, цифровые инфраструктуры не всегда соответствуют ожиданиям, что влечет за собой необходимость их

усовершенствования и развития для создания и распространения знаний о культурном наследии.

В современном контексте постановка и решение проблемы культурного наследия приводит к появлению исследований в области критического анализа культурного наследия. Ученые, работающие в этом направлении, рассматривают социальный и материальный характер проблем культурного наследования, демонстрируя социологические, экономические, политические и культурологические проблемы современности, которые находят отражение в принципах создания коллекций, процессах анализа и трансляции культурного наследия⁵⁷. В этом же ключе А. Лю, современный теоретик применения новых медиа для фиксации, коллекционирования, анализа и трансляции культурного наследия, утверждает, что исследования культуры в настоящее время тесно связаны с критическим анализом цифровых инфраструктур в области культурного наследия⁵⁸, поскольку именно в них мы можем наблюдать отражение современных социологических, политических и культурных процессов, влияющих на настоящее и будущее.

За последние двадцать лет мы наблюдаем значительный рост количества оцифрованного контента в области культурного наследия. Несмотря на то, что учреждения культуры, работающие с культурным наследием (музеи, архивы,

⁵⁷ Harvey, D. 1990. *The Condition of Postmodernity: An Enquiry into the Origins of Cultural Change*. Cambridge: Blackwell; Harvey David, C. 2001. *Heritage Pasts and Heritage Presents: Temporality, Meaning and the Scope of Heritage Studies*. *International Journal of Heritage Studies* 7 (4): 319–338; Harrison, Rodney 2013 *Heritage: Critical Approaches*. Routledge, Abingdon and New York; Harrison, R., ed. 2010. *Understanding the Politics of Heritage*. Manchester: Manchester University Press, with Open University; Bennett, T, Cameron, F, Dias, N, Dibley, B, Jacknis, I, Harrison R and McCarthy, C 2017 *Collecting, Ordering, Governing: Anthropology, Museums and Liberal Government*. Durham: Duke University Press; Smith, Laurajane (2006) *The Uses of Heritage*. London: Routledge; Smith, Laurajane and Natsuko Akagawa (eds) (2009) *Intangible Heritage*. London: Routledge; Gentry, K., & Smith, L. (2019). *Critical heritage studies and the legacies of the late-twentieth century heritage canon*. *International Journal of Heritage Studies*, 1–21.

⁵⁸ Liu, Alan. *Toward Critical Infrastructure Studies: Digital Humanities, New Media Studies, and the Culture of Infrastructure*. University of Connecticut. 23 February, 2017.

библиотеки), впервые столкнулись с необходимостью инвентаризации материалов с помощью информационных технологий и предмашинной подготовки стандартов описания предметов культурного наследия пятьдесят или шестьдесят лет назад⁵⁹, нарастающая потребность в цифровом контенте и оцифрованных коллекциях начала проявляться на рубеже XX и XXI веков⁶⁰. Огромную часть этих коллекций составляют цифровые изображения⁶¹. Как отмечает Дж. Шварц⁶², рост количества изображений в обращении начался в восемнадцатом веке после изобретения литографии. Этот рост продолжился в геометрической прогрессии в девятнадцатом и двадцатом веке, когда репродукция стала одним из базовых средств коммуникации, и в двадцать первом веке достиг огромных масштабов. Это произошло благодаря росту количества и качества цифровых изображений, увеличению объемов места хранения и сетевых технологий, позволяющих распространение изображений высокого разрешения⁶³ панорамных изображений и трехмерных моделей. Визуальный поворот в культуре, проявившийся во второй половине XX

⁵⁹ Chenhall, R. G. 1975. *Museum cataloging in the computer age*, Nashville, TN: American Association for State and Local History; R. B. Light, D. A. Roberts and J. D. Stewart (eds) *Museum Documentation Systems: developments and applications*, London, 1986; R. Parry. *Recording the Museum: Digital Heritage and the Technologies of Change*. Taylor & Francis, 2007; Parry, R. (ed.) 2007. *Museums in a Digital Age*, Leicester Readers in Museum Studies, Routledge.

⁶⁰ Besser, H., 1997. The transformation of the museum and the way it's perceived. In *The Wired Museum: Emerging Technology and Changing Paradigms*; K. Jones-Garmil, (ed.). American Association of Museums: Washington, D.C., 153-170; Hughes, L. *Digitizing collections: strategic issues for the information manager*, London: Facet Publishing, 2004.

⁶¹ Например, цифровая библиотека “Европеана” включает на момент написания диссертационного исследования около 28 миллионов изображений или около половины из 58 миллионов учетных записей в коллекциях “Европеаны”.

⁶² Schwartz, J.M. (2004) *Negotiating the visual turn: new perspectives on images and archives*. *Am Arch* 67:107–122.

⁶³ Lynch, C. A. (2002). *Digital collections, digital libraries and the digitization of cultural heritage information*. *First Monday*, 7(5).

века⁶⁴, привел к широкому распространению изображений. Это явление стало особенно важным для культуры в антропологическом смысле, как социальной сети, в которой используются определенные смыслы, образы и знаки, а не в смысле «высокой культуры», описанной М. Арнольдом⁶⁵, в то время как представление о культуре в антропологическом смысле восходит к книге «Первобытная культура» Э.Б. Тайлора⁶⁶. Н. Мирзоев утверждает, что культурология и исследования визуальной культуры делают больший акцент на культуре в антропологическом смысле, исследуя спектр проявлений смыслов, образов и систем знаков⁶⁷. Обращаясь к известной работе Л. Мановича⁶⁸, М. Дьюз показывает, что цифровая культура начала XXI века связана с общностью норм, ценностей, видов деятельности и ожиданий людей, объединенных этим видом культуры⁶⁹. Ремедиация визуальной культуры второй половины XX века в цифровой культуре становится важным процессом, в результате которого фиксируются те проявления спектра культур в антропологическом смысле, которые раньше оставались незамеченными в тени высоких культур фиксированного канона. Это явление связано с интересом к массовой культуре и теориям культурного релятивизма в конце XX века⁷⁰. М. Дьюз считает, что сочетание и повсеместное взаимодействие старых и новых медиа, а также индивидуализация творчества и проявлений культуры, выраженная в желании постоянного реконструирования цифрового контента и реме-

⁶⁴ Evans, J., & Hall, S. (Eds.). *Visual culture: The reader*. London: Sage and Open University Press, 1999; Mirzoeff, N. (1998). What is visual culture? In N. Mirzoeff (Ed.). *The Visual Culture Reader* (pp, 1-13). New York: Routledge.

⁶⁵ Arnold, M. *Culture and Anarchy*, Cambridge: Cambridge University Press, 1969.

⁶⁶ Tylor, E. B. *Primitive culture*. New York: Harper & Row, 1958. (Original work published 1871).

⁶⁷ Mirzoeff, N. What is visual culture? In N. Mirzoeff (Ed.). *The Visual Culture Reader* (pp, 1-13). New York: Routledge, 1998.

⁶⁸ Manovich, L. *The Language of New Media*. Cambridge, MA: MIT Press, 2001.

⁶⁹ Deuze, Mark (2006) 'Participation, Remediation, Bricolage: Considering Principal Components of a Digital Culture', *The Information Society* 22(2): 63–75.

⁷⁰ Elkins, J. *The Domain of Images*. Ithaca, NY: Cornell University Press, 1999.

диации ресурсов, созданных в доцифровую эпоху, являются особенностями новой цифровой культуры⁷¹. В рамках этой культуры концепты на любом уровне – географическом, историческом или иконографическом – воспринимаются как культурная среда, поддержанная связанными данными с возможностью исследовать концепт на более поверхностном или более глубоком уровне⁷². Конструирование культурных концептов зависит от объема данных, предложенных цифровыми ресурсами или агрегаторами и от того, какие данные предлагаются для исследования. При этом отбор данных проходит на уровне создания физических коллекций, ремедиации коллекций на этапе оцифровки предметов культуры и искусства, далее на уровне публикации данных, включения их в агрегаторы и возможности повторного использования и публикации на других платформах⁷³. Цифровая культура создает особые модели мира и взаимодействия с миром⁷⁴. Эти модели оказывают дополнительное влияние на пользователя, особенно в условиях, когда рекомендательные системы, основанные на обнаружении самых частых моделей и последующей настройке системы, усиливают этот эффект⁷⁵. В результате конструирование культурных концептов оказывается в руках тех, кто агрегирует, публикует, распространяет, связывает данные и лицензирует их для повторного использования. Таким образом цифровые коллекции в области культурного наследия становятся не только источником данных для исследования культуры, но и сами способы создания, агрегирования и распространения культурного контента становятся предметом культурологических исследований. Этот подход включен в направ-

⁷¹ Deuze, Mark (2006) 'Participation, Remediation, Bricolage: Considering Principal Components of a Digital Culture', *The Information Society* 22(2): 63–75.

⁷² Cameron, F. (2003). Digital Futures I: Museum collections, digital technologies, and the cultural construction of knowledge. *Curator*, 46, 325-340.

⁷³ Bode K. (2020). Why you can't model away bias, *Modern Language Quarterly* 81.1.

⁷⁴ Manovich, L. *The Language of New Media*. Cambridge, MA: MIT Press, 2001.

⁷⁵ Flyverbom M., Murray J. (2018). Datastructuring – Organizing and curating digital traces into action, *Big Data and Society* 5(2): 1–12; Zuboff S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: Profile Books.

ление исследований, которые рассматривают коллекции оцифрованных документов, опубликованных в цифровой среде как способ передачи знаний. Такой подход опирается на достижения палеографии, истории книги, искусствоведческие и культурологические подходы, предполагая, что новые цифровые платформы являются частью процесса создания инфраструктуры знания⁷⁶. В рамках этого подхода цифровые коллекции воспринимаются как первичные источники, которые документируют особенности производства и распространения знания в конкретную историческую эпоху.

Ремедиация визуальной культуры и перевод коллекций в цифровую форму

Таким образом цифровые коллекции музеев, библиотек и архивов становятся крайне необходимы для культурологических исследований и других исследований в области гуманитарных наук как источник визуальных данных⁷⁷ для разных типов пользователей⁷⁸, для исследования разных типов культур и для понимания и анализа особенностей разных инфраструктур создания, распространения и потребления знания. Важность коллекций обусловлена разнообразием источников, культур и способов ремедиации, которые они представляют. Быстрое увеличение количества цифровых изображений в архивах, музеях и библиотеках приводит к небывалым объемам данных: два датасета исследования, в которых источником данных служили изображения из голландских исторических газет, вместе составляют почти миллион изображений⁷⁹, коллекция изображений цифровой

⁷⁶ Mak, B. 2014. "Archaeology of a Digitization." *Journal of the Association for Information Science and Technology* 65.8 (2014): 1515–26

⁷⁷ Allen L., Frost H., Padilla Th., Potvin S., Roke E.R., Varner S. (2019). *The Collections as Data Framework: A Review from the Always Already Computational Project*. <https://tdl-ir.tdl.org/handle/2249.1/156364>

⁷⁸ Crane, G., & Wulfman, C. (2003). Towards a cultural heritage digital library. In L. Decambre, G. Henry, & C.C. Marshall (Eds.), *Proceedings of the 3th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp.75–86). New York, NY: ACM Press.

⁷⁹ Wevers, M., Smits, T. (2019). The visual digital turn: Using neural networks to study historical images. *Digital Scholarship in the Humanities*.

библиотеки «Европеана» составляет двадцать восемь миллионов учетных записей, а Государственный каталог Музейного фонда Российской Федерации⁸⁰ к настоящему времени опубликовал почти восемнадцать миллионов изображений или треть основного фонда российских музейных коллекций. Эти изображения появились в результате процессов цифровизации коллекций культурного наследия или оцифровки изображений. В разных контекстах оцифровке даются разные определения, но обычно под этим процессом понимают создание цифровых копий предметов коллекций культурного наследия⁸¹. Ф. Камерон утверждает, что помимо информационной ценности коллекций цифровых материалов, они обладают еще и социальной ценностью, так как способствуют демократизации процесса взаимодействия с коллекциями, давая возможность доступа, анализа и заинтересованного отношения. В этом случае пользователи становятся «автономными участниками, способными к действию в глобальной инфосфере»⁸². Идея Ф. Камерона заключается в том, что объекты цифрового культурного наследия отличаются от материальных предметов, и обоснована тем, что у них иной, компьютерный тип эстетики и они встроены в иные политические концепты и структуры. Несомненно, решения, которые принимаются по поводу массовой оцифровки коллекций в области культурного наследия, обусловлены разными культурными контекстами и влекут за собой ряд политических, юридических и культурологических осложнений и противоречий⁸³.

⁸⁰ URL: <https://goskatalog.ru/portal/#/collections>

⁸¹ Cameron, F. R. (2019). 'Theorising heritage collection digitisations in global computational infrastructures' In Hannah Lewi, Wally Smith, Dirk vom Lehn, Steven Cooke (eds) *The Routledge International Handbook of New Digital Practices in Galleries, Libraries, Archives, Museums and Heritage Sites*, Routledge.

⁸² Cameron, F. R. (2019). 'Theorising heritage collection digitisations in global computational infrastructures' In Hannah Lewi, Wally Smith, Dirk vom Lehn, Steven Cooke (eds) *The Routledge International Handbook of New Digital Practices in Galleries, Libraries, Archives, Museums and Heritage Sites*, Routledge.

⁸³ Thylstrup, N.B. *The politics of mass digitization*. The MIT Press, Cambridge, 2018.

Рутинные процессы оцифровки

Рутинный процесс оцифровки хорошо описан в работе Б. Тьерс, М. Тюлиг и К. Уотсон, связанной с естественнонаучными коллекциями, но имеющей аналогии и при оцифровке предметов культурного наследия⁸⁴. Технически оцифровка чаще всего осуществляется с помощью сканера или цифровой камеры (иногда больших установок для фотосъемки, снабженных дополнительным студийным освещением). В студиях, оснащенных соответствующим оборудованием, съемка проводится со скоростью 50-100 изображений в час. Процесс рутинной оцифровки настолько прост, что с ним справляются студенты, стажеры и волонтеры. Изображения хранятся на серверах, объемы хранения которых постоянно растут с изменением технологий и возможностей хранения. Данные об оцифрованных изображениях, которые вводятся в систему управления коллекциями, чаще всего создаются вручную, но иногда используются технологии оптического распознавания символов, особенно для стандартизованных печатных этикеток, разработанных до процесса оцифровки в печатную эпоху. Оцифровке в первую очередь подвергаются те объекты, которых нет в других учреждениях культуры, те, которые хорошо атрибутированы, и те, которые были включены в публикации учреждений. Однако более сложные культурные объекты, такие как клинописные таблички⁸⁵, наскальное искусство⁸⁶ или археологические объекты⁸⁷, могут подвер-

⁸⁴ Thiers, B.M., Tulig M.C., Watson K.A. Digitization of the new york botanical garden herbarium. *Brittonia*. 2016; 68(3):324–33.

⁸⁵ Willems, G., Verbiest, F., Moreau, W., Hameeuw, H., Van Lerberghe, K., Van Gool, L., 2005. Easy and cost-effective cuneiform digitizing. In *Short and Project Papers Proceedings of 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST2005)*, Mudge M., Ryan N., Scopigno R., (Eds.), Eurographics Association, pp. 73-80.

⁸⁶ Mudge, M., Malzbender, T., Schroer, C., Lum, M., 2006. New reflection transformation imaging methods for rock art and multiple-viewpoint display. In: Ioannides, M., Arnold, D., Niccolucci, F. (Eds.), *Proceedings of the 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST2006)*. Eurographics Association, pp. 195-200.

гаться оцифровке с использованием техник трехмерного моделирования, такими как лазерное сканирование, фотограмметрия и другие сложные техники⁸⁸. Предполагается, что в этом случае трехмерная модель объекта культурного наследия имеет такую же информационную ценность, как и реальный предмет. К сожалению, это почти всегда невозможно, учитывая характер создания цифрового изображения и технические ограничения, поэтому трехмерная модель чаще всего имеет ценность в рамках конкретной, хорошо определенной цели, проекта или исторического контекста⁸⁹. Для получения дополнительной информации и изучения исторических документов, которые выходят за пределы техник, доступных при рутинной оцифровке, используются более сложные и дорогостоящие технологии⁹⁰. Для сложной обработки изображений, исследования слоев изображения, которые предшествуют верхнему слою, а следовательно, реставрации произведений искусства и исторических документов, применяется спектральная (мульти-

⁸⁷ De Reu, J., Plets, G., Verhoeven, G., De Smedt, P., Bats, M., Cherretté, B., De Maeyer, W., Deconynck, J., Herremans, D., Laloo, P., Van Meirvenne, M., De Clercq, W., 2013. Towards a three-dimensional cost-effective registration of the archaeological heritage. *J. Archaeol. Sci.* 40 (2), 1108-1121.

⁸⁸ D. Koller, B. Frischer, and G. Humphreys, "Research challenges for digital archives of 3D cultural heritage models", in *ACM Journal on Computing and Cultural Heritage*, Vol. 2 Iss. 3, Dec. 2009.

⁸⁹ Hindmarch, J., Terras, M., Robson, S. (2020). "On virtual auras. The cultural heritage object in the age of 3D digital reproduction". In Hannah Lewi, Wally Smith, Dirk vom Lehn, Steven Cooke (eds) *The Routledge International Handbook of New Digital Practices in Galleries, Libraries, Archives, Museums and Heritage Sites*, Routledge.

⁹⁰ Chabries D. M., Booras S. W., Bearman G. H. (2003). Imaging the past: recent applications of multispectral imaging technology to deciphering manuscripts. *Antiquity*, 77(296): 359–72; Terras M. (2006). *Image to Interpretation: Intelligent Systems to Aid Historians in the Reading of the Vindolanda Texts*. Oxford Studies in Ancient Documents, Oxford University Press, Oxford; Tanner S., Bearman G. (2009). *Digitising the Dead Sea Scrolls: Archiving 2009*. Arlington, VA: The Society for Imaging Science and Technology, pp. 119–23.

пектральная или гиперспектральная) съемка⁹¹. Спектральная съемка и последующая обработка изображений применяются и при восстановлении поврежденных исторических документов⁹². Для создания контекстной информации и усиления эффекта при посещении археологических памятников создаются приложения виртуальной реальности. Приложения визуализируют реалистичную модель трехмерной среды на основе цифровых изображений, которой можно управлять либо физическими движениями пользователя, либо действиями, сходными с действиями во время компьютерной игры. Для создания приложения используется цифровая фотограмметрия или лазерное сканирование для получения нужного количества изображений. После нанесения текстуры данные импортируются в игровую среду (достаточно часто в игровую среду Unity⁹³), где они обрабатываются и подготавливаются к визуализации. В игровой среде также происходит программирование навигации и взаимодействия пользователя со средой модели⁹⁴. Для разработки мобильных приложений виртуальной реальности требуется зна-

⁹¹ Liang H: Advances in multispectral and hyperspectral imaging for archaeology and art conservation. *Appl Phys A Mater Sci Process.* 2012, 106: 309-323; S. Legrand, F. Vanmeert, G.V.d. Snickt, M. Alfeld, W.D. Nolf, J. Dik, K. Janssens Examination of historical paintings by state-of-the-art hyperspectral imaging methods: from scanning infra-red spectroscopy to computed X-ray laminography *Herit. Sci.*, 2 (2014), pp. 13-24.

⁹² Giacometti A., Campagnolo A., MacDonald L., Mahony S., Robson S., Weyrich T., Terras M., Gibson A. The value of critical destruction: Evaluating multispectral image processing methods for the analysis of primary historical texts. *Journal of Digital Scholarship in the Humanities*, Volume 32, Issue 1, April 2017, Pages 101–122; Миклас Х., Бреннер С., Саблатниг Р. — Мультиспектральная съемка для цифровой реставрации древних рукописей: устройства, методы и практические аспекты // *Историческая информатика.* – 2017. – № 3. – С. 116 - 134.

⁹³ <https://unity.com/ru>

⁹⁴ Kersten, T. P., Tschirschwitz, F., Deggim, S., & Lindstaedt, M. (2018). Virtual reality for cultural heritage monuments – from 3D data recording to immersive visualisation. In *Euro-Mediterranean Conference* (pp. 74–83). Springer, Cham; Пиков Н.О., Рудов И.Н., Ковалев А.А., Мельников В.Л., Перцева М.А. (2018). “Время разворачивать камни” // *Виртуальная археология (с воздуха, на земле, под водой и в музее)*, Издательство "Государственный Эрмитаж", с. 179-189.

чительно большее вложение времени и трудовых ресурсов, чем в случае создания трехмерных моделей или рутинной оцифровки.

Аннотирование изображений

Традиционные подходы к оцифровке включают не только создание цифровых копий, но и аннотирование изображений или работу по созданию метаданных⁹⁵. В литературе часто подчеркивается разнообразие стандартов метаданных в области описания предметов культурного наследия. Эта проблема приводит к тому, что данные сложно обнаружить, просмотреть, обеспечить их совместимость с другими данными и повторное использование в других проектах⁹⁶. Последние два условия предполагают, что метаданные представлены в открытом некоммерческом формате, что они представлены в машиночитаемой форме и их сопровождают лицензии, которые разрешают повторное использование метаданных. Можно выделить несколько подходов к аннотированию изображений или работе с метаданными. Аннотирование изображений или работа с метаданными могут быть стандартизованы и сведены к минимальному количеству информации с обязательным включением идентификатора⁹⁷. Альтернативный недавний подход сводится к расширению кратких аннотаций, созданных на первом этапе оцифровки, с помощью автоматизированного аннотирования, основанного на машинном обучении, а также участия специалистов и широкой публики в аннотировании

⁹⁵ Blagoderov, V., I. Kitching, L. Livermore, T. Simonsen, and V. Smith. 2012. No specimen left behind: industrial scale digitization of natural history collections. *ZooKeys* 209: 133–146.

⁹⁶ Poirier L., Fortun K., Costelloe-Kuehn B., Fortun M. (2020) Metadata, Digital Infrastructure, and the Data Ideologies of Cultural Anthropology. In: Crowder J., Fortun M., Besara R., Poirier L. (eds) *Anthropological Data in the Digital Age*. Palgrave Macmillan, Cham; Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., et al. (2016). The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Scientific Data*, 3(160018), 1–9.

⁹⁷ Blagoderov, V., I. Kitching, L. Livermore, T. Simonsen, and V. Smith. 2012. No specimen left behind: industrial scale digitization of natural history collections. *ZooKeys* 209: 133–146.

коллекций⁹⁸. Однако участие публики и привлечение волонтеров к аннотированию коллекций и переводу рукописного текста исторических документов в машиночитаемую форму является трудозатратным процессом, который требует много времени, хотя и порождает высококачественные результаты, необходимые для анализа явлений культуры⁹⁹. Поэтому наряду с оптическим распознаванием текста, техники, которая давно и устойчиво применяется при работе с культурным наследием¹⁰⁰, но связана с многочисленными ошибками и искажением смысла в процессе оцифровки и ремедиации¹⁰¹, часто используется автоматизированное распознавание рукописного текста, основанное на техниках машинного обучения¹⁰² и технике восстановления пропущенной информации на основе распознавания закономерностей в сходных объектах¹⁰³.

⁹⁸ Hedrick B, Heberling M, Meinecke E, Turner K, Grassa C, Park D, Kennedy J, Clarke J, Cook J, Blackburn D, Edwards S, Davis C (2019) Digitization and the future of natural history collections. *BioScience*, Volume 70, Issue 3, March 2020, Pages 243–251.

⁹⁹ . Weber, A., Ameryan, M., Wolstencorft, K., Stork, L., Heerlien, M., Schomaker, L.: Towards a digital infrastructure for illustrated handwritten archives. In: Loannides, M. (ed.) *Digital Cultural Heritage. Information Systems and Applications*, incl. Internet/Web, and HCI, vol. 10605, pp. 155–166. Springer International Publishing (April 2018).

¹⁰⁰ Govindan, V.K. and Shivaprasad, A.P. (1990), “Character recognition – a review”, *Pattern Recognition*, Vol. 23 No. 7, pp. 671-683; Ul-Hasan, A., Bukhari, S.S. and Dengel, A. (2016), “OCRoRACT: a sequence learning OCR system trained on isolated characters”, 12th IAPR Workshop on Document Analysis Systems (DAS), IEEE, pp. 174-79.

¹⁰¹ Jarlbrink, Johan, and Pelle Snickars. (2017). Cultural Heritage as Digital Noise: Nineteenth Century Newspapers in the Digital Archive. *Journal of Documentation* 73 (6): 1228–43.

¹⁰² Terras, M. (2005). Reading the Readers: Modelling Complex Humanities Processes to Build Cognitive Systems. *Literary and Linguistic Computing*, 20(1). pp. 41-59; Muehlberger, G., Seaward, L., Terras, M., Ares Oliveira, S., Bosch, V., Bryan, M., Colutto, S., Déjean, H., Diem, M., Fiel, S., Gatos, B., Greinoecker, A., Grüning, T., Hackl, G., Haukkovaara, V., Heyer, G., Hirvonen, L., Hodel, T., Jokinen, M., Kahle, P., Kallio, M., Kaplan, F., Kleber, F., Labahn, R., Lang, E., Laube, S., Leifert, G., Louloudis, G., McNicholl, R., Meunier, J., Michael, J., Mühlbauer, E., Philipp, N., Pratikakis, I., Puigcerver Pérez, J., Putz, H., Retsinas, G., Romero, V., Sablatnig, R., Sánchez, J.,

Использование цифровых коллекций

Использование цифровых коллекций имеет тесную связь с ожиданиями пользователей, функционирующих в рамках цифровой культуры. С ростом ожиданий пользователей, которые обращаются к цифровым ресурсам учреждений культуры, чтобы получить источник развлечения или вдохновения для собственной творческой деятельности, растет количество посещений тех ресурсов, которые отвечают этим ожиданиям. В литературе показана значительная стратификация пользователей цифровых ресурсов в области культурного наследия и подчеркивается, что свыше 70% посетителей пользуется цифровыми коллекциями музеев не в учебных или образовательных целях, а в целях развлечения или в поисках вдохновения для создания собственных произведений искусства¹⁰⁴. Рост количества посещений цифровых коллекций¹⁰⁵ может свидетельствовать о дальнейшей стратификации пользователей и увеличении количества пользователей, которые обращаются к ресурсам не в профессиональных целях. В литера-

Schofield, P., Sfikas, G., Sieber, C., Stamatopoulos, N., Strauß, T., Terbul, T., Toselli, A., Ulreich, B., Villegas, M., Vidal, E., Walcher, J., Weidemann, M., Wurster, H. and Zagoris, K. (2019), Transforming scholarship in the archives through handwritten text recognition: Transkribus as a case study, *Journal of Documentation*, Vol. 75 No. 5, pp. 954-976.

¹⁰³ Lendemer, J., Thiers, B., Monfils, A.K., Zaspel, J., Ellwood, E.R., Bentley, A., LeVan, K., Bates, J., Jennings, D., Contreras, D., Lagomarsino, L., Mabee, P., Ford, L.S., Guralnick, R., Gropp, R.E., Revelez, M., Cobb, N., Seltmann, K. & Aime, M.C. (2020) The extended specimen network: a strategy to enhance US biodiversity collections, promote research and education. *BioScience*, Volume 70, Issue 1, Pages 23–30.

¹⁰⁴ Walsh, D., Hall, M.M., Clough, P., Foster, J.: Characterising online museum users: a study of the national museums liverpool museum website. *International Journal on Digital Libraries* (Jul 2018); Elena Villaespesa (2019) Museum Collections and Online Users: Development of a Segmentation Model for the Metropolitan Museum of Art, *Visitor Studies*, 22:2, 233-252.

¹⁰⁵ Ежегодные отчеты Музея Метрополитан в Нью-Йорке показывают увеличение количества посещений сайта от 700 000 в месяц в 2017 году до 2,5 миллиона посещений в месяц в 2019 году. (подробнее см. <https://www.metmuseum.org/about-the-met/policies-and-documents/annual-reports>)

туре отмечается, что такие пользователи предпочитают функциональность сайта, связанную с возможностью исследовать коллекции случайным образом, не прибегая к строке поиска и не используя специфические поисковые термины, особенно если это касается цифровых коллекций произведений искусства¹⁰⁶. М. Уайтло отмечает важность визуального исследовательского интерфейса с возможностью просмотра коллекций, называя его «щедрым» интерфейсом в отличие от интерфейсов, предоставляющих только поисковую функциональность. М. Уайтло убедительно доказывает важность интерфейса при взаимодействии с пользователем, поскольку интерфейс является важным визуальным элементом медиации и конструирования представлений о коллекции и культурных концептах, которые эта коллекция представляет¹⁰⁷.

Особую категорию пользователей составляют исследователи, которые используют коллекции как источник данных. Процесс использования оцифрованных данных в исследованиях обычно связывают с возникновением цифровых гуманитарных наук¹⁰⁸, нового направления в междисциплинарных исследованиях. Огромное количество стандартизованных данных, которые появляются в учреждениях культуры благодаря процессам оцифровки и имеют отношение к культурному наследию, приводит к тому, что ученые имеют возможность воспользоваться компьютерными средствами анализа. Такие исследования проводятся в обла-

¹⁰⁶ Lopatovska, I., Bierlein, I., Lember, H. & Meyer, E. "Exploring requirements for online art collections." *Proceedings of the American Society for Information Science and Technology*, 50(1), pp.1–4.

¹⁰⁷ Whitelaw, M.: *Generous interfaces for digital cultural collections*. *Digital Humanities Quarterly* 9(1) (2015).

¹⁰⁸ S. Schreibman, R. Siemens, and J. Unsworth, editors. *A companion to Digital Humanities*. Blackwell, 2004; Schreibman, Susan, Ray Siemens, and John Unsworth. 2016. *A New Companion to Digital Humanities*. Chichester: Wiley-Blackwell; Gold, Matthew K., ed. *Debates in the Digital Humanities*. Minneapolis: U of Minnesota P, 2012.

ти компьютерного анализа текста¹⁰⁹, визуализации данных¹¹⁰, соотношения данных культурного наследия с культурными концептами пространства и географического местоположения с помощью географических информационных систем¹¹¹, анализа изображений¹¹², аудио данных¹¹³ и сетевого анализа¹¹⁴. Доступ к данным

¹⁰⁹ Jockers, M.L. *Macroanalysis: Digital Methods & Literary History*. University of Illinois Press, Champaign, Illinois, 2013; Underwood, T. *Distant Horizons: Digital Evidence and Literary Change*. U of Chicago P, 2019.

¹¹⁰ S. Janicke, G. Franzini, M. Cheema, and G. Scheuermann. On Close and "Distant Reading in Digital Humanities: A Survey and Future Challenges. In *Proc. of EuroVis—STARs*, pages 83–103, 2015; Drucker J (2014) *Graphesis: Visual Forms of Knowledge Production*. Cambridge, MA: Harvard University Press.

¹¹¹ Gregory, I., Healey, R. (2007). *Historical GIS: structuring, mapping and analysing geographies of the past*. *Progress in Human Geography*, 31, pp. 638–53; Paterson L.L., Gregory I.N. (2019) *Geographical Information Systems and Textual Sources*. In: *Representations of Poverty and Place*. Palgrave Macmillan, Cham;

¹¹² Besser H (1990) *Visual access to visual images: The UC Berkeley image database project*. *Library Trends* 38(4): 787–798; Jörgensen C (2003) *Image Retrieval: Theory and Research*. Lanham, MD: Scarecrow Press; Melissa Terras. *Digital Images for the Information Professional*. Ashgate, Aldershot, England, 2008; Di Lenardo I., Seguin B., and Kaplan F. (2016). *Visual patterns discovery in large databases of paintings*. *Proceedings of Digital Humanities 2016*, Krakow. <https://dh2016.adho.org/abstracts/348>; Junginger P., Ostendorf D., Vissirini B.A., Voloshina A., Kreiseler S., Dörk M. *Close-Up Cloud: Gaining A Sense Of Overview From Many Details*. Presented at *Digital Humanities 2019 Conference*, 9-12 July, 2019.

¹¹³ Müller, M. *Fundamentals of Music Processing*; Springer International Publishing: Cham, Switzerland, 2015; Gref, Michael, Köhler, Joachim, Leh, Almut. 2018. *Improved transcription and indexing of oral history interviews for digital humanities research*. *International Conference on Language Resources and Evaluation (LREC) 2018*, <http://publica.fraunhofer.de/dokumente/N-494202.html>, 30 Jan 2019; Cord Pagenstecher 2019. *Curating and Analyzing Oral History Collections*. Selected papers from the CLARIN Annual Conference 2018. *Linköping Electronic Conference Proceedings* 159: 144–151.

¹¹⁴ Knappett, C. (Ed.). (2013). *Network analysis in archaeology: New approaches to regional interaction*. Oxford: Oxford University Press; Eder, Maciej. 2015. "Visualization in Stylometry: Cluster

и возможности анализа данных предоставляют разные учреждения культуры, исследовательские центры, библиотеки, музеи и архивы, такие как Исследовательский центр Хатхи, консорциума учебных и научных библиотек университетов США¹¹⁵, Центр предоставления данных Национальной библиотеки Нидерландов¹¹⁶, Британская библиотека¹¹⁷, Музей Метрополитен в Нью-Йорке¹¹⁸, Отдел исторических газет Библиотеки Конгресса¹¹⁹. Несмотря на то, что использование коллекций учреждений в области культурного наследия как источника данных осложнено вопросами авторского права людей и институций, которые участвуют в создании и ремедиации контента, опыт предоставления доступа к данным и разработки компьютерных средств обнаружения данных, их сортировки и анализа становится частой практикой библиотек, музеев и архивов¹²⁰. Д. Синн демонстрирует рост использования цифровых проектов и цифровых данных (в значительной степени данных, опубликованных Библиотекой Конгресса) в исторических исследованиях на примере статей журнала *American Historical Review*, опубликованных с 2001 по 2010 годы¹²¹. Однако цитирование цифровых проектов и источников

Analysis Using Networks.” *Digital Scholarship in the Humanities* 30; Jackson, C., 2017. Using social network analysis to reveal unseen relationship in medieval Scotland. *Digital Scholarship in the Humanities*, 32(2), pp. 336-343; Tamper, M., Hyvönen, E., Leskinen, P.: Visualizing and analyzing networks of named entities in biographical dictionaries for digital humanities research. In: *Proceedings of CICLing 2019*, Springer-Verlag (2019).

¹¹⁵ <https://www.hathitrust.org/htrc>

¹¹⁶ <https://www.kb.nl/en/resources-research-guides/data-services-apis>

¹¹⁷ <https://www.bl.uk/catalogues-and-collections/digital-collections>

¹¹⁸ <https://www.metmuseum.org/blogs/now-at-the-met/2018/met-collection-api>

¹¹⁹ <https://chroniclingamerica.loc.gov/>

¹²⁰ Allen L., Frost H., Padilla Th., Potvin S., Roke E.R., Varner S. (2019). The Collections as Data Framework: A Review from the Always Already Computational Project. <https://tdl-ir.tdl.org/handle/2249.1/156364>

¹²¹ Sinn, D. (2012). Impact of digital archival collections on historical research.

Journal of the American Society for Information Science & Technology, 63(8), 1521–1537.

цифровых данных не является единственным способом взаимодействия с цифровыми коллекциями.

К. Фенлон¹²² описывает более сложные способы научной коммуникации, такие как цифровые академические издания¹²³, где цифровые коллекции являются одним из важных встроенных элементов, или компьютерные симуляции, представляющие исследовательскую среду для исследования и анализа сложных объектов культурного наследия¹²⁴. По утверждению К. Фенлон, процессы и результаты исследовательской работы в гуманитарных науках и изучении культурного наследия «существуют в новой медийной среде, насыщены данными, интерактивны, динамичны, связаны и подвержены постоянным эволюционным изменениям»¹²⁵. Это хорошо согласуется с утверждением П. Марти о том, что информационная инфраструктура соответствует тем институтам и организациям, которые ее создают, но одновременно с этим сама меняет эти институты и организации¹²⁶. Динамичная медийная среда, в которой проходит исследование культурного наследия, сама меняет исследователя и пользователя цифровых коллекций. Таким об-

¹²² Fenlon, K. (2019). "Modeling Digital Humanities Collections as Research Objects", Proceedings of ACM Joint Conference on Digital Libraries.

¹²³ Robinson, P. (2013). "Towards a Theory of Digital Editions." *Variants: The Journal of the European Society for Textual Scholarship* 10: 105–31; Schreibman, S. & C. Papadopoulos, 2019. *Textuality in 3D. Three-Dimensional (re)constructions as digital scholarly editions*. *International Journal of Digital Humanities Special issue on Digital Scholarly Editing*; Flanders, J., Siemens, R. *Considering the scholarly edition in the digital age: an engagement by the modern language association's committee on scholarly editions*. *Int J Digit Humanities* (2019).

¹²⁴ <https://mayaarch3d.org/en/>

¹²⁵ Fenlon, K. (2019). "Modeling Digital Humanities Collections as Research Objects", Proceedings of ACM Joint Conference on Digital Libraries.

¹²⁶ Marty, P. F. (1999). *Museum Informatics and Information Infrastructures: Supporting Collaboration across Intra-Museum Boundaries*. *Archives and Museum Informatics*, 13 1999, 169- 85.

разом создаются новые способы структурировать знание о культуре и мире, в котором функционирует культура¹²⁷.

Исследования, связанные с использованием цифровых коллекций и последующей оптимизацией процессов оцифровки, создания метаданных и публикации данных связаны со значительными трудностями. В первую очередь, трудность представляет включение уникального идентификатора, определяющего объект для систем агрегирования и отслеживания цифровых данных. Такой идентификатор может служить для связи однородных или похожих предметов и разных цифровых коллекций или для интеграции дополнительных данных географического или просопографического характера об объекте. Его следовало бы указывать в научных работах и образовательной деятельности и таким образом получать представление об использовании цифровых коллекций¹²⁸.

Повторное использование оцифрованного контента

Самые ранние работы, изучающие использование цифровых коллекций в области культурного наследия для профессиональной деятельности, восходят к 2008 году, когда были разработаны «Принципы оценки цифровых ресурсов в гуманитарных науках»¹²⁹. Авторы документа использовали библиометрический подход, оценивающий количество упоминаний цифровых ресурсов в научных публикациях. Отдавая должное эффективности библиометрического подхода, ав-

¹²⁷ Parry, R. (2007) Recording the Museum. Digital Heritage and the Technologies of Change. London: Routledge.

¹²⁸ Lendemer, J., Thiers, B., Monfils, A.K., Zaspel, J., Ellwood, E.R., Bentley, A., LeVan, K., Bates, J., Jennings, D., Contreras, D., Lagomarsino, L., Mabee, P., Ford, L.S., Guralnick, R., Gropp, R.E., Revelez, M., Cobb, N., Selmann, K. & Aime, M.C. (2020) The extended specimen network: a strategy to enhance US biodiversity collections, promote research and education. BioScience, Volume 70, Issue 1, Pages 23–30.

¹²⁹ Toolkit for the Impact of Digitized Scholarly Resources, <https://www.oii.ox.ac.uk/research/projects/tidsr/>, цитируется по A. Babeu 2011. Rome wasn't digitized in a day: building a cyberinfrastructure for digital Classics. Council on Library and Information Resources. Washington DC. <https://www.clir.org/wp-content/uploads/sites/13/pub150.pdf>

торы отмечают, что исследователи далеко не всегда ссылаются на цифровую коллекцию, даже если упоминают или используют цифровое изображение предмета или документа, доступ к которому они получили с помощью цифрового ресурса. А. Бабо отмечает, что в связи с этим объем реального использования цифровых коллекций может значительно превышать тот объем, который можно оценить с помощью библиометрических методов¹³⁰.

Цифровые коллекции используются посетителями как источник вдохновения и данных для последующей творческой работы как в профессиональной, так и в непрофессиональной области. Это значит, что пользователям цифровых ресурсов в области культурного наследия часто требуется скопировать и сохранить изображение для повторного использования в собственном проекте. Это особенно важно для исследовательской работы, которая предполагает создание собственных аннотированных коллекций изображений¹³¹, которые впоследствии будут опубликованы. В этом случае важность повторного использования заключается в том, что интерфейс нового проекта будет связан с научными задачами, вопросами и гипотезами нового исследования, которые чаще всего отличаются от задач, которые можно решить с помощью интерфейса исходного цифрового ресурса¹³². Интерес непрофессиональных пользователей, которым необходимы визуальные данные для собственной творческой деятельности, может быть обусловлен поиском культурной и этнической идентичности и онтологической защищенности¹³³. Людям необходимо чувствовать принадлежность к разным формам культуры, аутентичность и авторитет которых поддерживают учреждения культуры и органи-

¹³⁰ Babeu, A. (2011). "Rome wasn't digitized in a day:" building a cyberinfrastructure for digital Classics. Council on Library and Information Resources. Washington DC. <https://www.clir.org/wp-content/uploads/sites/13/pub150.pdf>

¹³¹ Fenlon, K. (2019). "Modeling Digital Humanities Collections as Research Objects", Proceedings of ACM Joint Conference on Digital Libraries.

¹³² Robinson, P. (2013). Five desiderata for scholarly editions in digital form. In Digital Humanities 2013: Conference Abstracts. University of Nebraska-Lincoln, 16-19 July 2013.

¹³³ Turner, B. (2001) 'The Erosion of Citizenship', British Journal of Sociology 52(2): 189–209.

зации, поддерживающие цифровые ресурсы в области культурного наследия, и получить возможность творчества, соединенного с культурным наследием этническими, культурными и онтологическими связями. Более того, интерес к повторному использованию цифровых ресурсов связан с интересом к редкой в доцифровую эпоху способностью по-новому представить мир вокруг нас с помощью инфраструктуры, в которой можно легко распространить варианты репрезентаций¹³⁴. Такая потребность в реконструировании и ремедиации цифровых ресурсов отвечает ожиданиям и нормам цифровой культуры¹³⁵.

Научные исследования, которые пользуются первичными историческими данными, то есть большая часть научных исследований в гуманитарных науках, зависят от открытого доступа к данным в той же мере, в которой зависят от открытого доступа к данным и естественные науки. Это значит, что большие объемы открытых данных, которые хранятся в библиотеках, музеях и архивах, способны изменить то, как ведутся исследования и способствовать развитию компьютерных инструментов для обнаружения закономерностей в данных. Концепция «открытого доступа», которой придерживается ряд учреждений культуры¹³⁶ для того, чтобы предоставить возможность повторного использования изображений из некоторых разделов своих коллекций, позволяет разным категориям пользователей копировать, изменять и распространять изображения, особенно те произведения, право копирования которых находится в общественном достоянии. В этом диссертационном исследовании «открытый доступ» понимается как возможность получить бесплатный и беспрепятственный доступ к изображениям, а также возможность их использовать, изменять и распространять. Это определе-

¹³⁴ Couldry, N. (2008). Mediatization or mediation: alternative understandings of the emergent space of digital storytelling. *New Media & Society* 10(3): 373–391.

¹³⁵ Deuze, M. (2006) 'Participation, Remediation, Bricolage: Considering Principal Components of a Digital Culture', *The Information Society* 22(2): 63–75.

¹³⁶ Kapsalis, E. (2016). *The Impact of Open Access on Galleries, Libraries, Museums, & Archives, Statewide Agricultural Land Use Baseline*, 2015.

ние следует формулировке, разработанной фондом «Открытое знание»¹³⁷ и применяется сообществом «Открытые библиотеки, музеи и архивы»¹³⁸. При этом применение лицензий, которые запрещают коммерческое использование или изменение изображений, делает коллекции закрытыми¹³⁹. Э. Капсалис определяет открытый доступ как возможность беспрепятственного использования материалов, которые находятся в общественном достоянии¹⁴⁰. Однако далеко не весь оцифрованный контент может быть доступен для повторного использования из-за ограничений, связанных с авторским правом и необходимостью лицензирования повторного использования. Право копирования творческого произведения переходит в общественное достояние через несколько десятков лет после смерти автора. Время, которое требуется для того, чтобы право копирования перешло в общественное достояние, регулируется национальными законодательствами. Право копирования творческих произведений считается общественным достоянием и для работ, созданных в то время, когда авторское право еще не существовало. Часто учреждения культуры ограничивают право копирования произведений в общественном достоянии¹⁴¹. Это решение часто обосновывается тем, что использование изображений в коммерческих целях должно оплачиваться, чтобы

¹³⁷ <http://opendefinition.org/>

¹³⁸ <https://medium.com/open-glam>; McCarthy, D., Uncovering the global picture of Open GLAM, Medium, 4 April, 2019, <https://medium.com/open-glam/uncovering-the-global-picture-of-open-glam-af364aadeeee>

¹³⁹ McCarthy, D., Uncovering the global picture of Open GLAM, Medium, 4 April, 2019, <https://medium.com/open-glam/uncovering-the-global-picture-of-open-glam-af364aadeeee>

¹⁴⁰ Kapsalis, E. (2016). The Impact of Open Access on Galleries, Libraries, Museums, & Archives, Statewide Agricultural Land Use Baseline, 2015.

¹⁴¹ Например, в соответствии с правилами, установленными Государственным Эрмитажем, необходимо получить письменное разрешение администрации музея для использования изображений высокого разрешения в коммерческих и научных публикациях. См. https://www.hermitagemuseum.org/wps/portal/hermitage/about/image_usage_policy/

финансировать расходы музеев на оцифровку изображений¹⁴². Однако представление о том, что создание цифровых изображений двумерных произведений искусства влечет за собой создание нового авторского права является спорным вопросом и часто оспаривается рядом комментаторов и организаций США, Великобритании и Европы¹⁴³. Особую проблему представляют произведения культуры и искусства, созданные в двадцатом веке, которые не могут быть включены в программы массовой оцифровки, которые поддерживаются крупными цифровыми коллекциями и агрегаторами. Этот пробел в создании и распространении знаний о культурном наследии часто называют «черной дырой двадцатого века». Публикацию и распространение знаний осложняет долгий процесс поиска правообладателя и бюрократические особенности лицензирования произведения для повторного использования, а в некоторых случаях полное отсутствие информации о правообладателе¹⁴⁴.

На момент написания диссертационного исследования свыше семисот музеев, библиотек и архивов во всем мире предоставляют открытый доступ к некоторым разделам своих коллекций¹⁴⁵, несмотря на многочисленные технические, финансовые, социальные и философские проблемы, с которыми сталкиваются учре-

¹⁴² Pachali, D. (2014). Open Access: How museums are opening their digital archives - Goethe Institut. <https://www.goethe.de/en/kul/bib/20365145.html>

¹⁴³ Intellectual Property Office (2015) 'Copyright Notice : digital images , photographs and the internet'.

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/481194/c-notice-201401.pdf; Petri, G. (2014). The Public Domain vs. the Museum: The Limits of Copyright and Reproductions of Two-Dimensional Works of Art, In *Journal of Conservation and Museum Studies*. 12 (1); Bielsten, S. (2006). *Permissions, A Survival Guide: Blunt Talk About Art as Intellectual Property*. Chicago, University of Chicago Press.

¹⁴⁴ Thylstrup, N.B. *The politics of mass digitization*. The MIT Press, Cambridge, 2018;

¹⁴⁵ McCarthy, D. and Wallace, A. (2019) Survey of GLAM open access policy and practice. https://docs.google.com/spreadsheets/d/1WPS-KJptUJ-o8SXtg00llcxq0IKJu8eO6Ege_GrLaNc/edit#gid=1216556120

ждения культуры при переходе к открытому доступу¹⁴⁶. Одна из причин, по которой ряд музеев, архивов и библиотек предоставляет данные цифровых коллекций для открытого доступа, заключается в том, что грань между коммерческим и некоммерческим использованием очень трудно провести. К тому же, отслеживание коммерческого использования требует значительного вложения финансовых и трудовых ресурсов¹⁴⁷. Те музеи, библиотеки и архивы, которые публикуют разделы цифровых коллекций с правом копирования в общественном достоянии для открытого доступа¹⁴⁸, обычно пользуются лицензиями Creative Commons¹⁴⁹. «Открытые лицензии Creative Commons дают сообществу набор простых инструментов, которые показывают, как можно распространять и повторно использовать данные и позволяют учреждениям культуры донести до максимально возможного количества людей знания о тех предметах, которые хранятся в библиотеках, музеях и архивах»¹⁵⁰. Такие лицензии дают возможность музеям, библиотекам и архивам распоряжаться авторским правом, которое принадлежит учреждению культуры, возникло в результате ремедиации аналогового объекта (цифровая съемка, создание файлов, работа над публикацией изображения и т. д.) и связано с исходным аналоговым объектом (например, цифровое изображение произведения искусства, которое находится в общественном достоянии), несмотря на то, что цифровая фотография произведения искусства может быть защищена авторским правом. Отказ от авторского права для цифровых фотографий произведений и объек-

¹⁴⁶ Kelly, K. (2013) Images of Works of Art in Museum Collections: The Experience of Open Access. <https://clir.wordpress.clir.org/wp-content/uploads/sites/6/pub157.pdf>

¹⁴⁷ Kelly, K. (2013) Images of Works of Art in Museum Collections: The Experience of Open Access. <https://clir.wordpress.clir.org/wp-content/uploads/sites/6/pub157.pdf>

¹⁴⁸ См., например, <https://www.metmuseum.org/about-the-met/policies-and-documents/image-resources>, <https://creativecommons.org/tag/british-library/>, <https://www.si.edu/Termsfuse>

¹⁴⁹ <https://creativecommons.org/>

¹⁵⁰ M. von Haller Groesbaek, (2014). 'GLAMorous remix: Openness and sharing for cultural institutions', in Sanderhoff, M. (ed.) Sharing is caring: Openness and sharing in the cultural heritage sector, pp. 141–153.

тов из общественного достояния (или признание того, что создание цифрового изображения не влечет за собой создание нового авторского права) и использование открытых лицензий становится практикой ряда музеев, архивов и библиотек. Это делается с целью распространения знаний о коллекциях, в том числе и с помощью использования сторонних платформ¹⁵¹.

Рассматривая повторное использование изображений с точки зрения конечного пользователя, Ф. Валеонти, М. Террас и А. Гудсон-Смит¹⁵² выделяют несколько препятствий. Во-первых, за исключением списка, составленного Д. МакКарти и А. Уоллас¹⁵³, не существует источника, агрегирующего информацию об учреждениях культуры, предоставляющего открытый доступ к своим коллекциям. Пользователи не могут определить, какие учреждения культуры дают возможность беспрепятственного доступа к изображениям с любыми целями. Поскольку открытый доступ предполагает в том числе коммерческие цели, это особенно важно для пользователей, которым необходимо реконструировать и ремедирировать изображения для создания нового продукта. Во-вторых, далеко не все изображения, предоставленные учреждениями культуры с открытым доступом, пригодны для повторного использования в образовательных/научных целях и с целью создания коммерческого продукта. Низкое качество изображений (черно-белые изображения или изображения с низким разрешением) приводит к тому, что они не годятся для повторного использования. В-третьих, некоторые музеи, предоставляющие открытый доступ к изображениям, например, Музей изобрази-

¹⁵¹ Kapsalis, E. (2020). 21st-Century Diffusion with Smithsonian Open Access. *Smithsonian*, 25 February 2020.

<https://www.si.edu/openaccess/updates/21st-century-diffusion>

¹⁵² Valeonti, F; Terras, M; Hudson-Smith, A. (2020) How open is OpenGLAM? Identifying barriers to commercial and non-commercial reuse of digitised art images. *Journal of Documentation*, 76 (1) pp. 1-26.

¹⁵³ McCarthy, D. and Wallace, A. (2019) Survey of GLAM open access policy and practice. https://docs.google.com/spreadsheets/d/1WPS-KJptUJ-o8Sxtg00llcxq0IKJu8eO6Ege_GrLaNc/edit#gid=1216556120

тельного искусства в Индианополисе (США)¹⁵⁴, требуют заполнить форму для того, чтобы контролировать публикацию изображений в других источниках и коммерческое использование изображений. Такой подход был, вероятно, оправдан в печатную эпоху, когда пользователю требовалось 50-70 изображений. В ситуации, когда пользователь разрабатывает проект, где необходимы сотни или тысячи изображений, этот подход осложняет открытый доступ. Несмотря на то, что отдельные музеи предоставляют открытый доступ к своим коллекциям с помощью интерфейса прикладного программирования (API), позволяющего осуществить доступ к данным и копирование изображений автоматически¹⁵⁵, повторное использование изображений объектов культурного наследия все еще осложнено вышеуказанными обстоятельствами.

Агрегаторы цифровых данных в области культурного наследия

Масштаб оцифровки предметов культурного наследия и количество данных, которые появляются в результате этого процесса как в виде цифровых изображений, так и в виде метаданных, требуют агрегирования данных для последующего обнаружения и использования в образовательных целях, научных исследованиях, а также с тем, чтобы удовлетворить тематические поисковые запросы широкой публики. Рассматривая проблему агрегирования данных на примере естественно-научных коллекций США, Дж. Лендерман и коллектив сотрудников, работающих над проектом агрегатора естественнонаучных коллекций, указывают на то, что для выполнения проектов такого типа необходимо объединить все источники данных, создать централизованную базу данных, обучить квалифицированных сотрудников и обеспечить долгосрочное финансирование¹⁵⁶. Особый акцент в этом

¹⁵⁴ URL: <https://discovernewfields.org/>

¹⁵⁵ URL: <https://www.programmableweb.com/category/museums/api>

¹⁵⁶ Lendemer, J., Thiers, B., Monfils, A.K., Zaspel, J., Ellwood, E.R., Bentley, A., LeVan, K., Bates, J., Jennings, D., Contreras, D., Lagomarsino, L., Mabee, P., Ford, L.S., Guralnick, R., Gropp, R.E., Revelez, M., Cobb, N., Seltmann, K. & Aime, M.C. (2020) The extended specimen network: a strategy to enhance US biodiversity collections, promote research and education. *BioScience*, Volume 70, Issue 1, Pages 23–30.

процессе делается на агрегировании метаданных (информации о данных, которая описывает цифровые ресурсы), поскольку многие типы данных (изображения, аудио, рукописные тексты) не являются машиночитаемыми и не способствуют использованию исходных цифровых ресурсов с помощью поисковых механизмов¹⁵⁷. Однако цифровые ресурсы, представляющие культурное наследие, хранятся в учреждениях разного типа и в разных странах, что ведет к разнообразию стандартов описания и несовместимости моделей данных, при этом часто каждое учреждение придерживается собственных стандартов и собственных моделей данных¹⁵⁸. Если библиотеки давно придерживаются общих стандартов метаданных, то архивы и музеи начали внедрять подобную практику совсем недавно. Ряд хорошо известных агрегаторов метаданных в области культурного наследия, такие как Europeana Collections¹⁵⁹, DigitalNZ¹⁶⁰, Trove, агрегатор культурного наследия, связанного с Австралией¹⁶¹ или Государственный каталог Музейного фонда Российской Федерации¹⁶² были созданы по инициативе крупных государственных организаций, в то время как агрегатор Google Arts and Culture возник по инициативе влиятельной коммерческой компании¹⁶³. Проблема совместимости данных приводит к тому, что агрегаторы данных о культурном наследии, такие как, например, Европеана, приводят исходные метаданные, полученные из учреждений культуры в разных странах, к единому формату и стандарту. Несмотря на то, что этот про-

¹⁵⁷ Freire, N.; Voorburg, R.; Cornelissen, R.; de Valk, S.; Meijers, E.; Isaac, A. Aggregation of Linked Data in the Cultural Heritage Domain: A Case Study in the Europeana Network. *Information* 2019, 10, 252.

¹⁵⁸ Peacock, D. (2007). The information revolution in museums. In *Museum informatics: People, information, and technology in museums*, Edited by: Marty, P.F. and Burton Jones, K. 59–76. London and New York: Routledge.

¹⁵⁹ URL: <https://www.europeana.eu/en>

¹⁶⁰ URL: <https://digitalnz.org/>

¹⁶¹ URL: <https://trove.nla.gov.au/general/about>

¹⁶² URL: <https://goskatalog.ru/portal/#/>

¹⁶³ URL: <https://artsandculture.google.com/>

цесс обеспечивает последовательное и совместимое представление данных, богатство и разнообразие исходных описаний и аннотаций пропадает в процессе конвертации¹⁶⁴. Несмотря на это, агрегаторы цифровых коллекций в области культурного наследия успешно сочетают растущий объем цифрового контента, простые способы доступа с помощью метаданных и функциональность интерпретации, анализа, образовательных элементов и выставочной деятельности для цифровых изображений произведений искусства и культуры. Такая деятельность, как указывает К. Линч, является исторической миссией многих учреждений культуры и, возможно, окажется в фокусе представителей многих сообществ пользователей за пределами музеев, архивов, библиотек и учреждений высшего образования¹⁶⁵. Еще одна функция агрегатора данных о предметах культурного наследия – обеспечить связь между метаданными и многочисленными справочными материалами, такими как данные о географических названиях, людях и событиях с помощью уникальных идентификаторов, присвоенных предмету, и возможности объединять данные о явлении, событии или персоналии в пространстве связанных данных для того, чтобы установить связи между коллекциями¹⁶⁶.

¹⁶⁴ de Boer, V., Wielemaker, J., van Gent, J., Hildebrand, M., Isaac, A., van Ossenbruggen, J., Schreiber, G.: Supporting linked data production for cultural heritage institutes: The amsterdam museum case study. In: *The Semantic Web: Research and Applications*, pp. 733–747. Springer (2012).

¹⁶⁵ Lynch, C. A. (2002). Digital collections, digital libraries and the digitization of cultural heritage information. *First Monday*, 7(5).

¹⁶⁶ Julia Marden, Carolyn Li-Madeo, Noreen Whysel, and Jeffrey Edelstein. 2013. Linked Open Data for cultural heritage: Evolution of an information technology. In *Proceedings of the 31st ACM International Conference on Design of Communication*. ACM, New York, NY, 107–112; Daquino, M.; Mambelli, F.; Peroni, S.; Tomasi, F.; Vitali, F. Enhancing semantic expressivity in the cultural heritage domain: Exposing the Zeri Photo Archive as Linked Open Data. *J. Comput. Cult. Herit.* 2017, 10, 21–42; de Boer, V., Wielemaker, J., van Gent, J., Hildebrand, M., Isaac, A., van Ossenbruggen, J., Schreiber, G.: Supporting linked data production for cultural heritage institutes: The amsterdam museum case study. In: *The Semantic Web: Research and Applications*, pp. 733–747. Springer (2012).

Таким образом, одна из функций оцифровки – «обеспечить динамические наборы данных взаимосвязанных ресурсов, которые все время обогащаются и дополняются с помощью научных исследований»¹⁶⁷. Однако в отличие от естественнонаучных коллекций с их четкой таксономической индексацией и возможностью интегрировать и совмещать метаданные разных коллекций, коллекции объектов культурного наследия представляют дополнительные трудности из-за сложности структуры метаданных и трудностей совместимости стандартизованных онтологий, словарей и классификаций. Так например, П. Секли с соавторами¹⁶⁸ обнаружил несовместимость фамилий авторов произведений, которые хранятся в Смитсоновском музее американского искусства с данными о персоналиях Википедии при сопоставлении данных музея и структурированных данных об именованных сущностях, которые хранятся в Википедии в виде онтологии¹⁶⁹. Авторы работы обращают внимание на необходимость создания библиотеки онтологий, представляющей руководство в области совместимости существующих стандартов. Помимо задач стандартизации данных, одной из самых сложных задач при работе над созданием агрегатора является задача заполнения необходимых полей там, где данные пропущены. Применяются техники восстановления данных с использованием оптического распознавания текста печатных аннотаций

¹⁶⁷ Lendemer, J., Thiers, B., Monfils, A.K., Zaspel, J., Ellwood, E.R., Bentley, A., LeVan, K., Bates, J., Jennings, D., Contreras, D., Lagomarsino, L., Mabee, P., Ford, L.S., Guralnick, R., Gropp, R.E., Revelez, M., Cobb, N., Seltmann, K. & Aime, M.C. (2020) The extended specimen network: a strategy to enhance US biodiversity collections, promote research and education. *BioScience*, Volume 70, Issue 1, Pages 23–30.

¹⁶⁸ Szekely, P. A., Knoblock, C. A., Yang, F., Zhu, X., Fink, E. E., Allen, R., & Goodlander, G. (2013). Connecting the smithsonian american art museum to the linked data cloud. In P. Cimiano, O. Corcho, V. Presutti, L. Hollink, & S. Rudolph (Eds.), *Proc. of ESWC 2013* (pp. 593–607). Springer volume 7882 of *Lecture Notes in Computer Science*.

¹⁶⁹ <https://wiki.dbpedia.org/about>

и поиск закономерностей в данных для переноса данных в случае совпадения части характеристик объектов¹⁷⁰.

Агрегаторы и тенденциозность, возникающая при ремедиации коллекций в области культурного наследия

Физические коллекции в области культурного наследия являются результатом конструирования и ремедиации доступных данных. Цифровые архивы, коллекции и агрегаторы данных наследуют традиции и особенности физических коллекций, но вместе с тем демонстрируют и своеобразие, присущее коллекциям цифровых данных. Польза цифровых коллекций для гуманитарных исследований иногда подвергается сомнению. Так например, К. Трампенер утверждает, что исследовательская работа в гуманитарных науках менее эффективна с помощью цифровых коллекций, поскольку подход, связанный с большими данными, не позволяет увидеть индивидуальность и эстетические качества литературного произведения¹⁷¹. Цифровые коллекции также связывают с заданными категориями поиска, которые вмещаются в индивидуальные подходы к изучению произведений искусства и культуры¹⁷². При этом значительная часть исследователей не ставит под сомнение пользу цифровых коллекций и больших данных в гуманитарных науках, но предлагает учитывать те их особенности, которые могут привести к необъективным выводам. При рассмотрении цифровых коллекций крайне важными являются те параметры, которые они унаследовали от физических коллек-

¹⁷⁰ Lendemer, J., Thiers, B., Monfils, A.K., Zaspel, J., Ellwood, E.R., Bentley, A., LeVan, K., Bates, J., Jennings, D., Contreras, D., Lagomarsino, L., Mabee, P., Ford, L.S., Guralnick, R., Gropp, R.E., Revelez, M., Cobb, N., Seltmann, K. & Aime, M.C. (2020) The extended specimen network: a strategy to enhance US biodiversity collections, promote research and education. *BioScience*, Volume 70, Issue 1, Pages 23–30.

¹⁷¹ Trumpener, K. (2009). Critical Response I. Paratext and Genre System: A Response to Franco Moretti. *Critical Inquiry* 36, no. 1, 159–71.

¹⁷² Saussy, H. Exquisite Cadavers Stitched from Fresh Nightmares: Of Memes, Hives, and Selfish Genes. In *Comparative Literature in an Age of Globalization*, edited by Haun Saussy, 3–42. Baltimore: Johns Hopkins University Press, 2004.

ций. Физические коллекции отражают историю создания и восприятия произведений искусства и культуры¹⁷³: в коллекциях оказались те произведения, которые распространялись по миру в многочисленных копиях, те, которые повлияли на последующую творческую деятельность, вошли в художественный канон. Последующая ремедиация оказывает такое же влияние на цифровые коллекции. Цифровые коллекции данных всегда конструированы из элементов доступных источников и отображают объективную картину только в некоторой степени. К. Бод утверждает, что оцифрованные документы, тексты и изображения не могут полностью отобразить масштаб, составные части и характеристики процессов, происходивших в культуре. Например, австралийская коллекция оцифрованных газет агрегатора Trove¹⁷⁴ включает только тридцать процентов газет, опубликованных в некоторые периоды XIX века. Принимая во внимание разные явления культуры, исключенные из анализа в этом случае, можно сказать, что цифровая коллекция способна усилить определенные характеристики и исключить другие особенности из последующего анализа¹⁷⁵. При работе с большими данными в гуманитарных исследованиях лакуны в цифровых коллекциях и присущую им тенденциозность можно преодолеть с помощью создания правильной выборки и сравнения данных¹⁷⁶. Однако, как утверждает К. Бод, масштабирование исследований и применение техник статистического анализа не освобождает гуманитарную науку и исследования, связанные с изучением культуры, от ограничений, вызванных политическими, техническими и социальными причинами, которые приводят к недостаточной репрезентативности цифровых коллекций. Эти ограничения и условия, при которых обобщения, сделанные при анализе больших данных о явлениях культуры, становятся верными, должны быть приняты во внимание в гуманитар-

¹⁷³ Bode, K. *A World of Fiction*. en. Ann Arbor, MI: University of Michigan Press, 2018.

¹⁷⁴ URL: <https://trove.nla.gov.au/newspaper/?q=>

¹⁷⁵ Bode, K. *A World of Fiction*. en. Ann Arbor, MI: University of Michigan Press, 2018.

¹⁷⁶ Underwood, T. *Distant Horizons: Digital Evidence and Literary Change*. U of Chicago P, 2019.

ных исследованиях, документированы и введены в научный оборот¹⁷⁷. В этом случае агрегаторы могут указать на ограничения, лакуны и эпистемические пробелы, которые связаны с «историческими условиями, культурными и институциональными практиками, экономическими факторами и технологическими процессами»¹⁷⁸. Деятельность по изучению репрезентативности и сбалансированности коллекции, способности прийти к объективным выводам на ее основе в результате анализа больших данных должна предшествовать количественным исследованиям и даже предоставлению данных широкой публике для распространения, изучения, создания новых продуктов и расширения цифрового канона.

Цифровой канон и расширение культурного канона

Ряд произведений искусства и культуры получают большее распространение по сравнению с другими произведениями. Их включают в коллекции библиотек, музеев и архивов, «их покупает широкая публика, исследования, связанные с этими произведениями, поддерживаются в рамках разнообразных программ финансовой поддержки, их цитируют ученые, ими восхищаются в университетах и за их пределами»¹⁷⁹. Такие произведения составляют основу так называемого культурного канона.

Цифровая эпоха приводит к ремедиации культурного канона и иным средствам распространения произведений, составляющих его основу. Меняется сама инфраструктура знания¹⁸⁰, при этом разрыв между цифровыми коллекциями и до-

¹⁷⁷ Bode, K. (2020) Why you can't model away bias. *Modern Language Quarterly* (2020) 81 (1): 95–124.

¹⁷⁸ Там же.

¹⁷⁹ Price, K. M. (2009). Digital Scholarship, Economics, and the American Literary Canon. *Literature Compass* 6.2: 274–90.

¹⁸⁰ Mak, B. 2014. Archaeology of a Digitization. *Journal of the Association for Information Science and Technology* 65.8 (2014): 1515–26; Feldman M.H. 2016. Rethinking the Canon of Ancient Near Eastern Art in the Internet Age, *Journal of Ancient Near Eastern History* 2016; 3(1): 57–79; Fraistat, N., Jones, S.E., Stahmer, C. (1998). "The Canon, The Web, and the Digitization of Romanticism." *Romanticism On the Net*, 10. <http://id.erudit.org/iderudit/005801ar>.

кументами, которые они представляют, демонстрирует онтологические лакуны и сам по себе представляет эпистемическую ценность¹⁸¹. Разрыв между цифровыми платформами и произведенным в культуре знанием в большой степени влияет на создание, фиксацию или расширение цифрового канона.

Проблеме ремедиации при расширении печатного канона уделяют внимание Н. Фрейстат, С. Джоунз и К. Стамер в работе «Канон, цифровая среда и оцифровка литературных произведений в эпоху романтизма». Они дают следующее определение канону: «Каким бы ни было определение «канона», в конечном итоге он всегда определяется тремя взаимосвязанными функциями: определенный (возможно, не слишком жестким и окончательным образом) список произведений, различные сообщества и организации, которые обсуждают, производят, распространяют и потребляют эти произведения, и материальные формы, в которых воплощены эти произведения и с помощью которых они распространяются. Большая часть аналитической литературы, связанной с каноном, касается самих произведений, меньшее внимание уделяется взаимодействию сообществ, которые производят и потребляют произведения, еще меньше внимания уделяется тому, как печать и печатный станок повлияли на создание канона»¹⁸². В периоды, когда меняется инфраструктура, поддерживающая создание, распространение и канонизацию знаний, списки произведений культуры и искусства становятся гибкими, включая в себя неожиданные области и направления развития культуры. Это связано не только со сменой технологий, но и со сменой позиций сообществ и организаций, которые «обсуждают, производят и распространяют» произведения. Так, в конце двадцатого века смена позиций институций, сообществ и организаций, связанная с большей открытостью мира, социальной и географической мобильностью и распространением визуальной культуры, привела к так называе-

¹⁸¹ Mak, B. 2014. Archaeology of a Digitization. *Journal of the Association for Information Science and Technology* 65.8: 1515–26.

¹⁸² Fraistat, N., Jones, S.E., Stahmer, C. (1998). "The Canon, The Web, and the Digitization of Romanticism." *Romanticism On the Net*, 10. <http://id.erudit.org/iderudit/005801ar>.

мым «войнам канона», в которых, как иногда утверждается, выиграли сторонники релятивизма и мультикультурализма¹⁸³.

В начале двадцать первого века Франко Моретти предложил выйти за пределы войн канона и включить в анализ не только литературу, написанную женщинами, меньшинствами, представителями рабочего класса и литературу за пределами западного канона, но все литературные произведения, когда-либо созданные в мире¹⁸⁴. Этот подход можно распространить на все произведения культуры, включенные в цифровые коллекции и агрегаторы. Эпистемические представления западной культуры при таком подходе расширяются, включая способы и инфраструктуру познания всех культур, представленных текстами, изображениями и документами в цифровых коллекциях. Однако при создании цифровых коллекций образуются лакуны, возникает недостаточная репрезентативность в случае одних культур и сообществ и избыточная репрезентативность в случае других культур или физических коллекций¹⁸⁵. Такого рода искажения и тенденциозность поддерживаются ограничениями, связанными с авторским правом, запретом на повторное использование в собственном интерфейсе или публикации¹⁸⁶ или проблемами публикации неатрибутированных произведений, созданных в двадцатом веке¹⁸⁷. Поскольку большую часть цифровых коллекций составляют

¹⁸³ Donadio, Rachel (2007). Revisiting the Canon Wars. *The New York Times*, September 16, 2007. <http://www.nytimes.com/2007/09/16/books/review/Donadiot.html?ref=review&pagewanted=all>

¹⁸⁴ Bode, K. (2020) Why you can't model away bias. *Modern Language Quarterly* (2020) 81 (1): 95–124.

¹⁸⁵ Там же; см. также Kizhner, I., Terras, M., Romyantsev, M., Khokhlova, V., Demeshkova, E., Rudov, I., Afanaseva, J. 'Digital Cultural Colonialism: measuring bias in aggregated digitized content held in Google Arts and Culture', *Digital Scholarship in the Humanities*, в печати.

¹⁸⁶ Price, Kenneth M. 'Digital Scholarship, Economics, and the American Literary Canon.' *Literature Compass* 6.2 (2009): 274–90.

¹⁸⁷ Martinez, M., & Terras, M. (2019). 'Not adopted': The UK orphan works licensing scheme and how the crisis of copyright in the cultural heritage sector restricts access to digital content. *Open Library of Humanities*, 5(1), 1–51.

неатрибутированные произведения, созданные в двадцатом веке, мы видим культурное наследие с современной точки зрения через призму современных канонов и принципов отбора произведений в физических коллекциях¹⁸⁸. Атрибутированные произведения доступны для анализа и легко объединяются с контекстом благодаря совместимым метаданным и протоколам связанных данных. Таким образом, мы можем утверждать, что цифровые коллекции и агрегаторы имеют тенденцию усиливать стереотипы, сформированные в результате создания культурного канона, и фиксируют печатный канон.

Развитие цифровой инфраструктуры в области культурного наследия

Цифровые инфраструктуры, поддерживающие создание и распространение знаний¹⁸⁹ в области культурного наследия, направлены на создание открытых репозитариев, способствующих распространению и повторному использованию данных¹⁹⁰. Некоторые примеры подобных репозитариев, ряд коллекций которых лицензирован для повторного использования, включают цифровую библиотеку Европеана¹⁹¹, Цифровую публичную библиотеку Америки¹⁹², агрегатор цифровых ресурсов в области культурного наследия Новой Зеландии¹⁹³, Платформу «Открытое наследие» Министерства культуры Франции¹⁹⁴, Инфраструктуру связанных данных для изучения культурного наследия, объединяющую ряд университетов и организаций Канады¹⁹⁵, платформу Trove, которая предоставляет доступ

¹⁸⁸ Bode, K. *A World of Fiction*. en. Ann Arbor, MI: University of Michigan Press, 2018.

¹⁸⁹ Mak, B. 2014. "Archaeology of a Digitization." *Journal of the Association for Information Science and Technology* 65.8 (2014): 1515–26

¹⁹⁰ Poirier L., Fortun K., Costelloe-Kuehn B., Fortun M. (2020) *Metadata, Digital Infrastructure, and the Data Ideologies of Cultural Anthropology*. In: Crowder J., Fortun M., Besara R., Poirier L. (eds) *Anthropological Data in the Digital Age*. Palgrave Macmillan, Cham.

¹⁹¹ URL: <https://www.europeana.eu/en>

¹⁹² URL: <https://dp.la/>

¹⁹³ URL: <https://digitalnz.org/>

¹⁹⁴ URL: <https://www.pop.culture.gouv.fr>

¹⁹⁵ URL: <https://lincsproject.ca/what-is-lincs/>

к метаданным цифровых коллекций в области культурного наследия Австралии¹⁹⁶. Действительно, цифровую библиотеку «Европеана» часто определяют как портал, который в конечном итоге сможет дать возможность использовать информацию о культурном наследии с помощью прикладного интерфейса программирования (API)¹⁹⁷. Открытое распространение данных способствует тому, что исследования приобретают дополнительную глубину и сложность, создаются коллаборации ученых, аудитория, которая пользуется данными, расширяется и стратифицируется¹⁹⁸. Цифровые инфраструктуры в гуманитарных науках способствуют укреплению коллабораций в гуманитарных науках, где цифровые исследования развиваются при участии специалистов из разных областей знаний. К тому же эпистемическая культура в цифровых гуманитарных науках часто требует моделирования данных и связей между данными, выделения тех черт в сложных объектах и произведениях культуры и искусства, которые могут быть исследованы цифровыми методами, и обнаружения связей между этими свойствами и внешним контекстом¹⁹⁹. Системы публикации и распространения тех данных, которые хранятся в разных базах данных, коллекциях и странах, позволяют

¹⁹⁶ URL: <https://trove.nla.gov.au/>

¹⁹⁷ Cesare Concordia, Stefan Gradmann, Sjoerd Siebinga (2010): Not just another portal, not just another digital library: A portrait of Europeana as an application program interface. In: International Federation of Library Associations and Institutions 36(1), pp. 61–69. (<http://dx.doi.org/10.1177/0340035209360764>) ; Doerr, M., Gradmann, S., Henicke, S., Isaac, A., Meghini, C., & Van de Sompel, H. (2012). The Europeana Data Model (EDM). Paper presented at the World Library and Information Congress: 76th IFLA General Conference and Assembly, Gothenburg, Sweden.

¹⁹⁸ Poirier L., Fortun K., Costelloe-Kuehn B., Fortun M. (2020) Metadata, Digital Infrastructure, and the Data Ideologies of Cultural Anthropology. In: Crowder J., Fortun M., Besara R., Poirier L. (eds) *Anthropological Data in the Digital Age*. Palgrave Macmillan, Cham.

¹⁹⁹ McCarty, Willard. *Humanities Computing*. London: Palgrave Macmillan, 2005; Flanders, Julia and Fotis Jannidis. (2016). Data Modeling. In *A New Companion to Digital Humanities*, edited by Susan Schreibman, Ray Siemens, and John Unsworth, 229–37. New York: Wiley & Sons.

связать разные характеристики модели и объединить данные о ряде объектов культурного наследия, представленные в разных форматах и на разных языках, либо с помощью совместимых метаданных²⁰⁰, либо с помощью онтологий, ориентированных на включение культурного наследия в цепь событий и отношения между событиями²⁰¹. Цифровые инфраструктуры демонстрируют возможность перехода от просмотрных и поисковых функций к аналитическим функциям и далее к функциям постановки задач и обнаружения закономерностей в данных. Ограничения агрегаторов, связанные с пробелами, лакунами и недостаточной или избыточной репрезентативностью части данных, приводят к необходимости критического анализа цифровых инфраструктур и агрегаторов как первичных источников данных в гуманитарных науках²⁰².

Цифровые инфраструктуры, создающие и распространяющие знание в области культурного наследия, обычно являются централизованными базами данных, которые администрируются крупной правительственной организацией или учреждением культуры. Так, например, цифровая библиотека Европеана со штаб-квартирой в Гааге администрируется в рамках одного из проектов Европейской Комиссии и финансируется Европейским Союзом, а также правительствами стран, которые являются членами Европейского Союза²⁰³. Платформа Trove предоставляет доступ к распределенному контенту, имеющему отношение к Австралии, администрируется Национальной библиотекой Австралии и финансируется правительством Австралии²⁰⁴. С другой стороны, Цифровая публичная биб-

²⁰⁰ См. напр. URL: <https://dublincore.org/>

²⁰¹ URL: <http://cidoc-crm.org/>

²⁰² Hyvonen, E.: Using the Semantic Web in Digital Humanities: Shift from data publishing to data-analysis and serendipitous knowledge discovery. *Semantic Web – Interoperability, Usability, Applicability* 11(1), 187–193 (2020).

²⁰³ URL: <https://pro.europeana.eu/about-us/governance>

²⁰⁴ Jones, Mike, and Deb Verhoeven. 2016. “Treasure Trove: Why Defunding Trove Leaves Australia Poorer.” *The Conversation*. https://openresearch-repository.anu.edu.au/bitstream/1885/200473/2/01_Jones_Treasure_Trove%253A_why_defunding_2016.pdf

лиотека Америки является объединением локальных центров, которые агрегируют распределенные данные американских библиотек и архивов. Хотя штаб-квартира этой цифровой инфраструктуры находится в одном из подразделений Гарвардского университета, этот агрегатор финансируется благодаря поддержке фондов и благотворительных организаций²⁰⁵. В 1986 году модульная система агрегатора метаданных/оцифрованного контента в области культурного наследия была предложена российским ученым Я.А. Шером, специалистом по развитию цифровых музейных коллекций²⁰⁶. Таким образом, модель цифровой инфраструктуры может включать как модульную систему, так и централизованную систему, которая администрируется подразделением крупной организации, работающей в области культурного наследия.

Цифровые коллекции и инфраструктуры направлены на публикацию оцифрованного контента с целью возможного влияния на международную аудиторию пользователей. Несоответствие масштаба оцифрованного контента реальному количеству неоцифрованных и неопубликованных произведений культуры и искусства поднимает важные вопросы о том, кто решает, какие произведения будут опубликованы, войдут в научный и образовательный оборот, расширят цифровой канон и будут служить источником вдохновения для широкой публики. Таким образом, проделанный обзор научной литературы демонстрирует спектр проблем развития цифровых коллекций и агрегаторов цифрового контента в области культурного наследия. Прежде всего стоит отметить, что в современных условиях цифровые коллекции воспринимаются не столько как информационный и маркетинговый буклет физической коллекции, сколько как публикационная платформа, задача которой - создание нового знания и распространение знаний. Такие публикационные платформы воспринимаются как динамические структуры, связываю-

²⁰⁵ URL: <https://pro.dp.la/about>

²⁰⁶ Sher, J. (1986). Museum documentation systems and computers: USSR experience. In Light, R.B., Roberts, D.A., and Stewart, J.D. (eds) *Museum Documentation Systems. Developments and Applications*. Butterworths: London, Boston, pp. 287-292.

шие информацию о произведениях и контекстах в единое целое и намечающие новые задачи, гипотезы и исследования в гуманитарных науках. Они направлены на ремедиацию и реконструирование знания, обнаружение новых контекстов и создание творческих ресурсов. С этим связана проблема повторного использования, ремедиации и возможности изменения ресурсов и проблема применения открытых лицензий для развития инфраструктуры. Не менее важной является проблема тенденциозности цифровых коллекций, когда одни культуры или сообщества представлены недостаточно полно, в то время как другие части коллекции будут всегда доминировать, особенно в ситуации анализа больших данных. Это ведет к тому, что цифровой канон окажется расширен за счет цифровых коллекций тех культур, которые имеют возможность и желание публиковать контент и лицензировать его для повторного использования. Публикационные платформы в области культурного наследия будут усилены рекомендательными системами, основанными на алгоритмах машинного обучения, что внесет дополнительные сложности в проблему прозрачности и доступности одного типа контента и отсутствия цифровых ресурсов, связанных с рядом культур и сообществ. Можно утверждать, что цифровые публикационные платформы должны восприниматься как первичные источники, которые отражают культурологические, политические и социальные проблемы нашей эпохи и показывают онтологические и эпистемические лакуны в восприятии культурной, этнической и социальной принадлежности.

Таким образом, анализ концептуальных и методических подходов к развитию современных цифровых технологий в области культурного наследия, определивших цифровой поворот для всех современных культурных процессов, позволил понять основные закономерности и тенденции, связанные с фиксацией, анализом и трансляцией культурного наследия на современном этапе. Можно увидеть, что критический анализ цифровых инфраструктур дает возможность провести культурологический анализ цифрового поворота в культурном наследии для того, чтобы понять возможности и ограничения цифровых технологий при анализе, публикации и распространении текстовых и визуальных материалов, де-

монстрирующих произведения культуры и искусства. Показано, что благодаря росту количества и качества цифровых изображений, увеличению объемов места хранения и сетевых технологий, позволяющих распространение изображений высокого разрешения, можно говорить о преобразующем влиянии цифровых технологий в области фиксации, анализа и трансляции культурного наследия. Можно утверждать, что анализ цифровых инфраструктур в области культурного наследия в ближайшее время может стать одним из важных направлений культурологического анализа.

1.2 Коллекции цифровых изображений и агрегаторы визуального контента в области культурного наследия

Цифровые коллекции и агрегаторы визуального контента в области культурного наследия демонстрируют ряд общих черт в способах создания и формах функционирования. Эти черты являются одним из основных факторов развития культурных закономерностей в современных процессах формирования цифровых коллекций и агрегаторов оцифрованного контента. Цифровые коллекции создаются как для академического сообщества, так и для общей публики. Обычно их целью является предоставление доступа к первичным архивным материалам и внедрение цифровых инструментов, технологий и программного обеспечения в проведение исследований в гуманитарных науках²⁰⁷. Важной задачей цифровых коллекций и агрегаторов является создание и распространение нового знания и, следовательно, включение цифровых материалов в дополнительные контексты²⁰⁸. Особенным преимуществом цифровых коллекций и агрегаторов является высокое качество изображений и возможность сравнения материалов со схожими данными, аннотированными с помощью тех же элементов словаря в системе управления коллекцией. Случается, что объемные произведения, такие как скульптура или археологические материалы, бывают представлены в цифровых коллекциях как трехмерные модели или панорамные изображения. Чрезвычайно важно то, что результаты анализа коллекций и новое знание, созданное в резуль-

²⁰⁷ Westerby, G., & Keegan, K. (2019). Digital Art History and the Museum: The Online Scholarly Collection Catalogues at the Art Institute of Chicago. *Visual Resources*, 1–14.

²⁰⁸ Simou N., Chortaras A., Stamou G., Kollias S. (2017) Enriching and Publishing Cultural Heritage as Linked Open Data. In: Ioannides M., Magnenat-Thalmann N., Papagiannakis G. (eds) *Mixed Reality and Gamification for Cultural Heritage*. Springer, Cham.

тате применения цифровых технологий, можно быстро донести до международной аудитории²⁰⁹.

Принципы создания и функционирования цифровых коллекций и агрегаторов в области культурного наследия рассматривали такие влиятельные ученые как К. Линч²¹⁰, П. Марти²¹¹, М. Террас²¹², Р. Парри²¹³, Г. Шрайбер²¹⁴, М. Бака²¹⁵, М. Ван Эрп²¹⁶.

Теоретический и методологический анализ цифровых коллекций и агрегаторов хорошо согласуется с исследованием цифровых коллекций как медийных форм, где исследователи продолжают традицию критического анализа культурного наследия, сложившуюся в двадцатом веке²¹⁷. В рамках этой традиции цифровые коллекции рассматриваются как отражение особенностей, и тенденций, сло-

²⁰⁹ Westerby, G., Keegan, K. (2019). Digital Art History and the Museum: The Online Scholarly Collection Catalogues at the Art Institute of Chicago. *Visual Resources*, 1–14.

²¹⁰ Lynch, C. A. (2002). Digital collections, digital libraries and the digitization of cultural heritage information. *First Monday*, 7(5).

²¹¹ Marty, P.F., and K. Burton Jones, eds. 2008. *Museum informatics: People, information and technology in museums*. New York: Routledge.

²¹² Terras, M. (2008) *Digital images for the information professional*. Ashgate, Burlington VT.

²¹³ Parry, R. *Recoding the Museum: Digital Heritage and the Technologies of Change*. Routledge, 2007.

²¹⁴ Dijkshoorn, C., Jongma, L., Aroyo, L., van Ossenbruggen, J., Schreiber, G., ter Weele, W., Wielemaker, J.: The Rijksmuseum Collection as Linked Data. *Semantic Web*. 9(2), 221–230 (2018);

²¹⁵ Baca, M., ed.: *Introduction to metadata*. The Getty Research Institute, Los Angeles (2008).

²¹⁶ Chiel van den Akker, Susan Legêne, Marieke Van Erp, Lora Aroyo, Roxane Segers, Lourens Van der Meij, Jacco van Ossenbruggen, Guus Schreiber, Bob Wielinga, Johan Oomen, and Geertje Jacobs. *Digital Hermeneutics: Agora and the Online Understanding of Cultural Heritage*. In *Proceedings of the 3rd International Conference on Web Science (WebSci 2011)*, pages 1–7, 2011.

²¹⁷ Smith, L. *The Uses of Heritage*. London: Routledge, 2006; Smith, Laurajane and Natsuko Akagawa (eds) (2009) *Intangible Heritage*. London: Routledge; Gentry, K., & Smith, L. (2019). Critical heritage studies and the legacies of the late-twentieth century heritage canon. *International Journal of Heritage Studies*, 1–21.

жившихся в доцифровую эпоху и усиленных с помощью современных рекомендательных систем.

Дж. ван Дайк, Т. Поэл и М. де Ваал²¹⁸ подвергают сомнению представления о прогрессивном влиянии онлайн-платформ и их важность для развития общества и задаются вопросом о том, «кто... должен нести ответственность за внедрение и укрепление социальных ценностей в обществе, где цифровые платформы имеют широкое распространение». Несмотря на важное значение онлайн-платформ в области культурного наследия для краудсорсинговых проектов, таких как Wikimedia Commons или Flickr Commons²¹⁹, количество исследований, в которых рассматриваются социальные последствия создания национальных и международных агрегаторов визуального культурного наследия, крайне незначительно. Й. Штиллер и В. Петрас²²⁰ отмечают, что среди сорока исследований, дающих обзор и оценку коллекций Европеаны²²¹, есть только две статьи о социальных конструктах, определяющих социальную ценность этого агрегатора европейского культурного наследия. В частности, Б. Вальтиссон²²² показывает, что Европеана продолжает традицию укрепления тех социальных конструктов, которые сформировались в музе-

²¹⁸ Van Dijck J, Poell T and Waal M de *The Platform Society*. New York: Oxford University Press, 2018.

²¹⁹ см., например, Kalfatovic, M. R., Kapsalis, E., Spiess, K. P., Van Camp, A., & Edson, M. (2008). Smithsonian team Flickr: A library, archives, and museums collaboration in Web 2.0 space. *Archival Science*, 8(4), 267–277; Oomen, J. & Aroyo, L. Crowdsourcing in the cultural heritage domain: opportunities and challenges. *Communities and Technologies 2011*, 138-149; Ridge, Mia (ed.) (2014). *Crowdsourcing our cultural heritage*. London: Ashgate Publishing, Ltd; Terras, M. 2016. “Crowdsourcing in the Digital Humanities.” In: *A New Companion to Digital Humanities*, edited by S. Schreibman, R. Siemens, and J. Unsworth, 420–438. Oxford: Blackwell.

²²⁰ Stiller, J., Petras, V.: Learning from digital library evaluations: The Europeana case. *ABI Technik* 38(1), 37–45 (2018).

²²¹ URL: <https://www.europeana.eu/portal/ru>

²²² Valtysson, B. (2012). “Europeana – The digital construction of Europe’s collective memory.” *Information, Communication & Society* 15,2: 151–170.

ях, библиотеках и архивах в девятнадцатом и двадцатом веке. И. Янкова, Е. Велей и К. Нушева²²³ вводят возможные требования к контенту Европейцы. В частности, предлагается получить более сбалансированный контент с минимальным количеством изображений предметов музеев, архивов и библиотек от каждого государства-члена Европейского Союза с тем, чтобы культурное наследие было представлено более равномерно и репрезентативно.

Объем цифровых коллекций и агрегаторов в области культурного наследия и то, в какой мере они представляют процессы развития культуры, являются двумя важными характеристиками, которые определяют основные способы и формы создания цифровых коллекций и функционирования агрегаторов оцифрованного контента в области культурного наследия. В то время как объем данных часто предлагают считать основной характеристикой, определяющей качество цифровой коллекции или агрегатора, репрезентативность оказывается характеристикой, которая может применяться к коллекциям цифровых данных с большими ограничениями, связанными с особенностями создания как физических, так и цифровых коллекций²²⁴.

Данное диссертационное исследование применяет понятия репрезентативности и сбалансированности цифровых коллекций данных, которые используются

²²³ Yankova, I. et al. (2015). The European Digital Library-Factor for Long-life learning in Arts and Cultural Studies. *Qualitative & Quantitative Methods in Libraries* 4,4: 965–971.

²²⁴ Egbert J. (2019). Corpus design and representativeness. In Sardinha T. B. and Pinto V. (eds.) *Multidimensional Analysis. Research Methods and Current Issues*. Bloomsbury Academic, pp. 27-42; Bode, K. (2020). 'The Archive,' in *The Cambridge Companion to Literature in the Digital Age* (Ed. Adam Hammond). Cambridge: Cambridge University Press; Bode, K. (2020) Why you can't model away bias. *Modern Language Quarterly* (2020) 81 (1); Tahmasebi, Nina, Niclas Hagen, Daniel Brodén, and Mats Malm. "A Convergence of Methodologies: Notes on Data-Intensive Humanities Research." In *Digital Humanities in the Nordic Countries 4th Conference*. Helsinki: Nina Tahmasebi, 2019.

в корпусной лингвистике²²⁵, где корпус является репрезентативным, когда он представляет «все составляющие части языкового разнообразия»²²⁶. Подробный обзор литературы в области репрезентативности и сбалансированности цифровых коллекций культурного наследия представлен в статье И. Кижнер и др²²⁷. Сбалансированный корпус должен следовать правилу: «пропорция, в которой конкретная часть представлена в корпусе, должна отражать соотношение этой части в данном типе языковых явлений или важность этой части в этом типе языковых явлений»²²⁸. С.Т. Грис объясняет, что сложность создания репрезентативного и сбалансированного корпуса заключается в том, что трудно понять пропорцию части в явлении реальной жизни, таком как язык или, в нашем случае, коллекции произведений искусства, где размер популяции или распределение частей неизвестны. Не менее сложно понять важность этих частей²²⁹. Например, ряд национальных музеев имеет обширные коллекции монет и бумажных денег. Так, например, нумизматическая коллекция Государственного исторического музея в Москве содержит 1,7 миллионов предметов и составляет почти 30% основной

²²⁵ См. например, Biber, D. (1993). Representativeness in Corpus Design. *Literary and Linguistic Computing*, 8, Issue 4, Oxford University Press, pp. 243-257; Leech, Geoffrey, 2007. New resources, or just better old ones? The Holy Grail of representativeness. In: Hundt, Marianne, Nesselhauf, Nadja, Biewer, Carolin (Eds.), *Corpus Linguistics and the Web*. Rodopi, Amsterdam, pp. 133-150.

²²⁶ Gries, S.T. 2009. What is corpus linguistics? *Language and Linguistics Compass* 3(5): 1225-1241, p. 1231.

²²⁷ Kizhner et al. Digital cultural colonialism: measuring bias in aggregated digitized content held in Google Arts and Culture, *Digital scholarship in the Humanities* (в печати).

²²⁸ Gries, S.T. 2009. What is corpus linguistics? *Language and Linguistics Compass* 3(5): 1225-1241, с. 1231.

²²⁹ Leech, G., 2007. New resources, or just better old ones? The Holy Grail of representativeness. In: Hundt, Marianne, Nesselhauf, Nadja, Biewer, Carolin (Eds.), *Corpus Linguistics and the Web*. Rodopi, Amsterdam, pp. 133-150; Корчагин К. М. «Поэзия XX века в поэтическом подкорпусе Национального корпуса русского языка: Проблема репрезентативности». *Труды Института русского языка им. В. В. Виноградова* 6 (2015б): 235–256

коллекции²³⁰. Какое пропорциональное соотношение нумизматической коллекции должен включать цифровой агрегатор предметов культурного наследия, чтобы он стал сбалансированным корпусом? Столь же важен вопрос, должны ли составители корпуса основываться на количестве произведенных текстов или на том, как тексты воспринимались и входили в обращение (количество тиражей, продаж, упоминаний в хрестоматиях и т. д.), то есть следует ли нам полагаться на количество текстов, произведенных в культуре, или на влияние текстов на дальнейшее историческое развитие языка²³¹.

Хотя каждый корпус «неизбежно нерепрезентативен»²³² и «Священный Грааль» репрезентативной и сбалансированной коллекции электронных текстов (или изображений) вряд ли будет достигнут²³³, лингвисты, которые рассматривали концепции репрезентативности и баланса в конце XX века, настаивали на подробном обсуждении методологических принципов: как получить материалы, включенные в корпус, и могут ли методы выборки влиять на достоверность

²³⁰ URL: <https://shm.ru/kolleksii-i-muzeynyy-kompleks/nauchno-khranitel'skie-otdely/otdely/otdel-numizmatiki/>

²³¹ Atkins, S., Clear, J., & Ostler, N. (1992). Corpus Design Criteria. *Journal of Literary and Linguistic Computing*, 7(1), 1-16; Clear, J. (1992). Corpus sampling. In G. Leitner (ed.) *New directions in English language corpora*, Mouton-de-Gruyter, Berlin, pp. 21 - 31; Biber, D. (1993). Representativeness in Corpus Design. *Literary and Linguistic Computing*, 8, Issue 4, Oxford University Press, pp. 243-257; Leech, G., 2007. New resources, or just better old ones? The Holy Grail of representativeness. In: Hundt, Marianne, Nesselhauf, Nadja, Biewer, Carolin (Eds.), *Corpus Linguistics and the Web*. Rodopi, Amsterdam, pp. 133-150

²³² Clear, J. (1992). Corpus sampling. In G. Leitner (ed.) *New directions in English language corpora*, Mouton-de-Gruyter, Berlin, p. 23; Leech, Geoffrey, 2007. New resources, or just better old ones? The Holy Grail of representativeness. In: Hundt, Marianne, Nesselhauf, Nadja, Biewer, Carolin (Eds.), *Corpus Linguistics and the Web*. Rodopi, Amsterdam, pp. 133-150

²³³ Leech, G. (2007). New resources, or just better old ones? The Holy Grail of representativeness. In: Hundt, Marianne, Nesselhauf, Nadja, Biewer, Carolin (Eds.), *Corpus Linguistics and the Web*. Rodopi, Amsterdam, pp. 133-150

результатов²³⁴. Однако недавнее исследование²³⁵ показывает, что большинство лингвистов, опубликовавших эмпирические исследования корпусов в лингвистических журналах в 2014 году, полагали, что наиболее важной характеристикой, определяющей корпус, был его размер, а не его репрезентативность. Это имеет важные последствия не только для лингвистических исследований, основанных на эмпирических исследованиях корпуса, но также и для множества исследований, в которых используются большие наборы данных, в том числе с цифровым культурным содержанием.

А. Данст, Р. Хартел и И. Лаудброк²³⁶ сообщают о методологических принципах, которые использовались при подготовке корпуса комиксов, и используют методы из области корпусной лингвистики для анализа других культурных форматов, таких как графические романы. Авторы согласны с К. Маекава, Х. Койсо, С. Фуруй и Х. Ишара²³⁷, которые утверждают, что случайная выборка – лучший способ создать репрезентативный корпус, они используют случайную выборку в качестве методологического принципа для подготовки корпуса графических новелл. Однако К. Боуд не считает, что случайная выборка ведет к репрезентативности как свойству, отражающему реальные процессы, характеризующие культуры

²³⁴ Woods, A., Fletcher, P., & Hughes, A. (1986). *Statistics in language studies*. Cambridge: Cambridge University Press, цитируется по Clear, J. (1992). *Corpus sampling*. In G. Leitner (ed.) *New directions in English language corpora*, Mouton-de-Gruyter, Berlin

²³⁵ Egbert, J. (2019). 'Corpus Design and Representativeness'. In Tony Berber Sardinha and Veirano Pinto (eds.) *Multidimensional Analysis. Research Methods and Current Issues*, pages 27-42, Bloomsbury Academic.

²³⁶ Dunst, A.; Hartel, R.; Laubrock, J. (2017). *The Graphic Narrative Corpus (GNC): Design, Annotation, and Analysis for the Digital Humanities*. In *Proceedings of the 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, Kyoto, Japan, 9–15 November 2017; pp. 15–20.

²³⁷ Maekawa, K., Hanae Koiso, Sadaoki Furui and Hitoshi Isahara (2000). *Spontaneous speech corpus of Japanese*. *Proceedings of the Second International Conference of Language Resources and Evaluation (LREC)*, 2, 947-952.

прошлого. Особенности распространения и трансляции произведений культуры и искусства, принципы подбора и организации материалов как физических коллекций, так и цифровых агрегаторов влияют на состав цифровой коллекции и препятствуют созданию репрезентативного набора данных.

Работа над корпусом произведений искусства для анализа в цифровых гуманитарных науках может привести к важным методологическим вопросам, которые включают в себя вопрос доли изображений предметов различных культур с целью расширения хрестоматийных канонов истории искусств²³⁸ и вопрос доли изображений предметов, которые выводят пользователя за пределы социальных конструкторов и точек зрения западной (имперской) культуры.

В научной литературе 1990-х годов появляются первые исследования о том, что мир, представленный учреждениями культуры предшествующей эпохи, являлся редуцированным, колонизированным миром²³⁹. В последующие десятилетия многочисленные исследования подчеркивали трудности представления мультикультурных, не единственных, а многообразных точек зрения в музее²⁴⁰. В этих

²³⁸ Earhart, A.E. (2012). Can information be unfettered?: Race and the new digital humanities canon. In M. K. Gold, editor, *Debates in the Digital Humanities*, pages 309–318. University of Minnesota Press, Minneapolis, London; Feldman, M.H. (2017). Rethinking the Canon of Near Eastern Art in the Internet Age, *Journal of Ancient Near Eastern History*. 3: 57-79; Kizhner, I., Terras, M., Rumyantsev, M., Sycheva, K., Rudov, I. Accessing Russian Culture Online: The scope of digitization in museums across Russia. *Digital Scholarship in the Humanities*, Volume 34, Issue 2, June 2019, Pages 350–367.

²³⁹ Boast, R. (2011) Neocolonial collaboration: museum as contact zone revisited. *Mus Anthropol* 34:56–70; Jones, A. L. (1993). Exploding Canons: the Anthropology of Museums. In: *Annual Review of Anthropology* 22 (1), p. 201–20.

²⁴⁰ Ashley, S. (2005). First Nations on View: Canadian Museums and Hybrid Representations of Culture. In *Hybrid Entities*, Annual Graduate Conference Hosted by the York/Ryerson Programme in Communication and Culture. Pp. 31–40. Canada: Rogers Communication Centre, York University; Knell, Simon J, Peter Aronsson & Arne Bugge Amundsen (eds), (2010). *National Museums: New Studies from Around the World*, London: Routledge; Boast, R. (2011) Neocolonial collaboration: museum as contact zone revisited. *Mus Anthropol* 34:56–70; Polm, M. 2016. 'Museum representations of Roman Britain and Roman London: a postcolonial perspective'. *Britannia* 47: 209-241.

исследованиях подчеркивались социальные и политические силы, определяющие политику развития коллекций в архиве и музее²⁴¹. Говоря о неокOLONIALИЗМЕ как о существенной особенности учреждения музея, где входят в контакт разные культуры и точки зрения, Р. Боуст утверждает, что «хотя вес разных культур нельзя сопоставить..., важно то, что в несопоставимом контексте выигрывают доминирующие культуры. Это то, что важно знать о месте пересечения и контакта культур»²⁴². Несмотря на утверждение Амита Суда²⁴³ о том, что цифровая платформа Google Arts and Culture (GA&C) открывает новую эру в изучении истории искусства, которая будет способствовать развитию исследований в этой области, ранее платформа GA&C критиковалась как база данных, которая не подходит для искусствоведческих исследований из-за ее популярного характера²⁴⁴ и корпус, где ряд культур недостаточно представлен и маргинализирован²⁴⁵. Несмотря на это, платформа GA&C является «крупным производителем культурного контента», при этом посещаемость сайта GA&C составила 50 миллионов человек

²⁴¹ Manoff, M. (2004). Theories of the Archive from Across the Disciplines. *portal: Libraries and the Academy* 4, no. 1: 9–25; Glinka, K, Meier, S and Dörk, M 2015 Visualising the » Un-seen «: Towards Critical Approaches and Strategies of Inclusion in Digital Cultural Heritage Interfaces. In: Busch, C and Sieck, J (Eds.) *Kultur und Informatik (XIII)*, pp. 105–18. Berlin: Cross Media, Werner Hülsbusch.

²⁴² Boast R (2011) Neocolonial collaboration: museum as contact zone revisited. *Mus Anthropol* 34, p. 64.

²⁴³ Sood, A. (2016). Every piece of art you've ever wanted to see - up, close and searchable. TED lecture. URL: https://www.ted.com/talks/amit_sood_every_piece_of_art_you_ve_ever_wanted_to_see_up_close_and_searchable

²⁴⁴ Mansfield, Elizabeth C. (2014) Google Art Project and Digital Scholarship in the Visual Arts. URL: <http://www.google.com/culturalinstitute/project/art-project>, *Visual Resources: An International Journal of Documentation*, 30:1, 110–117.

²⁴⁵ Pfisterer, U. (2018). Big Bang Art History. *Digital Art History*, No 3 (2018).

между 2011 и 2018 годами²⁴⁶. В контексте этого диссертационного исследования, это значит, что доступность произведений искусства способствует тому, что эта популярная платформа становится примером того, как агрегаторы визуального культурного контента должны быть структурированы и организованы в дальнейшем. Состав цифровой коллекции и выбор изображений и предметов культурного наследия, определяемые требованием сделать сюжеты и изображения понятными для широкой аудитории²⁴⁷, приводит, по утверждению У. Пфистерера²⁴⁸, к доминированию канонических форм искусства. Таким образом логика классификационных схем и способов фильтрации произведений, встроенная в физические коллекции, унаследованная от предыдущих эпох или связанная с популярными представлениями о культуре, оказывается доминирующей в крупных агрегаторах цифрового контента в области культурного наследия, таких как Google Arts and Culture или цифровая библиотека «Европеана»²⁴⁹.

Кроме этого в научной литературе о тенденциозности в лингвистических корпусах²⁵⁰, электронных коллекциях литературных текстов²⁵¹ или коллекциях

²⁴⁶ The Art Newspaper. (2018). 'How Google became a major producer of cultural content'. 19 January, 2018. URL: <https://www.theartnewspaper.com/feature/how-google-became-a-major-producer-of-cultural-content>

²⁴⁷ Там же.

²⁴⁸ Pfisterer, U. (2018). Big Bang Art History. Digital Art History, No 3 (2018).

²⁴⁹ Glinka, K, Meier, S and Dörk, M. (2015). Visualising the » Un-seen «: Towards Critical Approaches and Strategies of Inclusion in Digital Cultural Heritage Interfaces. In: Busch, C and Sieck, J (Eds.) Kultur und Informatik (XIII), pp. 105–18. Berlin: Cross Media, Werner Hülsbusch.

²⁵⁰ Caliskan A, Bryson JJ, Narayanan A. (2017). Semantics derived automatically from language corpora contain human-like biases. Science 356(6334):183–86.

²⁵¹ Grieve-Smith, A. (2019). The importance of sampling frames in representative historical corpora : a case study of Parisian theater. CogniTextes. Vol. 2019. URL: <https://journals.openedition.org/cognitextes/1671>.

музейных предметов²⁵² демонстрируется, как легко предвзятость, связанная с языком или выбором произведений искусства, проникает в базу данных и влияет на результаты анализа. К. Боуд отмечает, как трудно создать цифровую коллекцию данных, в которой не будет предвзятости и тенденциозности²⁵³. Исследователь предполагает, что, прежде чем перейти к количественному и статистическому анализу для последующей работы с большими данными и получения результатов анализа, необходимо проделать определенную методологическую работу, чтобы создать курируемую коллекцию произведений искусства, которая демонстрирует и объясняет выбор и интерпретацию включенных в нее текстовых документов. То, какие коллекции, произведения искусства, периоды или предметы, представляющие определенные географические регионы, оцифрованы, опубликованы, агрегированы и доступны для поиска, является свидетельством ограничений, изначально присущих культурным учреждениям, эпистемическим культурам, социальным и политическим системы. Этот эпистемологический выбор должен быть в явной форме продемонстрирован пользователю, при этом должны быть предложены интерпретации и объяснения выбора культурных явлений и изображений предметов, включенных в агрегатор для публикации и последующего распространения. С другой стороны Н. Тахмасеби, Н. Хаген, Д. Броден и М. Матт утверждают, что в ситуации большого количества данных о культурном наследии и последующего количественного анализа, демонстрация всех ограничений не представляется реализуемой в практическом развитии проекта²⁵⁴.

Таким образом, культурные исследования, основанные на цифровых данных, демонстрируют скорее тенденции и закономерности процесса конструиро-

²⁵² Huster, A. C. (2013). Assessing Systematic Bias in Museum Collections: A Case Study of Spindle Whorls. *Advances In Archaeological Practice*. 1(2):77-90.

²⁵³ Bode, K. (2020). 'Why You Can't Model Away Bias,' *Modern Language Quarterly* 81.1.

²⁵⁴ Tahmasebi, Nina, Niclas Hagen, Daniel Brodén, and Mats Malm. (2019). A Convergence of Methodologies: Notes on Data-Intensive Humanities Research. In *Digital Humanities in the Nordic Countries 4th Conference*. Helsinki: Nina Tahmasebi, 2019.

вания культурологического корпуса, а не тенденции и закономерности в развитии культуры²⁵⁵. Среди недостатков цифровых коллекций как инфраструктуры, способствующей развитию гуманитарных наук, отмечают то, что они не дают возможности критического анализа произведений и сводятся к большим наборам данных. М. Джокерс призывает использовать компьютерные инструменты анализа, чтобы открыть те закономерности, которые ранее были недоступны исследователям²⁵⁶. Однако К. Боуд считает, что цифровые коллекции и архивы необходимо рассматривать как динамические структуры, подверженные влиянию людей, которые их организуют и которые ими управляют и компьютерных технологий, которые используются для организации коллекций²⁵⁷. Несмотря на это анализ данных в гуманитарных науках редко принимает во внимание то, как источники данных влияют на их характеристики, обращая большее внимание на размер цифровой коллекции²⁵⁸. К. Боуд утверждает, что обосновывать достоинства коллекции культурологических данных размером этой коллекции было бы явным упрощением. Несомненным подтверждением связи цифровых коллекций с практиками подбора материалов для аналоговой коллекции культурных данных служит пример цифровой библиотеки *HathiTrust Digital Library*²⁵⁹, которая создавалась на основе университетских библиотек США, где подбор литературы характеризовался особенностями канона, учебных программ университетов²⁶⁰ и политики

²⁵⁵ Bode, K. (2020). 'The Archive,' in *The Cambridge Companion to Literature in the Digital Age* (Ed. Adam Hammond). Cambridge: Cambridge University Press.

²⁵⁶ Jockers, M. L. *Macroanalysis: Digital Methods and Literary History*. Urbana: University of Illinois Press, 2013.

²⁵⁷ Bode, K. (2020). 'The Archive,' in *The Cambridge Companion to Literature in the Digital Age* (Ed. Adam Hammond). Cambridge: Cambridge University Press.

²⁵⁸ Cordell, R. (2017). 'Q i-jtb the Raven': Taking Dirty OCR Seriously. *Book History* 20 (2017): 188–225.

²⁵⁹ URL: <https://www.hathitrust.org/>

²⁶⁰ Bode, K. (2020). 'The Archive,' in *The Cambridge Companion to Literature in the Digital Age* (Ed. Adam Hammond). Cambridge: Cambridge University Press.

развития коллекций крупных государственных библиотек США. Гендерная, языковая и культурная тенденциозность, встроенная в физические коллекции, мигрировала в агрегированный оцифрованный контент, созданный на основе уже имеющихся физических коллекций. Э. Ортега показывает, что «аналоговые и цифровые архивы... в конечном итоге формируют то знание, которое мы можем из них извлечь»²⁶¹. К. Бод также отмечает важность поисковых алгоритмов в процессе усиления тенденциозности результатов поиска в цифровой коллекции. Если авторы или ключевые слова, к которым чаще обращались пользователи появляются в верхних строчках поисковых результатов, вероятность того, что эти авторы и концепты будут фиксированы в новом цифровом каноне значительно выше²⁶². На эти процессы также влияет политика оцифровки, качество печатной продукции, которая подвергается оцифровке и законодательство, связанное с авторским правом, которое определяет, какие произведения искусства и культуры могут быть опубликованы и представлены широкому пользователю. Тенденциозность представления данных в цифровых коллекциях и агрегаторах оцифрованного контента усугубляется тем, что в проектах массовой оцифровки редакторы цифровых коллекций редко сообщают пользователям, какие принципы стоят за составом коллекции. Иногда отсутствие таких редакторских комментариев обосновано нежеланием коллектива, участвующего в создании проекта демонстрировать ошибки и компромиссы, которые стоят за составом опубликованных данных. Так или иначе до сих пор не разработаны процедуры, которые регулировали бы публикацию редакторских комментариев относительно состава и принципов подбора коллекции²⁶³. Научная литература в области влияния новых медиа на культурологические и гуманитарные исследования и принципы распростране-

²⁶¹ Ortega, É. (2018). Archives, Libraries, Collections, and Databases: A First Look at Digital Literary Studies in Mexico. *Hispanic Review* 86, no. 2 (2018): 229–47.

²⁶² Bode, K. (2020). The Archive, in *The Cambridge Companion to Literature in the Digital Age* (Ed. Adam Hammond). Cambridge: Cambridge University Press.

²⁶³ Там же.

ния и создания знания в области культурного наследия предлагает разные способы анализа такого влияния. Исследования, связанные с анализом цифровой инфраструктуры в области культурного наследия, называют «цифровой библиографией»²⁶⁴, «новой филологией»²⁶⁵, «сравнительными медийными исследованиями»²⁶⁶, «историей цифровой книги»²⁶⁷, «критическими исследованиями цифровой инфраструктуры»²⁶⁸.

Исследования, проделанные в недавние годы, предлагают подробный анализ коллекций, возникших в результате массовой оцифровки²⁶⁹, рассматривая условия их создания, принципы, которые стоят за организацией коллекций и те гипотезы, которые могут сформировать исследователи на их основе²⁷⁰. В этом отношении принципы отбора материалов и условия, определяющие большую или меньшую видимость цифровых материалов в условиях оцифровки и последующей публикации, сходны с процессами отбора и перевода литературного про-

²⁶⁴ Cordell, R. 'Q i-jtb the Raven': Taking Dirty OCR Seriously. *Book History* 20 (2017): 188–225.

²⁶⁵ McGann, J. *A New Republic of Letters: Memory and Scholarship in the Age of Digital Reproduction*. Chicago: University of Chicago Press, 2014.

²⁶⁶ Hayles, N. Katherine, and Jessica Pressman, ed. *Comparative Textual Media: Transforming the Humanities in the Postprint Era*. Minneapolis: University of Minnesota Press, 2013.

²⁶⁷ Daly, Liza, and Whitney Trettien, ed. *Digital Book History*, <https://digitalbookhistory.com/>

²⁶⁸ Liu, A. *Toward Critical Infrastructure Studies: Digital Humanities, New Media Studies, and the Culture of Infrastructure*. University of Connecticut. 23 February, 2017.

²⁶⁹ Fyfe, P. (2016). An Archaeology of Victorian Newspapers. *Victorian Periodicals Review* 49, no. 4 (2016): 546–77; Mak, B. (2014). Archaeology of a Digitization. *Journal of the Association for Information Science and Technology* 65, no. 8 (2014): 1515–26; Warren, C. N. *Historiography's Two Voices: Data Infrastructure and History at Scale in the Oxford Dictionary of National Biography (ODNB)*. *Journal of Cultural Analytics*, November 22, 2018, <http://culturalanalytics.org/article/11031>.

²⁷⁰ Bode, K. (2020). *The Archive*, in *The Cambridge Companion to Literature in the Digital Age* (Ed. Adam Hammond). Cambridge: Cambridge University Press.

изведения для того, чтобы оно получило международную известность.²⁷¹ Так или иначе, цифровые инфраструктуры культурного наследия настолько вписаны в культуру в настоящее время, что известный теоретик новых медиа в области культурного наследия А. Лю утверждает, что культурология в настоящее время является одним из способов критического анализа цифровых инфраструктур²⁷² с целью лучшего понимания культур и систем познания, представленных цифровым архивом, коллекцией или агрегатором. В другой работе А. Лю задает вопрос, возможно ли оценить разные типы репрезентативности в цифровой коллекции (социальную, политическую, лингвистическую, этническую, религиозную и т. д.) и измерить соотношение каждого типа с помощью объективной компьютерной модели, встроенной в систему управления коллекциями²⁷³. Предлагается сформировать подходы к классификациям, связанным с разными типами репрезентативности и порождающими культурное разнообразие. Хотя единая классификационная схема не представляется возможной, А. Лю допускает множественность подходов к классификациям материалов цифровых коллекций культурного наследия и предлагает рассматривать разные точки зрения на репрезентативность и разнообразие культур. При отсутствии необходимой репрезентативности в цифровой коллекции или агрегаторе А. Лю предлагает дополнять цифровую коллекцию материалами, обнаруженными в больших агрегаторах цифровых материалов, представляющих культурное наследие. Те типы трансляции культурного наследия, которые возникнут в результате такого подхода могут быть названы «научными изданиями многообразия культурного наследия»²⁷⁴. Таким образом цифро-

²⁷¹ Warren, C. N. “Historiography’s Two Voices: Data Infrastructure and History at Scale in the Oxford Dictionary of National Biography (ODNB).” *Journal of Cultural Analytics*, November 22, 2018, <http://culturalanalytics.org/article/11031>.

²⁷² Liu, A. “Toward Critical Infrastructure Studies: Digital Humanities, New Media Studies, and the Culture of Infrastructure.” University of Connecticut. 23 February, 2017.

²⁷³ Liu, A. (2020). Toward a Diversity Stack: Digital Humanities and Diversity as Technical Problem. *PMLA*, 135(1), 130–151.

²⁷⁴ Там же, с. 141.

вые технологии можно применять как способ создания коллекций и агрегаторов, отражающих «различия и сходство, неожиданные союзы, объединения, закономерности и связи, которые влияют на наше восприятие (культурной) идентичности»²⁷⁵.

Н. Тахмасеби показывает, что разнообразие и репрезентативность часто рассматриваются как способ избежать редукции в познании мира с помощью данных, представленных цифровыми коллекциями и агрегаторами в области культурного наследия: чем выше репрезентативность, тем ниже степень редукции²⁷⁶. Еще одна сложность применения цифровых технологий к коллекциям культурного наследия заключается в том, что оцифрованный текст или изображение не совпадают с оригинальным текстом или предметом. Применение компьютерных технологий, которые могут работать с узким кругом научных вопросов или гипотез и, как всякое программное обеспечение редуцируют подходы к пониманию моделируемых предметов и явлений²⁷⁷, приводит к следующей степени редукции в восприятии культурных процессов и явлений. Несмотря на то что модель предлагает удобные способы понимания и представления динамики реальных явлений и процессов, происходящих в культуре и обществе, значительная часть особенностей явлений и процессов исчезает при моделировании. Таким образом возникает вопрос, репрезентацией чего является модель, представленная в цифровой коллекции или агрегаторе изображений предметов культурного наследия²⁷⁸. Интерпретация полученных результатов обработки больших массивов данных коллекций в области культурного наследия ведет к подтверждению или опровержению

²⁷⁵ Там же.

²⁷⁶ Tahmasebi, Nina, Niclas Hagen, Daniel Brodén, and Mats Malm. A Convergence of Methodologies: Notes on Data-Intensive Humanities Research. In *Digital Humanities in the Nordic Countries 4th Conference*. Helsinki: Nina Tahmasebi, 2019.

²⁷⁷ Manovich, L. *Software takes command*. New York and London: Bloomsbury Academic, 2013.

²⁷⁸ Tahmasebi, Nina, Niclas Hagen, Daniel Brodén, and Mats Malm. A Convergence of Methodologies: Notes on Data-Intensive Humanities Research. In *Digital Humanities in the Nordic Countries 4th Conference*. Helsinki: Nina Tahmasebi, 2019.

гипотез и это тоже приводит к редукции сложности реального мира. Репрезентативность коллекций и агрегаторов определяется выбором текстов и изображений, которые представляют небольшую часть картины мира, в то время как использование цифровых методов, инструментов и программного обеспечения демонстрирует спектр точек зрения на эту часть реальности. Таким образом, как это часто утверждается в литературе по критическому анализу культурного наследия²⁷⁹, цифровые коллекции и агрегаторы являются свидетельством «онтологических лагун и эпистемических особенностей»²⁸⁰, характеризующих современную эпоху.

Современные цифровые коллекции и агрегаторы изображений в области культурного наследия, такие как цифровая библиотека «Европеана»²⁸¹ или консорциум фотоархивов «Фарос»²⁸² опираются на цифровые инфраструктуры, которые пользуются подходами открытых связанных данных, создавая возможности погружения изображений в ранее неизвестные контексты и формируя закономерности, которые не были исследованы ранее²⁸³. Цифровые коллекции и агрегаторы такого типа пользуются возможностями распознавания изображений и обнаружения одинаковых произведений искусства на разных фотографиях. Таким образом создается возможность исследования репрезентации разных авторов в архивах-участниках консорциума с последующим изучением формирования канона истории искусства в разных географических регионах. Работа с метаданными цифровых коллекций и агрегаторов позволяет проследить временную динамику и/или разницу между классификационными схемами, которые использовались в фото-

²⁷⁹ Smith, L. *The Uses of Heritage*. London: Routledge, 2006; Smith, Laurajane and Natsuko Akagawa (eds) (2009) *Intangible Heritage*. London: Routledge; Gentry, K., & Smith, L. (2019). Critical heritage studies and the legacies of the late-twentieth century heritage canon. *International Journal of Heritage Studies*, 1–21.

²⁸⁰ Bode, K. (2020) Why you can't model away bias. *Modern Language Quarterly* (2020) 81 (1).

²⁸¹ URL: <https://www.europeana.eu/en>

²⁸² URL: <http://pharosartresearch.org/>

²⁸³ Caraffa, C., Pugh, E., Stuber, T., & Ruby, L. (2020). PHAROS: A digital research space for photo archives. *Art Libraries Journal*, 45(1), 2-11.

архивах, расположенных в разных странах и позволяют определить особенности каждой классификационной схемы в сравнении с историческими условиями, в которых она создавалась. Кроме этого компьютерные алгоритмы способны указать и уточнить, какие представления лежат за классификационными схемами и принципами организации коллекций. Таким образом в основе создания агрегатора цифровых изображений в области культурного наследия лежат (классификационные) решения, унаследованные от физического архива и решения, связанные с созданием цифровой коллекции: какие стандарты оцифровки использовать и для каких изображений, как структурировать базу данных, как аннотировать изображения, к каким контекстам предоставлять ссылки, с какими объектами из других фотоархивов соединять изображения.

К. Боуд показывает, что мы никогда не взаимодействуем с культурами прошлого напрямую. Мы имеем дело с «некоторыми формами некоторых сохранившихся документов»²⁸⁴. Более того А. Пайпер утверждает, что моделирование культурных данных подчеркивает особенности конструирования знания и выделяет место наблюдателя внутри конструированных знаний²⁸⁵. В этой же работе А. Пайпер подчеркивает важность будущих исследований, которые рассмотрят, в какой мере результаты количественного анализа (и творческой работы общей публики, использующей части коллекций как следствие формирования нового цифрового канона) зависят от разных типов коллекций, а также взаимодействий между методологическими принципами исследований и принципами составления цифровых коллекций и агрегаторов²⁸⁶.

Поскольку цифровые коллекции и агрегаторы в области культурного наследия создаются, чтобы понять принципы и концепты культур на основе данных, полученных из материальных источников, основная сложность состоит в том, чтобы понять какая доля и какие элементы прошлого используются для конст-

²⁸⁴ Bode, K. (2020) Why you can't model away bias. *Modern Language Quarterly* (2020) 81 (1).

²⁸⁵ Piper, A. *Enumerations: Data and Literary Study*. Chicago: Chicago University Press, 2018.

²⁸⁶ Там же.

руирования культурного наследия и последующего анализа. Таким образом культурологические и гуманитарные исследования, учитывающие эти ограничения, влекут за собой дополнительные сложности по сравнению с работами, которые используют исключительно количественный анализ, поскольку демонстрируют нечеткость и неопределенность культурных процессов вместо того, чтобы представлять очевидные закономерности и тенденции²⁸⁷. Однако до сих пор количественные гуманитарные исследования с использованием цифровых технологий основывались на крупных коллекциях литературных текстов как на репрезентациях всего объема культурного наследия²⁸⁸. Тем не менее даже крупные цифровые коллекции произведений культуры и искусства не демонстрируют весь спектр культурного наследия, поскольку являются свидетельством особенностей коллекционирования и принципов подбора произведений, свойственных физическим коллекциям. Произведения из определенных регионов, произведения, созданные женщинами, а также ряд других произведений культуры и искусства или материальных свидетельств прошлого редко оказываются в физических, а следовательно, и в цифровых коллекциях. Таким образом формируется тенденциозность подбора материалов, основанная на исключении некоторых данных (*exclusion bias*)²⁸⁹. Тенденциозность такого рода окажется системным явлением и будет наблюдаться во многих цифровых коллекциях и агрегаторах произведений культурного наследия. Это значит, работы по анализу данных цифровых коллекций и агрегаторов не могут привести к воссозданию многих культур прошлого и такой анализ сталкивается с серьезными ограничениями.

²⁸⁷ Bode, K. (2020) Why you can't model away bias. *Modern Language Quarterly* (2020) 81 (1).

²⁸⁸ Bode, K. (2020) Why you can't model away bias. *Modern Language Quarterly* (2020) 81 (1); Ortega, É. (2018). Archives, Libraries, Collections, and Databases: A First Look at Digital Literary Studies in Mexico. *Hispanic Review* 86 no. 2: 229–47; Prescott, A. (2014). I'd Rather be a Librarian. *Cultural and Social History* 11 no. 3: 335–41.

²⁸⁹ Bode, K. (2020) Why you can't model away bias. *Modern Language Quarterly* (2020) 81 (1).

Таким образом, анализ основных способов и форм создания коллекций цифровых изображений и функционирования агрегаторов визуального контента в области культурного наследия позволил определить, что основные характеристики цифровых коллекций и агрегаторов включают объем оцифрованного контента и то, в какой мере цифровые коллекции репрезентируют культурное наследие стран, регионов и сообществ. Несомненно, еще одной важной характеристикой цифровой коллекции является возможность повторного бесплатного использования цифровых материалов тех произведений, которые перешли в открытое достояние. Можно утверждать, что репрезентативность цифровой коллекции или агрегатора является характеристикой, которая оказывается подверженной влиянию принципов создания физических коллекций и особенностей создания цифровой коллекции.

Параграф уточняет понятия репрезентативности и сбалансированности в отношении принципов создания цифровых коллекций и особенностей функционирования агрегаторов произведений культурного наследия. Показано, что введение в цифровую коллекцию доли произведений, которая выходит за пределы пропорции, принятой в хрестоматийном каноне произведений искусства и культуры, выводит пользователя за пределы социальных конструкторов западной культуры. Параграф демонстрирует, что логика классификационных схем и способов фильтрации произведений, унаследованная от учреждений культуры доцифровой эпохи или связанная с популярным характером цифровых собраний, оказывается доминирующей в крупных агрегаторах оцифрованного контента в области культурного наследия. Можно утверждать, что ограничения цифровых коллекций и агрегаторов, связанные с репрезентативностью и сбалансированностью коллекций, являются эпистемологическим выбором, принципы которого должны быть продемонстрированы пользователю. Таким образом, способы и формы создания цифровых коллекций и функционирования агрегаторов в области культурного наследия определяют культурные закономерности в современных процессах функционирования собраний цифровых материалов в области культурного наследия, а также влияют на результаты культурологического анализа цифровых коллекций.

2 МЕТОДЫ ОРГАНИЗАЦИИ И ПРЕДСТАВЛЕНИЯ ЦИФРОВЫХ КОЛЛЕКЦИЙ В ОБЛАСТИ КУЛЬТУРНОГО НАСЛЕДИЯ

2.1 Представление российских цифровых коллекций в области культурного наследия

Данный параграф посвящен исследованию представленности российских цифровых коллекций, формирующих современное культурное наследие. Параграф вводит опыт мониторинга представленности европейских цифровых коллекций на примере проекта ENUMERATE, инициированного Европейской Комиссией. Параграф также рассматривает предыдущий опыт исследования представленности российских цифровых коллекций и уточняет особенности влияния представленности цифровых коллекций на формирование нового цифрового канона в области культурного наследия. Вторая глава данного диссертационного исследования основана на применении эмпирических междисциплинарных методов. Для анализа представленности российских цифровых коллекций применяются количественные методы исследования пропорционального соотношения цифровых изображений по отношению к количеству музейных предметов и пропорционального соотношения изображений, опубликованных на сайтах музеев, использованные в европейском проекте ENUMERATE. Анализ современных практик изучения культурного наследия и понимание того, какая его часть будет известна пользователю в современных условиях, зависят от нашего знания представленности российских коллекций в современном информационном пространстве.

Значительный вклад в развитие культурного наследия формируется деятельностью учреждений культуры, в том числе благодаря участию более 15 000

музеев²⁹⁰ в Европе и 1068 музеев в США, аккредитованных Американским музейным альянсом²⁹¹ на момент написания этого исследования. Данные о европейских музеях поступают в Европейскую группу музейной статистики, созданную в 2002 году для того, чтобы собирать и публиковать данные, представленные в соответствии с единым стандартом²⁹². Продолжением и развитием этой деятельности стал проект ENUMERATE²⁹³, направленный на мониторинг представленности цифровых коллекций в области культурного наследия в Европе. Благодаря проекту нам известно процентное соотношение предметов из собраний музеев, библиотек и архивов, для которых существует цифровое изображение, в том числе опубликованное в сети.

Проект ENUMERATE является частью большого европейского проекта Европеана, цель которого – создание европейской цифровой библиотеки, предоставляющей единый доступ к каталогам музеев, библиотек и архивов Европы²⁹⁴. ENUMERATE продолжает работу проекта NUMERIC, который был начат в 2007 году для получения статистических данных о масштабах оцифровки в библиотеках Европы. Опыт проекта был распространен на музеи, библиотеки и архивы Европы проектом ENUMERATE в 2011 году, сбор данных происходил не с помощью министерств и профессиональных ассоциаций, как это делалось Европейской группой музейной статистики и проектом NUMERIC, а с помощью опроса сотрудников музеев, библиотек и архивов. Проект ENUMERATE начал работу в 2011 году и был поддержан Европейской Комиссией. Инициатором работы была

²⁹⁰ J. de Naan (2010). Museum statistics and cultural policy. EGMUS conference: Oslo (2010, oktober 21 – 2010, oktober 22). Данные представлены за 2004–2009 годы, в отчет не включены данные по числу музеев во Франции, Великобритании и Словакии.

²⁹¹ American Alliance of Museums, 2018. URL: <http://ww2.aam-us.org/resources/assessment-programs/accreditation/accredited-museums>

²⁹² European Group on Museum Statistics, 2018. URL: <https://www.egmus.eu/>

²⁹³ URL: <https://pro.europeana.eu/page/enumerate>

²⁹⁴ Petras, V., Hill, T., Stiller, J. and Gäde, M. (2017), Europeana – a search engine for digitized cultural heritage material, *Datenbank Spektrum*, Vol. 17 No. 1, pp. 41-46.

британская организация Collections Trust, созданная в 1977 году с целью разработки стандартов музейной документации. С 2014 года проектом ENUMERATE руководят Collections Trust и DEN Foundation (Голландский институт исследования культуры и цифрового развития)²⁹⁵. Ключевые показатели²⁹⁶, в рамках которых проект ENUMERATE собирает статистические данные включают:

- количество библиотек, музеев и архивов, в которых есть цифровые коллекции (84 %);
- наличие формальной стратегии работы с цифровыми коллекциями, одобренной администрацией учреждения (39 %);
- процентное соотношение оцифрованных коллекций (16 %);
- процентное соотношение коллекций, на которые не распространяется авторское право (общественное достояние), авторское право принадлежит стороннему лицу или организации, принадлежит учреждению, в котором хранятся коллекции, статус коллекций с точки зрения авторского права неизвестен;
- процентное соотношение объектов, для которых в базе данных учреждения содержится описание (метаданные) на уровне отдельного объекта (43 %);
- количество библиотек, музеев и архивов, в которых хранятся цифровые объекты, для которых нет аналогового оригинала и которые были созданы в цифровом виде, так называемые born-digital materials (57 % учреждений);
- процентное соотношение учреждений, публикующих коллекции в сети (32 %)²⁹⁷;

²⁹⁵ URL:<https://pro.europeana.eu/page/about-enumerate>

²⁹⁶ Показатели представлены по данным третьего опроса сотрудников библиотек, музеев и архивов (2015 год).

²⁹⁷ Данные опроса 2017 года (четвертого опроса, проведенного в рамках проекта) показывают рост показателя по сравнению с 2015 годом (дата третьего опроса). В 2017 году процентное соотношение учреждений, публикующих коллекции в сети, по отношению ко всем библиотекам, музеям и архивам, составило 25 %. Предполагается, что эти данные несколько завышены, поскольку ответы на вопросы исследования предоставили учреждения, которые заинтересованы в развитии цифровых коллекций.

– использование цифровых коллекций (библиотеки, музеи и архивы считают самыми важными направлениями научные исследования и использование коллекций в образовательных целях, коммерческое использование и предоставление изображений за деньги считаются европейскими библиотеками, музеями и архивами наименее важными направлениями использования цифровых коллекций);

– количество людей, которые участвуют в работе с цифровыми коллекциями в учреждении (в среднем 8 человек);

– количество музеев, библиотек и архивов, которые следуют международным стандартам долгосрочного хранения цифровых материалов (47 % учреждений работают с цифровыми коллекциями, не учитывая международные стандарты);

– количество цифровых объектов, опубликованных в сети европейскими музеями (7,5 %), библиотеками и архивами.

Проект ENUMERATE предполагает работу с библиотеками, музеями и архивами, которые хранят и изучают объекты культурного наследия для будущих поколений. В 2012 году, когда разрабатывались методологические принципы работы проекта, количество подобных учреждений в Европе составило 40–45 тысяч, в это число вошли 14 577 музеев. Школьные и публичные библиотеки не вошли в спектр исследования, поскольку в этом случае не соблюдался ведущий принцип работы: эти учреждения не соответствуют основному параметру (хранение и изучение объектов культурного наследия для будущих поколений). Исследование охватывало 27 стран Европейского Союза, предполагалось, что результатами опроса воспользуются Европейская Комиссия, правительства стран Европейского Союза, администрация библиотек, музеев, архивов и тех учреждений, где хранятся аудио и видео коллекции, сотрудники этих учреждений, организации, которые разрабатывают коммерческие продукты в области культурного наследия, исследователи, международные организации. Методологической основой проекта стала случайная выборка участников опроса (около 5 % от 42 970 учреждений, всего 1 421 учреждение). Однако авторы отчета по проекту отмечают, что поскольку национальные координаторы во многих странах настаивали на участии в проекте

ведущих национальных библиотек, музеев и архивов, выборка не стала случайной и, следовательно, возможна некоторая тенденциозность, связанная с неизбежными компромиссами²⁹⁸.

Отсутствие количественных данных о цифровых изображениях из коллекций Музейного фонда РФ привело к тому, что в 2014 году методология проекта ENUMERATE была использована в исследовании, проведенном И. Кижнер с соавторами²⁹⁹. В выборку были включены 440 музеев РФ, в 2014 году на адрес музейной администрации были отправлены письма с просьбой участвовать в опросе, опубликованном с помощью сетевой платформы. Адреса музеев были получены через портал «Музеи России»³⁰⁰. База данных портала включала 3 063 музея, характеристики музея включают указание количества посетителей в год. Для того чтобы обеспечить репрезентативность и сбалансированность выборки, музеи были разделены на три категории: большие (свыше 50000 посетителей в год), средние (от 15 000 до 50 000) и малые (менее 15 000 посетителей в год). В выборку вошли 130 музеев, представляющих Москву и Санкт-Петербург, 310 музеев, представляющих субъекты федерации (3–6 музеев от каждого субъекта федерации, представляющих большие, средние и малые музеи). Полученная выборка при условии достаточного количества ответов дала бы адекватное представление о масштабах оцифровки в Российской Федерации.

Ответы были получены от 6 % музеев в выборке. Это намного меньше 30 % учреждений, которые приняли участие в опросе о результатах оцифровки в биб-

²⁹⁸ Nauta, G. J. and de Niet, M. (2014). ENUMERATE: Conceptual Framework, April 2014, ENUMERATE, Europeana Pro. URL: https://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/ENUMERATE/deliverables/ENUMERATE-D2-12.pdf

²⁹⁹ Kizhner I., Terras M., Rumyantsev M. (2016). Museum digitization practices across Russia: Survey and web site exploration results. In Digital Humanities 2016: Conference Abstracts . Jagiellonian University & Pedagogical University, Kraków, pp. 600–2.

³⁰⁰ URL: <http://www.museum.ru/>

лиотеках США³⁰¹, и 51 % учреждений, которые дали ответы на вопросы о масштабах оцифровки в музеях, библиотеках и архивах Европы³⁰². Результаты опроса показали, что количество оцифрованных предметов Музейного фонда РФ составило в 2014 году 18,3 % Музейного фонда РФ, а количество предметов, изображения которых были опубликованы на музейных сайтах, составило 1–2 %. Данные о количестве изображений сравнивались с данными о количестве музейных предметов, опубликованными на сайте «Музеи России». В следующем исследовании 2016 года были использованы данные Министерства культуры РФ, были продемонстрированы средние данные о масштабах оцифровки музейных коллекций РФ для 2 300 музеев, показана разница между федеральными округами и субъектами федерации, выявлено количество изображений, опубликованных на музейных сайтах³⁰³. Эти результаты приводят к пониманию того, каким образом современные визуальные практики будут влиять на формирование нового культурного канона, какие произведения будут представлены в новом каноне и какие географические регионы повлияют на его формирование.

Особенности формирования нового культурного канона отражены в работах, опубликованных в последнее десятилетие двадцатого века и в первые годы нашего века³⁰⁴. Так, Питер Лим демонстрирует необходимость участия ученых

³⁰¹ Boock, M. and R. Vondracek. (2006). Organizing for Digitization: A Survey, *portal: Libraries and the Academy*, 6(2): 197-217.

³⁰² Poll, R. (2010). NUMERIC: statistics for the digitization of European cultural heritage, *Electronic Library and Information Systems*, 44(2): 122-31.

³⁰³ Kizhner, I., Terras, M., Rummyantsev, M., Sycheva, K., Rudov, I. (2018). Accessing Russian Culture Online; the scope of digitization in museums across Russia. *Digital Scholarship in the Humanities*. 19 September 2018.

³⁰⁴ Limb, P. (2007). The politics of digital" reform and revolution: towards mainstreaming and African control of African digitisation. *Innovation*, 34; Price K. M. (2009). Digital scholarship, economics, and the American literary canon. *Literature Compass*, 6(2): 274–90; Earhart A. (2012). Can information be unfettered? Race and the new digital humanities canon. In *Debates in the Digital Humanities*. Minneapolis, MN: University of Minnesota, pp. 309–18; Warwick C., Terras M., Nyhan J.

и сотрудников библиотек разных стран и континентов в процессе расширения культурного канона³⁰⁵. Кеннет Прайс показывает, что формирование нового культурного канона оказывается под влиянием тех художественных произведений и литературоведческих исследований, которые не касаются работ, созданных в двадцатом веке. Это происходит в связи с ограничениями обнаружения этих работ в свободном доступе, связанными с авторским правом. Новые художественные произведения не могут войти в канон, так как находятся в университетских библиотеках, доступ к которым ограничен для исследователей, университеты которых не получают достаточного финансирования³⁰⁶.

Музейные предметы, цифровые изображения которых оказываются основой нового цифрового канона, делают культуры, в большей степени представленные в информационном пространстве, основой последующего анализа. Таким образом, выводы гуманитарных наук касаются только тех данных, которые станут основой культурного канона, сформированного на основе новых визуальных практик.

Анализ того, как культурное наследие представлено в сетевой среде, представляет собой первый этап исследований эволюции культурного канона в новой цифровой среде. Визуальные практики представления культурного наследия в России могут быть исследованы с помощью анализа количества оцифрованных произведений и изображений, опубликованных в Интернете, изучения истории оцифровки культурного наследия в России, оценки юридических особенностей повторного использования и тиражирования изображений музейных предметов, важности многоязычных интерфейсов и метаданных и разницы между уровнем оцифровки в столичных центрах и в провинции.

(2012). *Digital Humanities in Practice*. London: Facet Publishing in association with UCL Digital Humanities Centre.

³⁰⁵ Limb, P. (2007). The politics of digital" reform and revolution: towards mainstreaming and African control of African digitisation. *Innovation*, 34

³⁰⁶ Price, K. M. (2009). Digital scholarship, economics, and the American literary canon. *Literature Compass*, 6(2): 274–90

Представления о повторном использовании изображений, открытых данных, связанных данных и использовании изображений в проектах из другой области приводят исследователей к мысли о том, что любой цифровой проект в гуманитарных науках не должен быть привязан к тому интерфейсу, где впервые были представлены данные. Научное сообщество, поддерживающее политику открытого доступа, рекомендует повторное использование изображений в разнообразных проектах в рамках лицензий, созданных сообществом, таких, например, как лицензии Creative Commons³⁰⁷.

Цифровые изображения и процесс оцифровки коллекций музеев, библиотек и архивов влияют на эволюцию культурного канона, затрагивая социальные аспекты репрезентации культурного наследия, участия разных культур в создании нового цифрового канона. Недооценка создания и публикации цифровых изображений, распространения их с помощью новых медиа и объединения изображений с помощью национальных и международных агрегаторов ведет к недостаточному пониманию новых визуальных практик как социальных и технологических инструментов³⁰⁸.

В настоящем диссертационном исследовании представлены данные о представленности российских цифровых коллекций, формирующих современное культурное наследие в сравнении с международными практиками представленности цифровых коллекций в области культурного наследия. Показаны сложности, возможности и последствия публикации российских цифровых коллекций музейных предметов в более широком европейском контексте. Очевидно, что те собрания цифровых изображений, которые опубликованы в Интернете и сопровождаются лицензиями, которые предоставляют возможность повторного использования изображения, окажутся в фокусе гуманитарных исследований. Это может

³⁰⁷ Robinson, P. (2013). Five desiderata for scholarly editions in digital form. In *Digital Humanities 2013: Conference Abstracts*. University of Nebraska-Lincoln, 16-19 July 2013.

³⁰⁸ Gooding P., Terras M., Warwick C. (2013). The myth of the new: mass digitization, distant reading and the future of the book. *Literary and Linguistic Computing*, 28(4): 629–39.

привести к тому, что исследования российской культуры и культур некоторых других регионов будут осложнены рядом ограничений и препятствий.

Росс Парри указывает на то, что создание цифровых коллекций в музеях является непростой задачей, которая осложняется разнообразием точек зрения на цели оцифровки³⁰⁹. Широкий спектр целей создания цифровых коллекций включает задачи учета музейных предметов, обеспечение доступа к коллекциям широкой публике, представление точки зрения хранителей и научных сотрудников (при этом научные описания музейного предмета могут не включать ряд полей учетной базы данных, но содержать информацию, которая выходит за пределы учетных целей), необходимость введения стандарта описания музейного предмета для машинной обработки³¹⁰. Анализ представленности российских цифровых коллекций показывает, что создатели цифровых музейных коллекций в России преследовали цели музейного учета, а не цели предоставления изображений для повторного использования в ресурсах гуманитарного характера и тем более не цели создания цифровых научных изданий³¹¹.

В России, как и многих других странах, основным мотивом деятельности по созданию цифровых коллекций была необходимость учета и обеспечения сохранности музейных предметов³¹². Создание машинных каталогов началось

³⁰⁹ Parry, R. *Recoding the Museum: Digital Heritage and the Technologies of Change*. London: Routledge, 2007.

³¹⁰ *Ibid.* p. 20.

³¹¹ Создание цифровых научных изданий является сложной задачей, которая требует высших редакторских навыков и глубоких знаний в области истории книгопечатания (McGann, 2013). Недавнее исследование показывает, что в мире существует не больше 300 цифровых научных изданий (Franzini et al., 2015).

³¹² Aseev Y., Sher J. (1983). Preface from the editors of the Russian edition. In Chenhall R. (ed.), *Museum Cataloging in the Computing Age*. Moscow: Mir, pp. 7–17; Chenhall R., Vance D. (2010). The world of (almost) unique objects. In Parry R. (ed.), *Museums in a Digital Age*. London; New York: Routledge, pp. 39–48; Williams D. (2010). A brief history of museum computerization. In Parry R. (ed.), *Museums in a Digital Age*. London, New York: Routledge, pp. 15–22; Navarette T. (2014). A

в 1970-е годы, а первые попытки предоставления доступа к музейным коллекциям относятся к концу XX и началу XXI века. Одна из первых важных инициатив по созданию новых визуальных практик и предоставлению доступа к музейным коллекциям относится к 1997 году, когда Государственный Эрмитаж³¹³ и коммерческая компания, работающая в области информационных технологий, International Business Machines (IBM) начали сотрудничать с целью демонстрации коллекций Эрмитажа в сети Интернет. Компания IBM³¹⁴ предоставила сканер, редкое и дорогое оборудование для тех лет, программное обеспечение, приложение для публикации данных в сети и дизайн интерфейса музейного сайта, опубликованного в 1999 году. Государственный Эрмитаж обладал уникальными возможностями создания цифровой коллекции, предназначенной для публикации, поскольку сочетал компетенции музейных научных сотрудников, способных предоставить качественное описание музейных предметов, высококачественные технологии оцифровки, предоставленные другой коммерческой компанией, и технологию IBM, предназначенную для разработки сайта. Взаимодействие такого крупного музея с коммерческой компанией было вполне типичным для первых этапов оцифровки в 1990-е годы. В тот период музеи получали большую пользу от крупномасштабного применения коммерческих технологий, а компании могли проводить экспериментальные разработки и зарабатывать репутацию, пользуясь известностью и достижениями проекта³¹⁵.

Необходимость преследовать две цели (учет музейных фондов и предоставление доступа к коллекциям) привела к созданию Государственного каталога Му-

History of Digitization: Dutch Museums. Ph.D. thesis, University of Amsterdam. <http://catalogus.boekman.nl/pub/P14-0752.pdf> (accessed 8 December 2017).

³¹³ URL: <https://www.hermitagemuseum.org/wps/portal/hermitage/>

³¹⁴ IBM. (2017). Hermitage museum project. URL: <https://www.research.ibm.com/haifa/projects/software/hermitage/> (accessed 12 December 2017).

³¹⁵ Terras M. (2011). The rise of digitization. In Rikowski R. (ed.), *Digitization Perspectives*. Rotterdam, Boston, Taipei: Sense Publishers, pp. 3–20.

зейного фонда Российской Федерации (Госкаталога МФ РФ)³¹⁶. Идея о необходимости создания Государственного каталога Музейного фонда была представлена в Федеральном законе «О музейном фонде в Российской Федерации и о музеях в Российской Федерации» в 1996 году. Государственный каталог создавался в первую очередь для учетных целей. Министерство культуры РФ сообщило о наличии нескольких реестров Госкаталога, в том числе о наличии реестра, который является «публичным порталом», предназначенным для ознакомления с музейными предметами в 2017 году³¹⁷. Предполагается, что наполнение Госкаталога будет завершено к 2026 году, когда все музейные предметы основного фонда (около 60 миллионов музейных предметов) будут включены в Госкаталог, а метаданные и изображения музейных предметов будут включены в реестр и опубликованы онлайн. Предоставление данных о государственной части Музейного фонда обязательно для всех государственных музеев. Предоставление и выполнение планов-графиков наполнения Государственного каталога контролирует Министерство культуры РФ³¹⁸. В Госкаталог МФ РФ входят три реестра. Первый реестр является обязательным для наполнения Реестром музеев, подведомственных Министерству культуры РФ, частные музеи могут быть включены в Реестр музеев на добровольной основе. В Реестр музеев вносят сведения о форме собственности, месте нахождения музея, органе государственной власти, который является учредителем музея, типе и профиле музея, адрес электронной почты, изображение

³¹⁶ Федеральный закон №54-ФЗ «О музейном фонде в Российской Федерации и о музеях в Российской Федерации, принят Государственной Думой 24 апреля 1996 года.

³¹⁷ Министерство культуры РФ, «Глава Минкультуры России призвал активнее заносить предметы музейного фонда в Госкаталог Российской Федерации», 9 июля 2017. URL: https://www.mkrf.ru/press/news/glava-minkultury-rossii-prizval-aktivnee-zanosit-p/?sphrase_id=2327199

³¹⁸ Министерство культуры РФ, «О предоставлении планов-графиков регистрации предметов в Государственном каталоге Музейного фонда РФ», 29 июня 2017. URL: <https://goskatalog.ru/portal/#/for-museums/docs?id=175>

здания, сведения об общем количестве музейных предметов и прочие данные³¹⁹. Второй реестр – Реестр музейных предметов, с помощью которого можно проследить за перемещением предмета в ситуации выставочной деятельности или реставрационных работ. Третий реестр дает возможность опубликовать изображение и описание музейного предмета в рамках публичного портала Госкаталога в просветительских целях³²⁰.

Руководство пользователя, представленное на сайте Госкаталога извещает музейных сотрудников о том, какую информацию нужно будет внести в учетную карточку музейного предмета³²¹. Помимо регистрационных и учетных данных список включает следующие поля: изображение предмета, автор/мастер, организация-изготовитель, место создания, место обнаружения, материал, техника, размеры, масса, способ поступления, источник поступления, время создания и некоторые другие данные.

Данное диссертационное исследование представляет динамику наполнения Госкаталога МФ РФ с 2016 года до 2020 года. Данные получены на сайте Госкаталога. На момент написания этой работы Госкаталог включает около 30 % основного фонда музейных предметов из государственной части музейных коллекций РФ (Таблица 1).

Таблица 1 – Динамика количества учетных записей с цифровыми изображениями, опубликованными на сайте Государственного каталога МФ РФ

³¹⁹ Министерство культуры РФ, «О регистрации в Реестре музеев Государственного каталога Музейного фонда Российской Федерации», 6 декабря 2017. URL: <https://goskatalog.ru/portal/#/for-museums/docs?id=185>

³²⁰ Министерство культуры РФ, «Глава Минкультуры России призвал активнее заносить предметы музейного фонда в Госкаталог Российской Федерации», 9 июля 2017. URL: https://www.mkrf.ru/press/news/glava-minkultury-rossii-prizval-aktivnee-zanosit-p/?sphrase_id=2327199

³²¹ Министерство культуры РФ «Руководство пользователя – музейного работника. pdf», 24 декабря 2018. URL: <https://goskatalog.ru/portal/#/for-museums/docs?id=227>

| Дата | Количество учетных записей с изображениями в Госкаталоге | Процентное соотношение по отношению к количеству музейных предметов в основном фонде Музейного фонда РФ (59,872,401 музейных предметов, данные на 2015 год) |
|--------------|--|---|
| Июль 2016 | 950,256 | 1.5 |
| Ноябрь 2016 | 1,200,000 | 2.0 |
| Январь 2017 | 1,496,317 | 2.5 |
| Май 2017 | 2,010,552 | 3.3 |
| Август 2017 | 2, 781,661 | 4.6 |
| Октябрь 2017 | 3,411,119 | 5.7 |
| Октябрь 2018 | 7,719,095 | 12.9 |
| Январь 2019 | 9,280,221 | 15.5 |
| Апрель 2019 | 10, 070,868 | 16,7 |
| Июнь 2020 | 18,657,517 | 31 |

Следующий вопрос, который стоит задать в рамках данного диссертационного исследования – как представлены предметы культурного наследия из Музейного фонда РФ в международных агрегаторах, демонстрирующих предметы из музеев, библиотек и архивов, таких как Европеана³²² и Google Arts and Culture³²³. В 2008–2009 гг пять музеев РФ³²⁴ выразили интерес к участию в цифровой

³²² На момент написания этого исследования в цифровая библиотека «Европеана» содержит 58 миллионов учетных записей, представляющих культурное наследие музеев, архивов и библиотек Европы.

³²³ URL: <https://artsandculture.google.com/> Google Arts and Culture – цифровая платформа, которая включает изображения объектов культурного наследия, в том числе музейных предметов.

библиотеке Европеана³²⁵. Между 2009 и 2011 годом Европеана опубликовала метаданные для 43839 предметов из этих музеев³²⁶. В момент написания этой диссертационной работы Европеана включает 404 учетных записи, предоставленные российскими музеями³²⁷, при этом изображение музейного предмета не предоставляется. Google Arts and Culture предоставляет доступ к изображениям и метаданным для 5 520 музейных предметов из российских коллекций, в том числе из частных музеев. Для того чтобы определить направления и мотивы создания цифровых коллекций в российских музеях в исторической перспективе, необходимы дальнейшие исследования. Необходимо выяснить причины почти полного отсутствия изображений российских предметов культурного наследия в международных агрегаторах. Следующий параграф второй главы диссертационного исследования продемонстрирует, что доступ к изображениям и метаданным с помощью сайтов российских музеев осложнен тем, что российские музеи представляют незначительное количество изображений на своих сайтах, и ограничениями, связанными с повторным использованием изображений. Это значит, что изображения объектов культурного наследия и музейных предметов не могут быть ис-

Проект был инициирован компанией Google в 2011 году для предоставления доступа к изображениям высокого качества (в высоком разрешении).

³²⁴ Государственная Третьяковская галерея, URL: <https://www.tretyakovgallery.ru/>, Саратовский государственный художественный музей имени А.Н. Радищева, URL: <http://radmuseumart.ru/>, Рыбинский музей-заповедник, URL: <https://www.rybmuseum.ru/ru/>, Чувашский государственный художественный музей, URL: <http://www.artmuseum.ru/page/about/>, Музей истории Казанского университета, URL: <https://kpfu.ru/museum-of-history>.

³²⁵ Браккер, Н.В. Российские учреждения культуры в проекте ATHENA // 12-ая ежегодная международная конференция EVA 2009 Москва. Москва, 30 ноября – 2 декабря 2009. URL: http://conf.evarussia.ru/eva2009/rus/reports/report_1774.html

³²⁶ Браккер, Н.В., Куйбышев Л.А. Европейская цифровая библиотека Europeana и российские учреждения культуры // ИФЛА. 2013. №1, (100).

³²⁷ Саратовский государственный художественный музей имени А.Н. Радищева, URL: <http://radmuseumart.ru/>, Государственный музей Л. Н. Толстого, URL: <http://tolstoymuseum.ru/>, Музей-диорама «Курская битва. Белгородское направление».

пользованы в научных и просветительских целях до конца 2025 года, когда, как предполагается, изображения 100% предметов Основного фонда музейных предметов Российской Федерации будут опубликованы в Госкаталоге МФ РФ. Эти обстоятельства являются важными для понимания того, как развивается гуманитарная наука и какие коллекции являются важными для научных исследований и формирования международных представлений о культурном каноне.

Работа с Госкаталогом МФ РФ является начальным этапом в понимании масштабов оцифровки музейных коллекций Российской Федерации, включая те коллекции, которые находятся в отдаленных регионах страны. В одной из предыдущих работ³²⁸ диссертанта было показано процентное соотношение предметов с цифровыми изображениями в российских музейных коллекциях в 2014 году. Работа также включала результаты исследования музейных сайтов и показывала процентное соотношение изображений предметов, опубликованных в Интернете. Мы получили ответы от 1,2 % российских музеев, 6 % музеев вошли в выборку исследования музейных сайтов. В результате были получены предварительные показатели, которые продемонстрировали, что в этот период масштабы оцифровки были ниже, чем в Европе – 18 % музейных коллекций по сравнению с 31 % аналоговых коллекций в Европе в 2015 году³²⁹. Процентное соотношение изображений, опубликованных онлайн с помощью музейных сайтов, было низким (1,5 %), но сопоставимым с европейскими показателями (7 %)³³⁰. Было проведено исследование масштабов оцифровки в стране, где существуют коллекции музейных предметов, представляющих разнообразие культурного и этнического насле-

³²⁸ Кижнер И. А., Террас, М., Румянцев М. В. Практики музейной оцифровки в России: результаты опроса и исследования музейных сайтов // Тезисы конференции Digital Humanities 2016. Ягеллонский университет, Педагогический университет им. Комиссии народного образования в Кракове. Краков, 2016.

³²⁹ Nauta, J. G., van den Heuvel, W. (2015). Survey report on digitization in European Cultural Institutions 2015. Europeana Pro. <http://pro.europeana.eu/enumerate/statistics/results>

³³⁰ Nauta, J. G., van den Heuvel, W. (2015). Survey report on digitization in European Cultural Institutions 2015. Europeana Pro. <http://pro.europeana.eu/enumerate/statistics/results>

дия. Ограничением исследования была маленькая выборка, к тому же мы не рассматривали качество коллекций, важность музейных объектов для гуманитарных исследований и качество цифровых изображений.

Нормы, связанные с работой с культурным наследием таковы, что общество предполагает, что пользователи будут запрашивать официальное письменное разрешение администрации музея в ситуации взаимодействия с визуальным искусством в цифровой среде³³¹. Степень свободы этого взаимодействия бывает разной в разных странах. Так, например, в одной из работ обсуждаются ограничения применения концепции «добросовестного использования» в США³³², в то время как в другой работе авторы предлагают реальные примеры ограничений использования изображений художественных произведений из коллекций музеев разных стран³³³. В России право настаивать на письменном разрешении администрации в случае использования изображений предметов музейных коллекций поддерживается законодательством РФ³³⁴. Это значит, что федеральное Министерство

³³¹ Bielstein, S. (2006). *Permissions, A Survival Guide: Blunt Talk about Art as Intellectual Property*. Chicago, IL: The University of Chicago Press; Whalen, M. (2009). *What's wrong with this picture? An examination of art historians' attitudes about electronic publishing opportunities and the consequences of their continuing love affair with print*. *Art Documentation: Bulletin of the Art Libraries Society of North America*, 28(2): 13–22; Petri G. (2014). *The public domain vs the museum: the limits of copyright and reproductions of two-dimensional works*. *Journal of Conservation and Museum Studies*, 12(1). <http://www.jcms-journal.com/articles/10.5334/jcms.1021217/print/>; Aufderheide P., Milosevic T., Bello B. (2016), *The impact of copyright permissions culture on the US visual arts community: the consequences of fear of fair use*. *New Media and Society*, 8(9): 2012–27.

³³² Aufderheide, P., Milosevic, T., Bello, B. (2016), *The impact of copyright permissions culture on the US visual arts community: the consequences of fear of fair use*. *New Media and Society*, 8(9): 2012–27.

³³³ Wallace, A., Deazley, R. (2016). *Display at your own risk: an experimental exhibition of digital cultural heritage*. URL: <http://displayatyourownrisk.org/publications/>

³³⁴ Федеральный закон «о Музейном фонде в Российской Федерации и музеях в Российской Федерации», принят Государственной Думой 24 апреля 1996 года. Статья 36 ФЗ сообщает, что «производство изобразительной, печатной, сувенирной и другой тиражированной продукции

культуры и региональные Министерства культуры РФ поддерживают музеи в праве требовать запроса разрешения при использовании изображений организациями и частными лицами. Например, Государственный Эрмитаж разрешает использование изображений в школьных и студенческих проектах, лекциях, диссертациях, выступлениях на конференциях. Однако публикация слайдов выступления на конференции с использованием изображений из коллекций Государственного Эрмитажа запрещается и считается нарушением законодательства. В этом случае необходимо обращаться с запросом на имя директора Государственного Эрмитажа и получать лицензию на изображение в том же порядке, в котором получают лицензию на коммерческое издание. Государственный Эрмитаж принимает решение о том, взимать или не взимать плату за изображение в каждом случае отдельно³³⁵. В одном из предыдущих исследований диссертанта было показано, что в случае, когда разрешение получено, изображение может использоваться в единственном проекте, его нельзя переносить на другую платформу или использовать с иными исследовательскими целями. Смена целей использования повлечет за собой составление нового договора и получение новой лицензии³³⁶.

Российские музеи не являются исключением из общей тенденции создания закрытых музейных коллекций. Недавнее исследование показало, что около 80 %

и товаров народного потребления с использованием изображений музейных предметов и музейных коллекций, зданий музеев, объектов, расположенных на территориях музеев, а также с использованием их названий и символики осуществляется с разрешения дирекций музеев». Второй закон, который регулирует использование изображений предметов культурного наследия - «Основы законодательства РФ о культуре» в редакции от 5 декабря 2017 года. Статья 53 «Основ законодательства РФ о культуре» разрешает воспроизводить объектов культуры и культурного достояния с разрешения владельцев.

³³⁵ URL: https://www.hermitagemuseum.org/wps/portal/hermitage/about/image_usage_policy

³³⁶ Kizhner, I., Stankevich, J., Rumyantsev, M., Makarchuk, I. (2016). Licensing images from Russian museums for an academic project. *Journal of Siberian Federal University. Humanities and Social Sciences*, 7(9), 1618–30. URL: http://elib.sfu-kras.ru/bitstream/handle/2311/20416/09_Kizhner.pdf?sequence=1

музеев в выборке из 175 учреждений из англоговорящих стран (США, Великобритания, Канада, Австралия и Новая Зеландия) требуют получения формального запроса при взаимодействии пользователей с произведениями визуального искусства³³⁷.

Недавнее исследование репутации художественных музеев³³⁸ включало 18 знаменитых музеев, представляющих разные страны и континенты. Одним из критериев оценки была возможность узнать о существовании музея. Диссертантом было проведено исследование документации музеев, для того чтобы выяснить, существует ли возможность повторного использования изображений, опубликованных на музейном сайте. Документы опубликованы на сайтах музеев и свидетельствуют о том, что две трети музеев не предоставляют открытый доступ (возможность беспрепятственного копирования изображений музейных предметов) к своим коллекциям.

Это показывает, что российские музеи не являются единственными учреждениями, которые препятствуют тому, чтобы изображения из их коллекций появились в научном обращении или способствовали созданию нового визуального канона.

Предполагается, что наполнение Государственного каталога Музейного фонда РФ будет завершено к 2026 году³³⁹. Вероятнее всего, это произойдет в рамках запланированного периода. Если это случится и если правила использования

³³⁷ Есалиева, Сандугаш. Правовое сопровождение цифровых коллекций произведений искусства. Выпускная квалификационная работа. Сибирский федеральный университет, Красноярск, 2017.

³³⁸ Van Riel, C., Heijndijk, P. (2017). Why people love art museums: a reputation study about the 18 most famous museums among visitors in 10 countries. Rotterdam School of Management, Erasmus University.

³³⁹ Министерство культуры РФ, «Глава Минкультуры России призвал активнее заносить предметы музейного фонда в Госкаталог Российской Федерации», 9 июля 2017. URL: https://www.mkrf.ru/press/news/glava-minkultury-rossii-prizval-aktivnee-zanosit-p/?sphrase_id=2327199

цифровых изображений для произведений, которые перешли в общественное достояние, изменятся и позволят использовать цифровые изображения в иных проектах, тогда российское культурное наследие будет доступно широкой международной аудитории. Если этого не произойдет, российское культурное наследие не будет адекватным образом представлено в информационном пространстве, и это приведет к недостаточному знанию о культурном наследии страны в глобальном масштабе в век, когда страны конкурируют за внимание пользователей с помощью цифровых каналов коммуникации.

Таким образом, в первом параграфе второй главы была рассмотрена представленность российских цифровых коллекций, формирующих культурное наследие. В параграфе уточняются ограничения, связанные с формами представленности российских цифровых коллекций, такими как представленность коллекций с помощью сайтов музеев, публикация цифровых изображений с помощью Государственного каталога Музейного фонда РФ, продемонстрированы юридические ограничения, препятствующие включению культурного наследия из российских музейных коллекций в новый цифровой канон произведений искусства и культуры. Данный параграф дополняет первую главу диссертационного исследования, показывая, что анализ культурного наследия в будущем возможен благодаря публикации изображений произведений культуры и искусства с помощью цифровых инфраструктур и современных способов представления культурного наследия. При этом возможность повторного использования цифровых изображений усиливает вероятность того, что культурное наследие, представленное в цифровых коллекциях, войдет в новый цифровой канон, окажется в учебных программах, научных исследованиях и окажет влияние на развитие представлений о культуре.

2.2 Методы организации агрегаторов оцифрованного контента в области культурного наследия (на примере цифровых коллекций Музейного фонда РФ и платформы Google Arts and Culture)

Данный параграф посвящен эмпирическому количественному исследованию цифровых коллекций Музейного фонда РФ и цифрового агрегатора Google Arts and Culture с целью определения складывающихся культурных закономерностей в современных процессах формирования цифровых коллекций и агрегаторов в области культурного наследия. В предыдущем параграфе были рассмотрены количественные исследования проекта ENUMERATE, которые составили методологическую основу исследования цифровых коллекций Музейного фонда РФ. Данный параграф сравнивает результаты представления цифровых коллекций учреждений культуры Европы и российских учреждений культуры в 2015 году. Показано, что цифровые коллекции Музейного фонда РФ создавались в первую очередь с целью учета и инвентаризации Музейного фонда. Параграф демонстрирует, что доля оцифрованных изображений, опубликованных на сайтах музеев РФ сравнима с долей изображений, представленных в Европе, но большая часть российского контента не годится для повторного использования из-за ограничений, связанных с соблюдением законодательства РФ. Вторая часть параграфа оценивает цифровые коллекции агрегатора изображений предметов культурного наследия Google Arts and Culture. Показано, что этот цифровой агрегатор действительно переживает период устойчивого роста с точки зрения количества изображений и учреждений культуры, которые принимают участие в развитии проекта. Продемонстрировано, что особенностью агрегатора является отсутствие сбалансированности коллекций разных стран и культур, при этом отдельные страны и учреждения доминируют в агрегаторе. Выявлено, что цифровой анализ культурного наследия оказывается возможен только благодаря тем учреждениям культуры, которые публикуют большое количество изображений, дающих представление об их коллекциях и при этом используют открытые лицензии для предоставления бесплатного по-

вторного использования изображений. Такие массивы больших данных могут быть крайне редки за пределами ограниченного количества западных цифровых коллекций, и это может привести к тому, что ряд культур и сообществ будут исключены из исследований больших данных в гуманитарных науках.

Параграф включает две части. В первой части эмпирического исследования будут рассмотрены цифровые коллекции Музейного фонда РФ. Во второй части исследования будет представлен анализ агрегатора цифровых изображений Google Arts and Culture.

Целью эмпирического количественного исследования цифровых коллекций РФ было сравнение данных, полученных в результате анализа российских музейных коллекций, с данными проекта ENUMERATE. В рамках проекта ENUMERATE были получены показатели оцифровки музейных, библиотечных и архивных коллекций в Европе между 2011 и 2015 годами. Сравнивая данные оцифровки музейных коллекций в России с данными исследования ENUMERATE, можно было выяснить, действительно ли масштаб оцифровки коллекций в России соответствует объему подобной деятельности в Европе. Мы использовали данные исследования ENUMERATE за 2015 год³⁴⁰. В опросе участвовали 355 музеев из 20 стран Европы.

Исследование проводилось во второй половине 2016 года и весной 2017 года на кафедре информационных технологий в креативных и культурных индустриях Гуманитарного института СФУ³⁴¹.

Для того чтобы определить количественное соотношение цифровых изображений в российских музейных коллекциях, был рассмотрен масштаб оцифровки музейных коллекций РФ, используя статистическую отчетность (Форма N

³⁴⁰ Nauta J. G., van den Heuvel W. (2015). Survey report on digitization in European Cultural Institutions 2015. Europeana Pro. <http://pro.europeana.eu/enumerate/statistics/results>

³⁴¹ Inna Kizhner, Melissa Terras, Maxim Rumyantsev, Kristina Sycheva, Ivan Rudov, Accessing Russian culture online: The scope of digitization in museums across Russia, *Digital Scholarship in the Humanities*, Volume 34, Issue 2, June 2019, Pages 350–367.

8-НК)³⁴², которая была отправлена в Министерство культуры РФ в 2015 году. Отчет предоставили 2367 музеев. В результате анализа статистических отчетов были получены средние показатели по стране и средние показатели по восьми федеральным округам. Таким образом можно выяснить географическое распределение и сравнить показатели по регионам.

Данные статистической отчетности Министерства культуры о деятельности музеев РФ в 2015 году были получены летом 2016 года по запросу, поступившему в Министерство культуры РФ из приемной проектора по учебной работе Сибирского федерального университета. Данные были получены в виде таблицы формата Excel, которая представляла ответы администрации 2635 музеев РФ на вопросы Формы № 8-НК. Насколько нам известно, такие подробные данные ранее никогда не использовались, чтобы изучать масштабы оцифровки на региональном и международном уровнях.

Из полученной таблицы была удалена информация, которая не имела отношения к процессам создания цифровых коллекций или информация о галереях, предназначенных для временных выставок. В результате были получены данные о цифровых коллекциях 2 367 музеев. Были проведены анализ и сравнение данных о количестве предметов в каждом музее, количестве учетных записей, внесенных в базу данных музея, где присутствовали цифровые изображения музейного предмета, количество учетных записей, для которых изображения опубликованы на сайте музея и наличие интерфейса музейного сайта на английском языке (в последнем случае данные были получены с помощью исследования музейных сайтов всех музеев, которые вошли в таблицу).

Процентное соотношение предметов с цифровыми изображениями по сравнению с общим количеством музейных предметов составило 14 %. Это немного по сравнению со средним количеством оцифрованных музейных предметов в Ев-

³⁴² Министерство культуры РФ ввело форму статистической отчетности для музеев (Форма N 8-НК) в 2003 году. Форма N 8-НК доступна на сайте «Статистическая отчетность отрасли» ГИВЦ Министерства культуры РФ. URL: <https://stat.mkrf.ru/forms/>

ропе, где, по данным опроса ENUMERATE, этот показатель составил 31 % в 2015 году³⁴³. Масштаб оцифровки музейных коллекций оказался очень разным в зависимости от географического региона, при этом самые низкие показатели были продемонстрированы в Дальневосточном федеральном округе, Приволжском федеральном округе и Северо-Кавказском федеральном округе. Самый высокий уровень оцифровки был отмечен в Санкт-Петербурге. Масштаб оцифровки был совершенно разным в федеральных округах, при этом минимальный уровень (6 % аналоговых коллекций) был отмечен в Дальневосточном федеральном округе, а максимальный уровень (25 %) в Северо-Западном федеральном округе близ Санкт-Петербурга. Этот разрыв может привести к тому, что коллекции провинциальных музеев будут недостаточно представлены даже тогда, когда Госкаталог будет почти полностью заполнен в 2026 году (музейные предметы отдаленных округов не будут оцифрованы в полной мере).

Отчет о состоянии цифровых коллекций музеев, библиотек и архивов Европы³⁴⁴ показывает, что сотрудники культурных организаций считают, что оцифрованы должны быть 86% коллекций. Это значит, что в случае европейских музеев и музеев Санкт-Петербурга в 2015 году была оцифрована треть коллекций. В Сибири, Приволжском федеральном округе и Дальневосточном федеральном округе цифровые коллекции включали значительно меньшую долю музейных предметов.

Сравнительный анализ цифровых коллекций Москвы и Санкт-Петербурга привел к интересному и неожиданному результату. Доля оцифрованных предметов в музейных коллекциях Санкт-Петербурга оказалась гораздо выше, чем в среднем по Российской Федерации и значительно выше, чем в музейных коллекциях Москвы. Совместный проект ИВМ и Эрмитажа, который мог повлиять на развитие деятельности такого рода в Санкт-Петербурге, является только час-

³⁴³ Nauta J. G., van den Heuvel W. (2015). Survey report on digitization in European Cultural Institutions 2015. Europeana Pro. <http://pro.europeana.eu/enumerate/statistics/results>

³⁴⁴ Nauta J. G., van den Heuvel W. (2015). Survey report on digitization in European Cultural Institutions 2015. Europeana Pro. <http://pro.europeana.eu/enumerate/statistics/results>

тичным объяснением. Значительный объем цифровых коллекций в этом городе и Северо-Западном федеральном округе, вероятно, является следствием взаимодействия музейного сообщества и Академии наук СССР в 1970-е годы³⁴⁵, взаимодействие музеев с крупными российскими и международными компаниями, включая IBM, и работа компании КАМИС³⁴⁶, направленной на автоматизацию учета, хранения и каталогизации музейных коллекций. Штаб-квартира компании КАМИС находится в Санкт-Петербурге³⁴⁷.

Можно отметить, что в 2015 году в Российской Федерации существовали цифровые коллекции музейных предметов, но масштаб оцифровки за пределами Северо-Западного федерального округа был значительно ниже уровня создания цифровых коллекций в Европе.

Представляется особенно важной оценка соотношения цифровых коллекций музейных предметов и количества предметов, опубликованных на сайтах музеев. Создатели цифровых коллекций руководствуются в первую очередь задачами учета, в то время как обществу в целом интересен доступ к цифровым коллекциям для просветительских целей, научных исследований, распространения изображений и создания новых визуальных практик. Результаты диссертационного исследования показывают, что в 2015 году музеи Санкт-Петербурга с рекордным уровнем цифровых коллекций в 36 % (выше европейского уровня) опубликовали на сайтах музеев изображения 0,93 % музейных предметов. Музеи Уральского федерального округа оцифровали 18 % коллекций (второе место в РФ) и опубликовали на сайтах музеев изображения 3,2 % предметов. Культурное наследие в этом регионе оказывается максимально доступным пользователям, в то время как музеи Сибирского федерального округа демонстрируют пользователям самое ма-

³⁴⁵ Шер, Я.А. Первые шаги отдела музейной информатики в Эрмитаже (1975-1985 гг) // Информационные технологии в музее. № 2. Государственный Эрмитаж, Санкт-Петербург, 2006. URL: <http://kronk.spb.ru/library/sher-yaa-2006.htm>

³⁴⁶ URL: <http://kamis.ru/kamis/>

³⁴⁷ URL: <http://kamis.ru/kontakty/>

лое количество изображений в РФ. В этом контексте действительно уместным оказывается вопрос о том, существуют ли сибирские музеи как источники данных для просветительской деятельности и научных исследований.

Результаты диссертационного исследования показывают, что в 2015 году цифровые коллекции в России существовали главным образом для учетных целей. В то же время возможность увидеть российское культурное наследие, получить доступ к изображениям для научных исследований и ввести изображения предметов из коллекций российских музеев в международный канон возможно только благодаря наличию обширных цифровых коллекций и публикации изображений онлайн. Учитывая то, что многочисленные коллекции музейных изображений доступны в сети, основная часть культурного наследия Российской Федерации рисковала остаться недоступной для общей публики и научных исследований как на российском, так и на международном уровне.

Сайты музеев Российской Федерации были проанализированы с точки зрения наличия информации о цифровых коллекциях на английском языке. При сравнении музеев Москвы, Санкт-Петербурга и соседних регионов с провинциальными регионами можно отметить, что цифровые коллекции музеев Сибири, Дальнего Востока и Северо-Кавказского федерального округа доступны международным пользователям меньше всего. В музеях Москвы, Санкт-Петербурга и Северо-Западного федерального округа интерфейс сайта музея действительно переведен на английский язык. Почти половина музеев Москвы предлагает английскую версию сайта музея, но только четверть всех музеев (16 музеев из 64) соединяют изображения с сайта музея с английским интерфейсом. Пятнадцать музеев из всех музеев РФ (0,63 % от общего количества музеев) предлагают метаданные на английском языке. В Москве метаданные на английском языке есть на сайтах Государственного музея изобразительных искусств им. А.С. Пушкина³⁴⁸, Государственной Третьяковской галереи³⁴⁹, Политехнического музея³⁵⁰, Музеев

³⁴⁸ URL: <https://pushkinmuseum.art/>

³⁴⁹ URL: <https://www.tretyakovgallery.ru/en/>

Московского Кремля³⁵¹. Похожая ситуация использования английской версии сайта для привлечения физических посетителей и очевидные сложности в предоставлении доступа к цифровым коллекциям для международных посетителей характеризует музеи Санкт-Петербурга. В то время как 25 музеев Санкт-Петербурга предлагают интерфейс музейного сайта на английском языке, только три основных музея (Государственный Эрмитаж, Музей истории Санкт-Петербурга и Русский музей) переводят метаданные на английский язык для того, чтобы у международной аудитории пользователей была возможность взаимодействия с отдельным экспонатом.

Таким образом, в 2015 году российские музеи воспринимали работу по созданию цифровых коллекций как необходимую задачу, которая поддерживает цели учета и хранения. Это подтверждается значительной разницей между объемами цифровых коллекций и количеством изображений, опубликованных на сайтах музеев, особенно в таких продвинутых регионах, как Санкт-Петербург и Северо-Западный федеральный округ.

Ограничением диссертационного исследования является то, что мы не рассматриваем, какие художественные работы были выбраны методом отбора самых предпочтительных/известных/популярных произведений для представления на цифровых платформах³⁵², что влияет на принятие решений о том, какие художественные произведения будут оцифрованы и какой эффект это оказывает на восприятие культуры. Диссертационное исследование также не рассматривает проблему качества изображений, опубликованных в сети, оставляя в стороне вопрос о том, как качество (например, высокое разрешение или эффективная цветовая калибровка) влияют на восприятие изображения. Среднее качество может

³⁵⁰ <https://polymus.ru/eng/museum/collection.html>

³⁵¹ <https://www.kreml.ru/en-Us/museums-moscow-kremlin/>

³⁵² Besser H. (1997). The changing role of photographic collections with the advent of digitization. In Katherine J. G. (ed.), *The Wired Museum*. Washington: American Association of Museums, pp. 115–127.

в свою очередь обеспечить контроль над распространением изображений и, с другой стороны, соответствовать ожиданиям пользователей, которые взаимодействуют с изображениями в текущей цифровой среде, где пользователь предполагает определенные стандарты качества.

Диссертационное исследование показывает, что цифровые коллекции действительно существовали в России в 2015 году, при этом около трети цифровых коллекций опубликованы на сайте Госкаталога в момент написания диссертационного исследования. Они представлены как метаданными, так и оцифрованными изображениями. Однако мы не можем сказать, что за пределами Госкаталога они представлены в Интернете достаточно широко, чтобы продемонстрировать репрезентативным образом все культурное наследие РФ, учитывая его географическое и этнографическое разнообразие. Мы можем подтвердить результаты нашего предыдущего исследования о процентном соотношении музейных объектов с цифровыми изображениями³⁵³. Разница между результатами этого исследования и исследования с использованием статистических данных Министерства культуры РФ³⁵⁴ двух исследований составляет 4 % (18 % и 14 %). Однако результаты нашего предыдущего исследования могли содержать погрешность выборки, так как музеи, которые приняли участие в опросе, могут проявлять больший интерес и понимание важности оцифровки и, следовательно, уделять большее внимание привлечению финансовой и административной поддержки этой деятельности.

При сравнении результатов нашего анализа с данными проекта ENUMERATE, который был направлен на «исследование масштаба оцифровки в Европе»³⁵⁵, где ряд вопросов касался процентного соотношения музейных предметов

³⁵³ Kizhner I., Terras M., Rummyantsev M. (2016a). Museum digitization practices across Russia: Survey and web site exploration results. In *Digital Humanities 2016: Conference Abstracts*. Jagiellonian University & Pedagogical University, Kraków, pp. 600–2.

³⁵⁴ Kizhner, I., Terras, M., Rummyantsev, M., Sycheva, K., Rudov, I. (2018). Accessing Russian Culture Online; the scope of digitization in museums across Russia. *Digital Scholarship in the Humanities*. 19 September 2018.

³⁵⁵ Europeana. (2017). Enumerate. URL: <https://pro.europeana.eu/tags/enumerate>

с цифровыми изображениями в музейных коллекциях³⁵⁶, можно отметить, что 14 % (средний результат по РФ в 2015 году) значительно ниже, чем результаты, полученные проектом ENUMERATE в том же году (31 % музейных коллекций). Проект ENUMERATE дает возможность сравнить данные, полученные в результате опроса сотрудников музеев, библиотек и архивов. Отчет о проекте показывает более высокое процентное соотношение для цифровых изображений в музеях по сравнению с библиотеками (19 %) и архивами (13 %)³⁵⁷. Мы не можем провести похожее сравнение, чтобы получить подобное сравнение для результатов оцифровки культурного наследия в Российской Федерации, поскольку у нас нет доступных данных по оцифровке в библиотеках и архивах. Результаты исследования процентного соотношения цифровых коллекций в Санкт-Петербурге (36 %) выше средних результатов по Европе. Процентное соотношение изображений, доступных с помощью музейных сайтов, составило в 2015 году 1,5 %. Это ниже, чем средний результат, который продемонстрирован с помощью проекта ENUMERATE (24 % цифровых коллекций и 7,5 % европейских аналоговых коллекций). Однако результаты проекта ENUMERATE включали как цифровые коллекции аналоговых предметов, так и оригинальные цифровые объекты, что несколько осложняет сравнение³⁵⁸. Большой вес цифровых коллекций в Северо-Западном федеральном округе можно частично объяснить наличием квалифицированных кадров в этом регионе, историческими связями с технологическими компаниями, инфраструктурой. Исторические причины влияния квалифицированных музейных сотрудников из Санкт-Петербурга, центра Северо-Западного федерального округа, включая связи с крупными российскими и международными компаниями, такими как IBM и КАМИС: Музейные системы, также являются важными.

³⁵⁶ Nauta J. G., van den Heuvel W. (2015). Survey report on digitization in European Cultural Institutions 2015. Europeana Pro. URL: <http://pro.europeana.eu/enumerate/statistics/results>

³⁵⁷ Nauta J. G., van den Heuvel W. (2015). Survey report on digitization in European Cultural Institutions 2015. Europeana Pro. URL: <http://pro.europeana.eu/enumerate/statistics/results>

³⁵⁸ Europeana. (2017). Enumerate. URL: <https://pro.europeana.eu/tags/enumerate>

Можно заметить, что Северо-Западный федеральный округ имеет значительный опыт формирования цифровых коллекций. Однако этот факт находится в противоречии с процентным соотношением цифровых изображений, опубликованных на сайтах музеев. Соотношение составляет 1,32 % от количества аналоговых музейных предметов для Северо-Западного федерального округа и еще более низкое соотношение (0,93 %) для Санкт-Петербурга. Эта цифра в два раза ниже, чем в среднем по России и равна процентному соотношению цифровых коллекций, опубликованных на музейных сайтах Дальнего Востока. Несмотря на то что оцифровка музейных предметов действительно ведется, эти изображения не публикуются на сайтах музеев, вопреки принципам работы с открытыми данными в области научного сотрудничества и культурного наследия, которые, как мы можем наблюдать, внедряются в Европе и США³⁵⁹. Возможное объяснение заключается в том, что центральные музеи Москвы и Санкт-Петербурга хранят миллионы музейных предметов. Еще одно объяснение отказа публиковать цифровые коллекции - желание привлечь посетителей в физический музей. Музейные сайты, которые представляют метаданные на английском языке, существуют для некоторых самых известных музеев с важными коллекциями, изображения которых давно доступны в международных печатных источниках (Государственная Третьяковская галерея, Русский музей, Музеи Московского Кремля, Государственный Эрмитаж, Государственный музей изящных искусств им. А. С. Пушкина).

Влияние на отбор тех музейных предметов, которые создают основу цифровых коллекций, в большой мере формировало содержание работ в области оцифровки в РФ, начиная с 1980-х годов. Так, в этот период администрация Государственного Эрмитажа была заинтересована в создании системы управления коллек-

³⁵⁹ Boyle, J. (2010). *The Public Domain, Enclosing the Commons of the Mind*. New Haven, CT: Yale University Press; Borgman, C. (2015). *Big Data, Little Data, No Data*. Cambridge, MA: MIT Press; Terras M. (2015). *Opening Access to Collections: the making and using of open digitized cultural content*. *Online Information Review*, 39(4): 50–69; European Commission. (2016), *The European cloud initiative*. *Digital single market*. <https://ec.europa.eu/digital-single-market/en/%20european-cloud-initiative>

циями для собрания европейской живописи³⁶⁰. Намерение администрации Эрмитажа перенести знаменитые работы из печатных материалов в цифровые коллекции можно легко объяснить и понять с точки зрения позиционирования Эрмитажа как учреждения, которое поддерживает ключевые европейские ценности и хранит богатое культурное наследие Европы. Еще одним возможным объяснением того, что цифровые коллекции, опубликованные в сети, часто не выходят за рамки печатного канона, может быть желание музейной администрации сохранить контроль над распространением содержания коллекций. Понятие контроля обсуждалось в контексте лицензирования изображений из коллекций музеев США в начале 21 века³⁶¹. Это желание, вероятно, характерно для администраций музеев во всем мире, и российские музеи не являются исключением. Борьба с ситуацией, когда разрешения на распространение изображений являются обязательными в учреждениях, которые хранят произведения визуального искусства³⁶², несмотря на то, что многие из них перешли в общественное достояние и могут быть опубликованы без ограничений³⁶³, что происходит в некоторых музеях в некоторых странах мира³⁶⁴, осложняется значительным сопротивлением музейной админист-

³⁶⁰ Шер, Я.А. Первые шаги отдела музейной информатики в Эрмитаже (1975-1985 гг) // Информационные технологии в музее. № 2. Государственный Эрмитаж, Санкт-Петербург, 2006. URL: <http://kronk.spb.ru/library/sher-yaa-2006.htm>

³⁶¹ Kelly, K. (2013). Images of Works of Art in Museum Collections: The Experience of Open Access. Prepared for the Andrew W. Mellon Foundation. URL: http://archiv.ub.uni-heidelberg.de/artdok/4998/1/Kelly_Images_of_works_of_art_in_museum_collections_2013.pdf

³⁶² Bielstein, S. (2006). Permissions, A Survival Guide: Blunt Talk about Art as Intellectual Property. Chicago, IL: The University of Chicago Press.

³⁶³ Petri, G. (2014). The public domain vs the museum: the limits of copyright and reproductions of two-dimensional works. *Journal of Conservation and Museum Studies*, 12(1). URL: <http://www.jcms-journal.com/articles/10.5334/jcms.1021217/print/>

³⁶⁴ Aufderheide, P., Milosevic, T., Bello, B. (2016), The impact of copyright permissions culture on the US visual arts community: the consequences of fear of fair use. *New Media and Society*, 8(9): 2012–27.

рации, когда музеи предполагают, что «запрос на предоставление разрешения является обязательным»³⁶⁵. Эта точка зрения поддерживается в российских музеях законодательством РФ³⁶⁶.

Результаты диссертационного исследования ставят под сомнение восприятие музейных коллекций во всем мире как «видимых и легко доступных»³⁶⁷. В то время как уровень доступа к музейным коллекциям в Европе, ограниченный техническими сложностями и проблемами, связанными с авторским правом³⁶⁸, остается достаточно высоким во многих странах Европы, это не значит, что этот уровень достигнут во всем мире. Хотя многое известно о той группе музеев, которая публикует значительную долю своих коллекций в сети³⁶⁹ и о европейских музеях, в которых есть цифровые коллекции³⁷⁰, необходимы дальнейшие исследования, которые помогут определить долю тех музеев, которые способны распространять информацию о своих коллекциях через цифровые платформы в международном масштабе.

³⁶⁵ Aufderheide P., Milosevic T., Bello B. (2016), The impact of copyright permissions culture on the US visual arts community: the consequences of fear of fair use. *New Media and Society* , 8(9): 2012–27.

³⁶⁶ Kizhner I., Stankevich J., Rumyantsev M., Makarchuk I. (2016). Licensing images from Russian museums for an academic project. *Journal of Siberian Federal University. Humanities and Social Sciences* , 7(9), 1618–30. URL: http://elib.sfu-kras.ru/bitstream/handle/2311/20416/09_Kizhner.pdf?sequence=1

³⁶⁷ Salamon-Cindori, B., Tot, M., Zivkovic, D. (2014). Digitization: challenges for Croatian museums. *Qualitative and Quantitative Methods in Libraries* , 1: 213–220.

³⁶⁸ Taylor, J., Gibson, L. K. (2016). Digitization, digital interaction and social media: embedded barriers to democratic heritage. *International Journal of Heritage Studies* , 23(5), 1–13.

³⁶⁹ Aufderheide, P., Milosevic, T., Bello, B. (2016), The impact of copyright permissions culture on the US visual arts community: the consequences of fear of fair use. *New Media and Society* , 8(9): 2012–27.

³⁷⁰ Nauta, J. G., van den Heuvel, W. (2015). Survey report on digitization in European Cultural Institutions 2015. *Europeana Pro*. URL: <http://pro.europeana.eu/enumerate/statistics/results>

Если мировые коллекции визуального искусства будут продолжать оставаться невидимыми и недоступными, создание корпуса визуальных произведений для исследований в области истории искусства³⁷¹ и применение науки о больших данных к исследованиям в этой области³⁷² будет ограничено произведениями искусства из западных музеев. Дальнейшие шаги в области разработки моделей визуальных данных, изменение наборов параметров и извлечение новых неожиданных параметров в наборах больших визуальных данных будут ограничены наборами доступных данных³⁷³, и анализ очевидным образом будет содержать погрешности, вызванные преобладанием данных из западных музеев.

Масштаб усилий, направленных на создание цифровых коллекций в Европе и других частях мира, требует объединения проектов в области оцифровки и научных исследований, направленных на исследование эволюции культурного канона в цифровую эпоху. Тенденциям культурологического анализа двадцатого века, которые связаны с колониальными и политическими установками девятнадцатого века³⁷⁴, придут на смену новые отношения, вызванные двадцать первым веком. Культура ремикса³⁷⁵, которая предполагает повторное использование художественных работ в своих произведениях, требует разработки теорий и тщательного анализа практик для создания нового восприятия культуры. Необходимы дальнейшие исследования для того, чтобы понять способы эволюции культурного канона, и определения того, каким образом он отличается от материала печатных публикаций. Определение критериев, по которым художественное произведение

³⁷¹ Drucker, J. (2013), Is there a 'digital' art history? *Visual Resources* , 29(1–2): 5–13.

³⁷² Manovich, L. (2015). Data science and digital art history. *International Journal for Digital Art History* , 1, pp. 12–35.

³⁷³ Manovich, L. (2016), Cultural analytics, social computing and digital humanities. In Schafer M. T., van Es K. (eds), *The Datafied Society: Studying Culture Through Data* . Amsterdam: Amsterdam University Press, pp. 55–69.

³⁷⁴ Said E. W. (1993). *Culture and Imperialism* . New York: Random House.

³⁷⁵ Lessig L. (2008). *Remix: Making Art and Commerce Thrive in the Hybrid Economy* . New York: Penguin Press.

попадает в цифровые коллекции, публикуется на сайте культурного учреждения, легко обнаруживается с помощью системы навигации и соединяется с релевантной информацией из других источников - важный научный вопрос, который должен быть поставлен в будущих научных исследованиях. Вероятно, было бы полезно провести исследования, подобные данной диссертационной работе, для сравнения результатов исследования ENUMERATE с оценкой деятельности в области музейной оцифровки в других географических регионах для того, чтобы оценить, действительно ли европейские и североамериканские цифровые коллекции доминируют как коллекции культурного наследия, представленные на цифровых платформах.

Данное диссертационное исследование представляет первый взгляд на состояние российских цифровых коллекций в региональном и национальном масштабах, в исследовании сообщается об уровне цифровых коллекций в разных регионах России. Таким образом мы можем сделать вклад в исследование понятия цифрового канона и утверждать, что в цифровом пространстве культурный канон, представленный в художественных произведениях, следует расширить. Наше исследование поддерживает недавнюю критическую оценку деятельности по созданию цифровых коллекций, которая не сопровождается определением контекста предметов культурного наследия и недостаточна для того, чтобы создать новые знания в гуманитарных науках³⁷⁶. В контексте российского культурного наследия отставание в публикации изображений на сайтах музеев может быть использовано для создания сети текстов и изображений, предлагающих необходимый исторический контекст, который объединяет шедевры и другие визуальные произведения из разных географических и социальных сфер и связывает их друг с дру-

³⁷⁶ Hitchcock T. (2013). Confronting the digital or how academic history writing lost the plot. *Cultural and Social History*, 10: 9–23; Gregory I., Atkinson P., Hardie A., Joulain-Jay A., Kershaw D., Porter C., Rayson P., Rupp C. J. (2016). From digital resources to historical scholarship with the British library 19th century newspaper collection. *Journal of Siberian Federal University. Humanities and Social Sciences*, 9(4): 994–1006.

гом. Необходимы государственные программы, которые могут предложить рекомендации относительно того, как сайты музеев РФ и/или Государственный каталог Музейного фонда РФ должны организовать систему навигации для поиска и просмотра визуальных произведений для создания инфраструктуры, которая будет поддерживать исследования в области гуманитарных исследований. Будущие исследования необходимы также для понимания масштаба оцифровки и известности цифровых коллекций культурного наследия, которые содержатся в библиотеках и архивах РФ, для понимания того, как широкая международная аудитория может получить доступ к национальному культурному наследию. Только увеличивая масштаб деятельности по созданию цифровых коллекций и расширяя возможности ознакомления с цифровыми коллекциями культурного наследия в России, можно добиться того, что они станут известны российской и международной аудитории.

Исследование коллекций цифрового агрегатора Google Arts and Culture

За последние два десятилетия визуальные данные в области культурного наследия, хранящиеся в музеях, архивах и библиотеках, стали доступны для развлечения, образования и исследований, по крайней мере, в некоторых странах. Можно отметить значительное увеличение количества предметов, хранящихся в музеях и художественных галереях, которые можно увидеть в Интернете³⁷⁷. Можно также отметить рост количества учреждений, которые ввели открытое лицензирование для этих данных³⁷⁸. При этом проблемы устройства лингвистиче-

³⁷⁷ Rogers N (2016). Museum drawers go digital. *Science* 352:762–765; Wang, M.C. (2016). Digitization and the Depoliticization of Museum Access. *The Situation at the National Palace Museum. Advances in Computer Science: an International Journal*, Vol. 5, Issue 3, No.21, May 2016; Hall, S. (2017). Memories of style: Digitizing the Yorkshire Fashion Archive. *Art Libraries Journal*, 42(3), 157-161

³⁷⁸ Kapsalis, Effie. (2016). “The Impact of Open Access on Galleries, Libraries, Museums and Archives”. Smithsonian Emerging Leaders Development Program. http://siarchives.si.edu/sites/default/files/pdfs/2016_03_10_OpenCollections_Public.pdf; Tallon, L. (2019). Sparking Global Connections to Art Through Open Data and Artificial Intelligence.

ских корпусов³⁷⁹ и устранения ошибок при обработке больших корпусных данных в компьютерной лингвистике³⁸⁰ или цифровых литературных исследованиях³⁸¹ хорошо известны. Однако методы организации визуальных коллекций в культурном наследии для применения количественного анализа и методов машинного обучения изучены в гораздо меньшей степени.

Один из способов понять масштаб и характер этой проблемы – воспользоваться примером крупного агрегатора данных, платформы Google Arts and Culture (GA&C). Вопрос исследования заключается в том, насколько репрезентативной

Blogs/Now at the Met. The Metropolitan Museum of Art, 4 February 2019. <https://www.metmuseum.org/blogs/now-at-the-met/2019/met-microsoft-mit-art-open-data-artificial-intelligence>; McCarthy, D. (2019). Open Access Scope in Open GLAM, Medium, 2 May 2019. <https://medium.com/open-glam/open-access-scope-in-open-glam-70461bec2bca>

³⁷⁹ Atkins, S., Clear, J., & Ostler, N. (1992). Corpus Design Criteria. *Journal of Literary and Linguistic Computing*, 7(1), 1-16; Biber, D. (1993). Representativeness in Corpus Design. *Literary and Linguistic Computing*, 8, Issue 4, Oxford University Press, pp. 243-257; Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. New York: Cambridge University Press; Maekawa, Kikuo, Hanae Koiso, Sadaoki Furui and Hitoshi Isahara (2000). Spontaneous speech corpus of Japanese. *Proceedings of the Second International Conference of Language Resources and Evaluation (LREC)*, 2, 947-952.

³⁸⁰ Clear, J. (1992). Corpus sampling. In G. Leitner (ed.) *New directions in English language corpora*, Mouton-de-Gruyter, Berlin, pp. 21 - 31; Atkins, S., Clear, J., & Ostler, N. (1992). Corpus Design Criteria. *Journal of Literary and Linguistic Computing*, 7(1), 1-16; Leech, Geoffrey, 2007. New resources, or just better old ones? The Holy Grail of representativeness. In: Hundt, Marianne, Nesselhauf, Nadja, Biewer, Carolin (Eds.), *Corpus Linguistics and the Web*. Rodopi, Amsterdam, pp. 133-150; Gries, S.T. 2009. What is corpus linguistics? *Language and Linguistics Compass* 3(5): 1–17.

³⁸¹ Jockers, M.L. (2013). *Macroanalysis*. Champaign, IL: University of Illinois Press; Bode, K. and Tara Murphy. 'Methods and Canons: An Interdisciplinary Excursion.' *Advancing Digital Humanities: Research, Methods, Theories*. Eds. Paul Longley Arthur and Katherine Bode. London: Palgrave Macmillan, 2014. 175-93; Корчагин К. М. «Поэзия XX века в поэтическом подкорпусе Национального корпуса русского языка: Проблема репрезентативности». *Труды Института русского языка им. В. В. Виноградова* 6 (20156): 235–256; Bode, K. (2020). 'Why You Can't Model Away Bias,' *Modern Language Quarterly* 81.1.

и сбалансированной является платформа Google Arts and Culture (GA&C) как один из самых крупных агрегаторов оцифрованных данных о культурном наследии всего мира³⁸². Исследование проводилось осенью 2018 года и весной 2019 года в кафедре информационных технологий в креативных и культурных индустриях СФУ³⁸³. Результаты исследования представлены в статье И. Кижнер, М. Террас, М. Румянцева и др.³⁸⁴ С 2011 года Google Arts and Culture предоставляет широкой публике изображения произведений искусства из крупных галерей и художественных музеев с высоким разрешением: в настоящий момент на платформе размещено около шести миллионов изображений произведений искусства из 1 780 музеев по всему миру. Задача платформы Google Arts and Culture – предоставить возможность «по-новому увидеть для себя произведения искусства, международные коллекции и истории, ...увидеть произведения искусства в мельчайших деталях и дать возможность распространения этих произведений»³⁸⁵, цель GA&C – демократизация и распространение международного цифрового наследия, а его статус делает его идеальным местом для изучения того, какая культура поддерживается, распространяется и доступна для онлайн-аудитории. Этот значительный проект, финансируемый большим технологическим гигантом со штаб-квартирой в США, может в результате продемонстрировать тенденциозность представления данных, что повлечет за собой непреднамеренные последствия. Выбор GA&C для исследования определяется тем фактом, что это международная платформа, созданная влиятельным коммерческим гигантом, который оказал огромное влияние на поиск, сортировку, классификацию и фильтрацию информа-

³⁸² URL: <https://artsandculture.google.com/?hl=en>

³⁸³ Kizhner, I., Terras, M., Rumyantsev, M., Khokhlova, V., Demeshkova, E., Rudov, I., Afanaseva, J. Measuring Bias in Aggregated Digitised Content: a Case Study on Google Arts and Culture, *Digital Scholarship in the Humanities* (принято в печать).

³⁸⁴ Kizhner, I. et al. Digital Cultural Colonialism: measuring bias in aggregated digitized content held in Google Arts and Culture. *Digital Scholarship in the Humanities* (в печати).

³⁸⁵ URL: <https://www.google.com/culturalinstitute/about/>

ции по всему миру. Платформа включает большую коллекцию, которая на момент написания диссертационного исследования насчитывает около шести миллионов изображений и охватывает почти все 200 стран из списка стран-членов ООН³⁸⁶.

Цифровые изображения произведений искусства и инфраструктура, поддерживающая их хранение и агрегирование, важны по ряду социальных, политических и этических причин. Во-первых, они предоставляют доступ к большому количеству значимых произведений искусства, в том числе тех, которые не выставлены в физических музеях (например, рисунки и гравюры), и позволяют нам подробно изучить каждый музейный объект. Это имеет важные последствия для тех пользователей, которые живут в отдаленных местах, где посещение известных физических музеев связано с рядом ограничений. Однако на структуру агрегатора влияют частные интересы, упомянутые в научной литературе³⁸⁷. К тому же, онлайн-агрегаторы используют данные, полученные от музеев, архивов и библиотек, где коллекции были составлены под влиянием определенных принципов отбора материалов.

Задача данного раздела параграфа - понять, как крупные коллекции разных стран представлены в Google Arts and Culture. В частности, рассматривается количество учреждений, представляющих каждую страну в Google Arts and Culture, количество изображений, представленных каждым учреждением, и количество «иностранных» музейных предметов в музеях других стран, когда географическое название в рамках другой страны указано в метаданных. Диссертационное исследование показывает распределение изображений по времени, жанрам искусства и доминирующим художникам для двух стран, России и Франции. Эти данные

³⁸⁶ URL: <https://www.un.org/en/member-states/index.html>

³⁸⁷ Mansfield, E. C. (2014) Google Art Project and Digital Scholarship in the Visual Arts. URL: <http://www.google.com/culturalinstitute/project/art-project>, Visual Resources: An International Journal of Documentation, 30:1, 110-117; Van Dijck J, Poell T and Waal M de (2018). *The Platform Society*. New York: Oxford University Press.

сравниваются с данными, опубликованными национальными агрегаторами цифровых изображений предметов культурного наследия этих двух стран.

В параграфе вводится понятие репрезентативных и сбалансированных цифровых коллекций, заимствованное из корпусной лингвистики и применяемое к агрегированному визуальному контенту. Показано, как важно продемонстрировать ограничения электронных коллекций данных. При этом в научной литературе корпус ранее определялся как «подверженный ошибкам и тенденциозному представлению данных (skewed)»³⁸⁸ или «неизбежно нерепрезентативный (inevitably biased)»³⁸⁹. Литература показывает, что предвзятость корпуса усугубляется особенностями отбора произведений, которые свойственны тем галереям, музеям, архивам и библиотекам, которые участвуют в составлении корпуса³⁹⁰ или особенностями, вызванными социальными обстоятельствами³⁹¹.

Затем параграф рассматривает, как Google Arts and Culture, один из самых крупных агрегаторов визуальных данных в области культурного наследия, организует свою цифровую коллекцию. Параграф демонстрирует тенденциозность представления отдельных культур и недостаточную репрезентативность (сбалансированность) коллекций ряда стран и культурных учреждений. Показано, что эта особенность распространяется на музейные предметы других стран в национальных коллекциях, например, предметы египетского искусства в европейских музе-

³⁸⁸ Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton;

³⁸⁹ Clear, J. (1992). *Corpus sampling*. In G. Leitner (ed.) *New directions in English language corpora*, Mouton-de-Gruyter, Berlin, pp. 21 - 31, см. также Raineri, S. and Debras, C. 'Corpora and Representativeness: Where to go from now?' *CogniTextes*. Vol. 2019. <https://journals.openedition.org/cognitextes/1671>; Bode, K. (2020). 'Why You Can't Model Away Bias,' *Modern Language Quarterly* 81.1.

³⁹⁰ Huster, A. C. (2013). *Assessing Systematic Bias in Museum Collections: A Case Study of Spindle Whorls*. *Advances In Archaeological Practice*. 1(2):77-90.

³⁹¹ Корчагин К. М. «Поэзия XX века в поэтическом подкорпусе Национального корпуса русского языка: Проблема репрезентативности». *Труды Института русского языка им. В. В. Виноградова* 6 (2015б): 235–256.

ях. Мы видим большие расхождения в количестве объектов из пяти ведущих стран, где количество представленных предметов является максимальным (США, Великобритания, Нидерланды, Италия и Южная Корея) и остального мира. Более того, страны, представленные GA&C предметами, хранящимися в учреждениях других стран, не демонстрируют сбалансированного участия. Первые девять стран в этом списке хранят около 40% коллекций, представленных в GA&C.

Затем параграф рассматривает географическое распределение, виды искусства и периоды времени для предметов, опубликованных в двух крупных коллекциях GA&C: коллекциях России и коллекциях Франции. Эти данные сравниваются с данными, представленными национальными музейными агрегаторами двух стран. Можно заметить, что визуальный контент представлен в значительно большей степени культурными учреждениями столиц обеих стран, при этом искусство провинциальных регионов представлено в меньшей степени. Искусство, представленное GA&C, в большей мере касается современного искусства и фотографических изображений XIX-XX веков. При этом особенно представлены необычные явления и изменения в визуальном каноне, демонстрируются характерные особенности экзотических культур. Этот результат получен с помощью сравнения видов искусства и периодов времени, представленных в GA&C с национальными агрегаторами, а также с помощью изучения литературы, которая касается пяти авторов, произведения которых встречаются чаще всего для коллекции каждой страны (Россия и Франция).

Эти результаты обсуждаются в ключе подтверждения цифрового колониализма, свойственного агрегированному цифровому контенту в области культурного наследия, при этом в диссертационном исследовании предполагается, что колониальный подход был унаследован от печатной эпохи, предшествующей эпохе цифрового контента. Позиция цифрового культурного колониализма противоречит взглядам, открыто выраженным руководителем проекта GA&C Амитом

Судом³⁹². Иными причинами отсутствия равноправия в представленности данных в GA&C может быть прозрачность данных, доступность музейных API-интерфейсов и политика открытого доступа для некоторых стран и участвующих учреждений, которые облегчают работу цифровых платформ и значительно усиливают возможность получения контента. Эти эффекты могут быть усилены алгоритмами поисковых систем. Диссертационное исследование рекомендует сопровождать агрегированные данные описанием принципов отбора и статистическими данными, демонстрирующими распределение агрегированного контента, чтобы не создавать у пользователя неправильное представление о возможностях платформ.

Основным источником данных для анализа новых визуальных практик в этой части диссертационного исследования и в работе И. Кижнер и др.³⁹³ были данные самой платформы GA&C. Платформа публикует изображения предметов культурного наследия, видео и панорамы с высоким разрешением. При этом поля базы данных GA&C и последующий интерфейс пользователя включают географические регионы, к которым относятся произведения искусства и культуры, такие как Соединенные Штаты Америки, Китай, Франция, Российская Федерация, Париж или Лондон. Google Arts and Culture – издательская платформа для агрегированного оцифрованного контента, созданная в 2011 году Институтом культуры компании Гугл³⁹⁴. Можно отметить устойчивый рост количества оцифрованного контента в период между 2009 и 2015 годами и резкое увеличение числа учреждений, начиная с 2015 года. Этот рост можно объяснить широкой популярностью

³⁹² Sood, A. (2016). Every piece of art you've ever wanted to see - up, close and searchable. TED lecture. URL: https://www.ted.com/talks/amit_sood_every_piece_of_art_you_ve_ever_wanted_to_see_up_close_and_searchable

³⁹³ Kizhner, I. et al. Digital Cultural Colonialism: measuring bias in aggregated digitized content held in Google Arts and Culture. *Digital Scholarship in the Humanities* (в печати).

³⁹⁴ <https://www.google.com/culturalinstitute/about/>

платформ, агрегирующих цифровой контент, таких, например, как Курсера³⁹⁵ примерно в это время или тем, что некоторые учреждения начали распространять свои коллекции, следуя политике открытого доступа. Результаты этого параграфа основаны на анализе полной коллекции двумерных изображений с метаданными, доступной на веб-сайте GA&C (около 5 500 000 изображений предметов культурного наследия) осенью 2018 года и весной 2019 года. Все приведенные ниже количественные оценки основаны на анализе цифровых коллекций учреждений культуры, таких, например, как журнал LIFE или Opera National de Paris, опубликованных GA&C) для каждой страны из списка государств-членов ООН³⁹⁶. Для анализа данных о музейных объектах, представляющих ряд культур, в учреждениях других стран мы использовали раздел «Откройте для себя это место», в котором перечислены все предметы с упоминанием страны (или города в этой стране) в метаданных или свободных текстах, сопровождающих изображения. Категория «Места» в составе меню цифровой платформы GA&C дает возможность выбрать конкретное место (например, Францию) и рассмотреть, как произведения искусства этой страны представлены цифровой платформой. Анализ категории «Места» позволил нам увидеть, как коллекции отдельных стран представлены с помощью GA&C.

Было определено количество и размер коллекций, представленных в GA&C, и страна, в которой хранятся коллекции. Анализируя тот раздел сайта, который представляет коллекции страны, мы можем определить количество коллекций (например, 87 коллекций для Франции по состоянию на осень 2018 года). В каждом разделе, посвященном определенной коллекции, отображается количество элементов в коллекции. Суммируя количество коллекций в стране и количество произведений из каждой коллекции, мы получаем количественные оценки для анализа и сравнения. По состоянию на осень 2018 года платформа GA&C

³⁹⁵ URL: <https://www.coursera.org/>

³⁹⁶ URL: <https://www.un.org/en/member-states/index.html>

включала около 5 500 000 изображений произведений искусства и культуры из 1 780 учреждений культуры, представляющих многие страны мира.

Количество коллекций и произведений искусства и культуры, представленных цифровой платформой, само по себе не имеет особого значения, если мы не знаем количество учреждений культуры, включающих цифровые изображения предметов в свои базы данных, количество оцифрованных произведений, которые хранятся в базах данных учреждений культуры или количество изображений произведений культуры в стране. Однако можно проверить количество оцифрованных музейных предметов, изображения которых хранятся в базах данных учреждений культуры или национальных агрегаторах для некоторых стран, поскольку изображения и метаданные включены в государственные статистические базы данных. Для дальнейшего анализа мы выбрали коллекции России и Франции, представленные на цифровой платформе Google Arts and Culture. Коллекции этих стран были выбраны потому, что в обеих странах есть впечатляющее количество музеев и музейных предметов – около 2 300 музеев и 80 миллионов музейных предметов для России³⁹⁷ и 1 224 учреждения, 121 миллион музейных предметов для Франции³⁹⁸. Музеи обеих стран представлены в исследовании, рассматривающем восемнадцать самых влиятельных музеев мира³⁹⁹. Кроме того, статистические данные о фондах и оцифрованных коллекциях этих музеев доступны из правительственных источников, таких как Государственный каталог музейных коллекций Российской Федерации⁴⁰⁰ или Открытая платформа наследия Мини-

³⁹⁷ Kizhner, I., Terras, M., Rummyantsev, M., Sycheva, K., Rudov, I. Accessing Russian Culture Online: The scope of digitization in museums across Russia. *Digital Scholarship in the Humanities*, Volume 34, Issue 2, June 2019, Pages 350–367.

³⁹⁸ URL: <http://www2.culture.gouv.fr/documentation/joconde/fr/apropos/presentation-joconde.htm>

³⁹⁹ Van Riel C., Heijndijk P. (2017). Why people love art museums: a reputation study about the 18 most famous museums among visitors in 10 countries. Rotterdam School of Management, Erasmus University.

⁴⁰⁰ URL: <https://goskatalog.ru/portal/#/>

стерства культуры Франции⁴⁰¹. Эти данные позволяют сравнить официальное количество оцифрованных материалов с тем, как эти коллекции представлены платформой GA&C для того, чтобы продемонстрировать, что количество материалов, представленное платформой, значительно уступает количеству, которое представлено в опубликованных официальных источниках.

Основная часть изображений в Google Arts and Culture представляет произведения искусства и культуры, которые хранятся в учреждениях, расположенных в США. Эта часть изображений значительно превосходит все остальные национальные коллекции. Тем не менее, Google Arts and Culture демонстрирует французские и русские коллекции в достаточной степени, чтобы получить представление об их богатстве и разнообразии. Эти страны занимают промежуточное положение в списке стран с наибольшим количеством коллекций и произведений в GA&C. Министерства культуры обеих стран поддерживают базы данных оцифрованных произведений искусства, которые хранятся в федеральных, региональных и муниципальных музеях. Базы данных опубликованы онлайн, а поисковая система позволяет установить количество музейных предметов разных жанров и периодов времени, представленных в Государственном каталоге МФ РФ или французской Открытой платформе наследия. Это позволяет нам узнать, насколько репрезентативен GA&C с точки зрения географического распределения или распределения по времени и жанрам. В диссертационном исследовании используется качественная методология для изучения принципов подбора контента GA&C для обеих стран. Были проанализированы тексты, опубликованные на платформе GA&C и написанные научными сотрудниками французских и российских музеев. Кроме этого, мы использовали стандартные процедуры поиска информации⁴⁰², чтобы найти доказательства того, что определяет подбор контента

⁴⁰¹ URL: <https://www.pop.culture.gouv.fr/>

⁴⁰² Ellis, D. 1989. 'A Behavioural Model for Information Retrieval System Design', *Journal of Information Science*, 15(4–5), 237–247.

для учреждений, представивших свои коллекции для публикации с помощью платформы GA&C.

Было определено количество коллекций и произведений искусства всех российских учреждений культуры, опубликованных GA&C. Следующим этапом исследования был анализ того, какое количество российских музеев оцифровало (полностью или частично) свои коллекции. Этот этап был необходим для того, чтобы сравнить количество изображений, представленных в GA&C, с количеством оцифрованных предметов, представленных в базах данных музеев, предоставляющих статистическую отчетность в Министерство культуры РФ. GA&C включает произведения искусства, которые хранятся в государственных музеях. Туда также входят коллекции частных музеев, коллекции отдельных художников и коллекции других культурных учреждений, таких как театры, научно-исследовательские институты, архивы фотожурналов и т. д. Число этих учреждений может быть равно, а может даже превосходить число государственных музеев, которые предоставляют статистическую отчетность и количество оцифрованных экспонатов которых известно. В июне 2019 года количество произведений искусства в российских коллекциях GA&C из фондов государственных музеев составляло 2 844 изображения, а количество произведений искусства из других учреждений культуры – 2 784 изображения. Это диссертационное исследование включает анализ произведений искусства и культуры, которые хранятся в государственных музеях. Таким образом, мы можем сравнить географическое распределение российских коллекций в GA&C с базой данных оцифрованных коллекций государственных музеев, предоставленной Государственным каталогом Музейного фонда РФ⁴⁰³. Уровень оцифровки музейных коллекций в России в июне

⁴⁰³ Государственный каталог Музейного фонда РФ – это обширная база данных всех музейных предметов в стране, собранная с целью инвентаризации музейных коллекций. К 2026 году в базу данных должны войти 100 % музейных предметов Основного фонда всех российских музеев, подотчетных Министерству культуры (около 60 миллионов предметов). Доступ к базе данных осуществляется через поисковую систему с двумя видами поиска: по типам коллекций и по на-

2019 года известен, поскольку в Государственном каталоге Музейного фонда РФ, опубликованном Министерством культуры РФ, приведены данные о количестве доступных объектов с изображениями.

В диссертационном исследовании проведена работа с данными о российских произведениях искусства из государственных музеев РФ, представленными на цифровой платформе GA&C. Было определено географическое положение каждого музея, местоположение музеев было представлено на карте, чтобы сравнить данные о географическом распределении российских музеев из GA&C с географическим распределением музеев, представленных в Государственном каталоге МФ РФ. Список музеев и их географическое распределение были обнаружены на сайте Государственного каталога МФ РФ. Была составлена таблица, которая включала отметки о каждом предмете, оцифрованное изображение которого представлено в российских коллекциях GA&C. Таблица включала тип объекта, период создания и автора произведения. Схема классификации для типов объектов, таких как картины, скульптуры, документы или нумизматические коллекции, присутствует в Государственном каталоге и является одним из двух способов, с помощью которых можно производить поиск в базе данных. Это позволило сравнить количественные показатели для типов предметов культуры и искусства в нашей электронной таблице и количественные показатели в классификационных разделах Государственного каталога, чтобы выяснить, насколько репрезентативной является выборка платформы Google Arts and Culture. Были определены типы предметов для произведений искусства, опубликованных GA&C, при этом использовались метаданные и произвольный текст, сопровождающий изображения. В этой работе иногда приходилось полагаться на собственное суждение, поскольку GA&C не использует стандартное описание произведений искусства

званию музея. В феврале 2020 года база данных насчитывала свыше 17 миллионов предметов или более 25 % российских Основного фонда музеев РФ.

(стандарт метаданных)⁴⁰⁴ и данные о классификационном типе предмета были не всегда доступны. На следующем этапе исследования необходимо было определить авторов, фамилии которых чаще всего встречались в российских коллекциях GA&C. Использование электронной таблицы позволило определить 10 наиболее часто встречающихся фамилий. Невозможно было сравнить эти данные с данными Государственного каталога, поскольку эта издательская платформа не предоставляет список наиболее частых имен создателей или авторов. Однако эти данные демонстрировали тип и период в развитии искусства, который, как предполагали редакторы платформы и научные сотрудники тех учреждений культуры, которые работали с GA&C, вызовет наибольший интерес у пользователя. В целом, на первом этапе были изучены 49 российских коллекций и 5 628 изображений. После того как были исключены частные музеи и те учреждения культуры, которые не давали возможность сравнить представленную выборку с официальными статистическими данными, был проведен анализ 32 российских коллекций, в таблицу были внесены данные для 2 844 предметов искусства и культуры, представленных на платформе GA&C.

Анализ количественных показателей, связанных с коллекциями и произведениями искусства тех французских учреждений культуры, которые представили данные для публикации в GA&C, проводился осенью 2018 года. Было определено географическое распределение музеев, которые отправили цифровые изображения произведений искусства из своих коллекций для публикации в GA&C. Были использованы электронные таблицы Excel, которые включали такие данные, как имя автора, период времени, тип коллекции и доступность метаданных. Чтобы сравнить географическое положение музеев, представивших свои коллекции в GA&C, с официальными данными о музеях Франции, мы работали со списком французских музеев, которые внесены в реестр Министерства культуры Фран-

⁴⁰⁴ Mansfield, E. C. (2014) Google Art Project and Digital Scholarship in the Visual Arts. URL: <http://www.google.com/culturalinstitute/project/art-project>, *Visual Resources: An International Journal of Documentation*, 30:1, 110–117.

ции⁴⁰⁵. Сравнение данных по авторам и периодам времени стало возможным благодаря тому, что Министерство культуры Франции опубликовало базу данных «Джоконда», которая включает изображения около 400 000 оцифрованных музейных предметов из французских музеев, подотчетных французскому Министерству культуры⁴⁰⁶. Классификационные типы предметов, представленные в базе данных Joconde, были слишком многочисленными для дальнейшей работы и анализа, поэтому этот тип сравнительного анализа был исключен из исследования. В целом, на втором этапе анализа весной 2019 года мы работали с 94 коллекциями и 19 024 музейными предметами, цифровые копии которых были опубликованы на платформе GA&C. После сравнения списка французских культурных учреждений в GA&C и списка французских национальных музеев были исключены те музеи и коллекции не музейных учреждений, которых не было в списке. Таким образом, дальнейший сравнительный анализ позволил определить, насколько репрезентативной оказалась выборка GA&C с точки зрения географического распределения национальных музеев, а также распределения по периодам времени и авторам). В результате были проанализированы 6 575 музейных предметов из 21 коллекции, которые вошли в исследование для определения репрезентативности и сбалансированности национальных коллекций агрегатора визуального контента, представляющего культурное наследие разных стран.

Результаты анализа показывают значительный разрыв между коллекциями пяти стран, представляющих максимальное количество изображений предметов культуры и искусства коллекции в GA&C. Можно отметить, что предметы из США, Великобритании, Нидерландов, Италии и Южной Кореи составляют 93,4 % контента, публикуемого платформой GA&C, причем доля предметов, которые находятся в учреждениях культуры США, составляет 82 % изображений, опубликованных платформой. Доля остальных 190 стран из списка ООН составляет 6,6 %. Это означает, что вероятность получения изображений предметов, ко-

⁴⁰⁵ URL: <http://www2.culture.gouv.fr/documentation/museo/index.htm>

⁴⁰⁶ URL: <https://www.pop.culture.gouv.fr>

торые хранятся в учреждениях культуры стран с максимальным количеством изображений для нейтральных поисковых терминов, таких как «шахматы», «перчатки» или «украшения», намного выше, чем для изображений, поступающих из стран остального мира, при этом культуры стран из списка с максимальным количеством изображений (особенно США) получают возможность привлечь большее внимание пользователей и повлиять на развитие их представлений об историческом развитии культуры и искусства.

Еще один подход к изучению репрезентативности и сбалансированности цифрового агрегатора – рассмотреть распределение коллекций на уровне учреждений культуры, которые публикуют изображения хранящихся в них предметов с помощью платформы GA&C. Можно отметить, что количество изображений коллекции фотографического журнала LIFE⁴⁰⁷ в пятнадцать раз превышает количество изображений следующей коллекции в списке. Контент фотожурнала составляет 76 % предметов культуры, опубликованных с помощью платформы GA&C. Это явление демонстрирует, как включение одной большой коллекции может исказить репрезентативность всего агрегатора. Количество изображений, представленных самой крупной коллекцией агрегатора, означает, что доля изображений из этой коллекции по каждому запросу, связанному с двадцатым веком, таким как «Маргарет Тэтчер», «фашизм» или «колхоз», будет выше, чем доля изображений, представленных другими учреждениями культуры и/или другими странами. Таким образом, представление о двадцатом веке создается через призму редакторов конкретного фотожурнала, представляющих культуру конкретной страны.

Хотя результаты этой части диссертационного исследования показывают огромные различия в количестве объектов для разных стран на уровне географиче-

⁴⁰⁷ Журнал LIFE выходил еженедельно с 1936 по 1972 год. Он является богатым источником визуальных данных в области социальной истории США и многих стран мира, подробнее см. Grady, J. (2007). Advertising images as social indicators: depictions of blacks in LIFE magazine, 1936–2000. *Visual Studies*. 22(3): 211-239.

ческого распределения учреждений культуры, мы не можем сказать, что GA&C не является репрезентативным корпусом, так как произведения искусства, созданные в одних странах и оказавшиеся в учреждениях культуры и музеях других стран (например, деревянная скульптура из Анголы во французском музее, представляющем искусство Африки, Азии и Океании⁴⁰⁸), можно обнаружить практически для каждой страны. Это возможно с помощью раздела платформы GA&C «Откройте для себя это место». Однако результаты исследования показывают, что даже для этой группы количественные показатели первой страны в списке (США) настолько превосходят другие страны, что одно из четырех изображений, которые пользователь может увидеть в агрегаторе, с большой вероятностью будет включать метаданные, относящиеся к этой стране, если только GA&C не меняет эти результаты с помощью алгоритмов, которые направлены на предоставление пользователю более сбалансированной картины.

Даже если исключить коллекцию фотожурнала LIFE из анализа, связанного с репрезентативным представлением учреждений культуры в GA&C, мы получим аналогичные результаты, поскольку коллекции США являются самыми многочисленными, особенно если учесть, что новые коллекции учреждений культуры и музеев США постоянно добавляются к платформе⁴⁰⁹. Данные диссертационного исследования показывают, что только 72 страны из 195 стран из списка ООН публикуют образцы коллекций своих музеев с помощью GA&C. 123 страны не взаимодействуют с агрегатором и не публикуют с его помощью коллекции своих музеев, архивов и библиотек, что означает, что GA&C не является репрезентативным или сбалансированным агрегатором с точки зрения представления учреждений культуры. Если рассматривать произведения искусства и культуры в коллек-

⁴⁰⁸ URL: <https://artsandculture.google.com/asset/king-playing-the-sanza/bwHHZEiYNmt8bA>

⁴⁰⁹ По состоянию на август 2019 года в GA&C появились новые визуальные коллекции Национального управления по авиации и исследованию космического пространства (НАСА), которые включают около 160 000 изображений, при этом цифровые коллекции США в GA&C оказываются самыми многочисленными, даже если мы исключаем коллекцию фотожурнала LIFE.

циях других стран, можно отметить, что контент GA&C является репрезентативным и, возможно, более сбалансированным. Можно заметить, что почти все страны из списка ООН представлены в метаданных или текстах, сопровождающих изображение, или относятся к городу, в котором родился художник. Однако количество объектов, относящихся к первым девяти странам, то есть около двух миллионов изображений произведений искусства, показывают, что мы не можем сказать, что платформа GA&C представляет собой сбалансированный корпус визуального контента, поскольку девять стран представляют около 40% визуального контента платформы.

При подробном анализе сайтов пяти музеев и учреждений культуры, предоставивших максимальное количество изображений для публикации на платформе GA&C, можно отметить, что три из них - Музей естествознания в Лондоне⁴¹⁰, Музей Метрополитен в Нью-Йорке⁴¹¹ и Рейксмузеум в Амстердаме⁴¹² предлагают получить доступ к данным и загрузить их с помощью API⁴¹³. Это означает, что музеи могут создавать новые партнерские отношения с технологическими компаниями, включая Google Arts and Culture, в частности, для изучения того, как использовать визуальные данные для исследовательских, промышленных и коммерческих приложений⁴¹⁴, а также для лицензирования использования изображе-

⁴¹⁰ URL: <https://www.nhm.ac.uk/>

⁴¹¹ URL: <https://www.metmuseum.org/>

⁴¹² URL: <https://www.rijksmuseum.nl/en>

⁴¹³ URL: <https://data.nhm.ac.uk/about/download>, <https://www.metmuseum.org/blogs/now-at-the-met/2018/met-collection-api>, <https://www.rijksmuseum.nl/en/api>

⁴¹⁴ Tallon, L. (2018). Scaling the Mission. The Met Collection API. Blogs/Now at the Met. The Metropolitan Museum of Art. 25 October, 2018. <https://www.metmuseum.org/blogs/now-at-the-met/2018/met-collection-api>, Valeonti, F., Terras, M., Hudson-Smith, A., Zarkali, Ch. (2018). Exploring new business models for monetizing digitization beyond image licensing to promote adoption of OpenGalm. EVA Berlin 2018 Conference, 7 - 9 November 2018. <https://discovery.ucl.ac.uk/id/eprint/10075085/1/Exploring%20new%20Business%20Models.pdf>

ний при изготовлении товаров и предоставлении услуг⁴¹⁵. Таким образом, стремление к открытому лицензированию и OpenGlam в некоторых странах (преимущественно в США и Европе), направленное на выполнение миссии учреждений культуры и достижение образовательных и просветительских целей, может привести к тому, что в новом цифровом каноне коллекции этих стран и учреждений займут основное место.

Таким образом, результаты диссертационного исследования показывают, что GA&C не является сбалансированным корпусом данных с точки зрения географического распределения для всех стран из списка ООН. Далее, диссертационное исследование рассматривает географическое распределение учреждений культуры, представленных в GA&C для двух стран, России и Франции. По состоянию на весну 2019 года коллекции российских федеральных, региональных и муниципальных музеев в GA&C были представлены 32 учреждениями культуры и 2 844 предметами. При сравнении их географического положения с местоположениями 2 189 музеев, зарегистрированных в Государственном каталоге музейных коллекций РФ, можно обнаружить, что российские учреждения культуры, представленные в GA&C, чаще всего находятся в центральной части России и в Москве. GA&C представляет семь российских регионов или 6 % из восьмидесяти пяти регионов страны. Однако географическое распределение российских музеев с оцифрованным контентом, опубликованном в Государственном каталоге МФ РФ, значительно шире и охватывает практически все регионы Российской Федерации. Это значит, что, несмотря на наличие оцифрованных коллекций в музеях страны, на международном уровне представлено небольшое количество предметов из коллекций центральных музеев.

Чтобы выяснить, является ли российская коллекция агрегатора репрезентативной с точки зрения периодов развития культуры и искусства, рассматривался

⁴¹⁵ Kapsalis, E. (2016). "The Impact of Open Access on Galleries, Libraries, Museums and Archives". Smithsonian Emerging Leaders Development Program. URL: http://siarchives.si.edu/sites/default/files/pdfs/2016_03_10_OpenCollections_Public.pdf

промежуток времени, когда было создано произведение, опубликованное на платформе GA&C. Государственный каталог МФ РФ не предоставляет возможности сравнивать распределения предметов, опубликованных на этой платформе по периодам времени. Результаты, представленные в данном диссертационном исследовании, показывают распределение периодов времени для предметов из российских коллекций, опубликованных на платформе GA&C. Произведения искусства двадцатого века составляют почти 60 % коллекции, в то время как предметы, созданные значительно ранее (до двенадцатого века), представлены менее чем 1 % всех изображений. Произведения живописи, фотографии и декоративно-прикладное искусство составляют половину изображений из российских музеев в GA&C, в то время как оружие, предметы естественнонаучных коллекций и скульптура представлены в меньшем количестве по сравнению с Государственным каталогом МФ РФ. Живопись представлена в GA&C большим количеством произведений по сравнению с Государственным каталогом. Результаты для пяти авторов с максимальным количеством произведений в GA&C показывают, что это художники двадцатого века, дизайнеры тканей и костюмов, поэты и фотографы. Одним из самых частотных российских авторов в российских коллекциях GA&C является Лев Самойлович Бакст, дизайнер костюмов и декораций для Русского балета Дягилева, который в начале двадцатого века повлиял на развитие моды в Европе и Северной Америке. Газетные статьи, освещающие творчество Бакста и рассказывающие о его работах, относятся к серии вырезок, опубликованных в GA&C как часть контента, представленного для Года моды в 2016 году⁴¹⁶. Газетные вырезки, посвященные Баксту, составляют основную часть (40 %) музейной коллекции. Авторы статьи, сопровождающей оцифрованные материалы, пи-

⁴¹⁶ Shumanova, I., Zoya Shergina, Tatiana Nikolaeva, Anna Grishina, Maria Konstanistova, Maria Krivenko, Katrin Krylova, Elena Rymshina, Elizaveta Yakovleva, Svetlana Suvorova, Igor Khlystov, Elena Misuchenko. (2019). 'The Collector of Success: Leon Bakst'. Google Arts and Culture. URL: <https://artsandculture.google.com/exhibit/KgICSXEITyGILA>

шут: «Специально для проекта «Год моды» мы просмотрели архив Бакста, чтобы найти газетные вырезки, посвященные этому явлению»⁴¹⁷.

В список пяти самых частых авторов, представленных в GA&C, входит Владимир Маяковский, чьи записные книжки демонстрируются платформой. Следующей российской коллекцией с максимальным количеством произведений одного автора являются художественные фотографии, сделанные в начале XX века. Научный сотрудник музея Ф. И. Тютчева в Мураново пишет: «Фотографии дают нам представление о неувядающей красоте и величии жизни ушедшей эпохи. Некоторые из работ фотографов-любителей настолько высокого качества, что их можно по праву назвать подлинными произведениями искусства»⁴¹⁸.

Анализ российских коллекций GA&C показывает, что распределение произведений искусства и культуры не является равномерным, при этом произведения, созданные в двадцатом веке, представлены в значительно большем количестве. Пользователи с большей вероятностью увидят в российской части коллекций GA&C картины и фотографии двадцатого века по сравнению с произведениями других жанров или периодов времени.

В данном параграфе было показано, что коллекции Франции, представленные в GA&C, сопоставимы с коллекциями России по ряду параметров. Анализ, представленный в диссертационном исследовании, подтверждает, что результаты, полученные для России, не зависят от страны, а скорее являются редакционными принципами, которые определяют то, как платформа GA&C представляет цифровой контент.

⁴¹⁷ Shumanova, I., Zoya Shergina, Tatiana Nikolaeva, Anna Grishina, Maria Konstanistova, Maria Krivenko, Katrin Krylova, Elena Rymshina, Elizaveta Yakovleva, Svetlana Suvorova, Igor Khlystov, Elena Misuchenko. (2019). 'The Collector of Success: Leon Bakst'. Google Arts and Culture. URL: <https://artsandculture.google.com/exhibit/KgICSXEITyGILA>

⁴¹⁸ URL: <https://artsandculture.google.com/exhibit/DAJSWSShnjg7LQ>

Поскольку незарегистрированные музеи⁴¹⁹ и не музейные учреждения культуры не рассматривались в данном исследовании, анализ проводился для 21 музея и 6 575 музейных предметов, опубликованных с помощью платформы GA&C. Следует отметить, что музеи, расположенные в Париже, представлены значительно лучше по сравнению с региональными музеями, особенно теми, которые расположены далеко от Парижа. Как можно заметить, Иль-де-Франс, центральный французский регион, в котором находится Париж, представлен одиннадцатью музеями или 52,4 % музеев в выборке. Париж представлен восемью музеями или 38,1 % всех изображений. Франция как страна, включающая разные провинции, недостаточно представлена, поскольку в коллекцию GA&C входят музеи из восьми провинций или 44,5% из восемнадцати провинций страны. В целом, можно отметить тенденцию, которую мы уже наблюдали в российских коллекциях в GA&C: представленные музеи расположены в столице или в регионе, который включает столицу. Провинции, расположенные на расстоянии от столицы, представлены недостаточным количеством изображений или не представлены совсем.

Данный параграф диссертационного исследования демонстрирует также распределение произведений искусства по периодам времени во французских коллекциях GA&C и в базе данных Joconde. Результаты для GA&C ясно показывают преобладающее количество произведений, созданных в двадцатом веке (41,7 %), за которым следует девятнадцатый век (36,5 %). Неожиданный результат – большое количество произведений двадцать первого века (почти 14 %) во французских коллекциях GA&C, учитывая, что в исследование не были включены незарегистрированные музеи и личные галереи. В базе данных Joconde можно отметить преобладание произведений искусства XIX века (45,7 %) и значительно меньший процент произведений искусства XXI века (0,3 %). В то время как все произведения искусства, созданные до семнадцатого века, составляют около 7 % в базе данных Joconde, их количество составляет 1,1 % во французских коллекциях GA&C. Это означает, что пользователи, ожидающие увидеть фран-

⁴¹⁹ Музеи, не включенные в Реестр Министерства культуры Франции.

цузское искусство, созданное до XVII века, обнаружат только около 70 произведений искусства того периода, в то время как во французской коллекции GA&C представлены почти 1 000 произведений искусства и культуры двадцать первого века и около 2 700 произведений искусства и культуры двадцатого века.

Пять французских авторов с максимальным количеством произведений, представленных в коллекциях, опубликованных Google Arts and Culture, включают дизайнера двадцатого века и фотографов девятнадцатого века. Работы французского дизайнера Маргаритты Порракиа представлены 1 200 изображениями, или почти 20 % французских коллекций, публикуемых платформой GA&C. Они датируются первой половиной двадцатого века, и, вероятно, редакторы GA&C включили их благодаря работе над коллекциями, посвященными Году моды в 2016 году. В то же время кураторы музея, возможно, представили работы, которые продемонстрировали изменение в каноне моды, связанном с культурными заимствованиями в начале двадцатого века, в частности, связанные с ориентализмом или «конструированием других географических пространств»⁴²⁰. Иллюстрируя свой рассказ о предметах Музея декоративного искусства в Париже работами Маргаритты Порракиа, научный сотрудник музея С. Имберт пишет: «По мере развития международной торговли многие товары из Азии и Ближнего Востока появились в Европе и повлияли на стиль одежды. Такой культурный обмен обогатил декоративный репертуар западной моды и способствовал ее обновлению»⁴²¹.

Вторым автором с максимальным количеством изображений французских коллекций GA&C является Пьер Тремо, французский фотограф девятнадцатого века, который фиксировал культуру Африки и Азии и который, как предполагает К. Аддельман-Франкель⁴²², был частью феномена колониальной культуры. Одна-

⁴²⁰ Said, EW. (2000). Invention, memory, and place. *Critical Inquiry*, 26: 175–192.

⁴²¹ Imbert, C. (2019). ‘Transfert(s): Western fashion meets the East’. Google Arts and Culture. URL: <https://artsandculture.google.com/exhibit/EwISenSdmLJ8IQ>

⁴²² Addleman-Frankel, K. (2018). The Experience of Elsewhere: Photography in the Travelogues of Pierre Trémaux. *photographies*, 11: 31–56.

ко если проанализировать частотность фамилий авторов для изображений, представленных некоторыми музеями для российской и французской коллекций GA&C, такими, например, как Государственный Эрмитаж в Санкт-Петербурге или Музей изобразительных искусств в Лионе, мы увидим, что они представляют канон европейского искусства, который мы можем найти в печатных публикациях по истории искусства по всему миру.

На вопрос, предоставляют ли редакторы GA&C рекомендации относительно того, какой тип контента необходим для публикации на сайте, сложно ответить однозначно. Вообще говоря, редакторы GA&C могут руководствоваться предположением, что пользователям сайта необходим популярный и востребованный контент и что учреждения культуры, предоставляющие изображения, должны научиться «работать с другой аудиторией»⁴²³. Иногда редакторам требуется контент, который соответствует определенному проекту или теме, например, проекту «Год моды». Мы не можем прийти к заключению, что каждый музей получает рекомендации относительно того, какие изображения необходимы агрегатору, но процесс отбора контента, похоже, соответствует наблюдению Э. К. Мэнсфилд⁴²⁴, что GA&C создает «одну из тех инфраструктур знаний»⁴²⁵, которые не имеют отношения к научным исследованиям. При этом методы отбора изображений не публикуются или не известны широкой публике.

⁴²³ The Art Newspaper. (2018). How Google became a major producer of cultural content. 19 January, 2018. URL: <https://www.theartnewspaper.com/feature/how-google-became-a-major-producer-of-cultural-content> (accessed 6 March 2020)

⁴²⁴ Mansfield, E. C. (2014) Google Art Project and Digital Scholarship in the Visual Arts. URL: <http://www.google.com/culturalinstitute/project/art-project>, Visual Resources: An International Journal of Documentation, 30 (1): 110-117.

⁴²⁵ Mak, B. (2014). Archaeology of a Digitization. Journal of the Association for Information Science and Technology, 65(8): 1515–26.

В результате анализа, представленного в данном параграфе диссертационного исследования и в работе И. Кижнер и др.⁴²⁶, можно выделить две основные проблемы и обозначить ограничения исследования. Во-первых, подразумевают ли результаты исследования новый цифровой культурный колониализм, когда агрегаторы подчеркивают и выделяют определенные культуры и периоды за счет других культур? Во-вторых, какими должны быть принципы организации больших наборов данных оцифрованного визуального контента, представляющего культурное наследие.

Хотя результаты данного параграфа показывают, что ряд стран представлен GA&C большим количеством изображений, мы не знаем, предпринимались ли редакторами GA&C попытки сделать корпус более репрезентативным при отображении результатов поиска. Мы также не знаем, насколько это зависит от типа поиска. Представлены ли термины, относящиеся к истории двадцатого века, такие как «Маргарет Тэтчер», «холокост» или «колхозы», большим количеством изображений, которые хранятся в учреждениях культуры небольшого количества стран? Относится ли это к более широким терминам, таким как «лошадь», «чаша» или «ювелирное украшение»? Насколько разнообразными будут результаты для более широких терминов? Они будут включать изображения предметов, которые хранятся в музеях и учреждениях культуры, которые представлены в GA&C максимальным количеством изображений, или в этом списке появятся изображения предметов из музеев других стран? Каково будет процентное соотношение для каждой страны в этом случае? В исследовании представлены результаты по количеству изображений, опубликованных в GA&C учреждениями культуры разных стран. Исследование не касается вопроса о том, что происходит с данными, когда они отображаются в качестве результатов поиска и делает ли это агрегатор более репрезентативным и сбалансированным. Алгоритм поиска так же важен, как и контент, хотя остается черным ящиком (а сами алгоритмы поиска

⁴²⁶ Kizhner et al. Digital Cultural Colonialism: measuring bias in aggregated digitized content held in Google Arts and Culture. *Digital Scholarship in the Humanities* (в печати).

демонстрируют проявления предрассудков и тенденциозности), что показано в некоторых исследованиях⁴²⁷.

Результаты исследования, связанные с распределением цифрового контента по времени, жанрам искусства и авторам с максимальным количеством произведений, относятся только к изображениям произведений искусства из Франции и России. Их недостаточно для того чтобы делать заключения о распределении содержания агрегатора в целом. С другой стороны, LIFE Photo Collection составляет 76 % контента GA&C, что может привести к тому, что для некоторых результатов поиска приоритет будет отдан фотографической журналистике 20-го века, представленной через призму восприятия фотографов США.

В исследовании не показано, в какие страны и учреждения культуры редакторы GA&C направили запросы о предоставлении изображений и каково процентное соотношение этих запросов для разных стран. В исследовании также не рассматривается, что могло повлиять на решение редакторов о включении изображений в агрегатор. Неизвестно географическое распределение ответов на запросы о предоставлении изображений и не можем сравнить его с нашими результатами по географическому распределению коллекций Google Arts and Culture. Исследование также не касалось возможных планов редакторов GA&C, связанных с решением проблемы отсутствия сбалансированности в агрегаторе. Сравнение результатов для французской и российской коллекций проводилось при анализе национальных агрегаторов изображений предметов культурного наследия, опубликованных Министерством культуры РФ и Министерством культуры Франции. В то время как Государственный каталог музейных коллекций РФ, представляющий на момент написания диссертационного исследования 20 % музейных предметов страны из всех регионов России, можно считать репрезентативной выборкой, французская база данных музейных предметов включает около 400 000 изображений или 0,3 % музейных предметов французских музейных кол-

⁴²⁷ Noble S.U. (2018). Algorithms of oppression: How search engines reinforce racism. New York: NYU Press.

лекций, которые включают более 121 миллиона музейных предметов⁴²⁸. Такая выборка не может считаться достаточной для того, чтобы сделать заключение о ситуации, наблюдаемой для французских музейных коллекций. Однако это не меняет нашего вывода о том, что GA&C отображает больше произведений искусства двадцатого века, чем предметов, которые представляют какие-либо другие периоды, или что они в основном представляют собой произведения живописи, графики, фотографического и декоративно-прикладного искусства для обеих стран в выборке.

Результаты данного параграфа диссертационного исследования подтверждают выводы, представленные в работах, направленных на изучение постколониальной культуры и репрезентацию культуры в музеях и иных учреждениях. Таким образом, выводы диссертационного исследования переносят представления неокOLONиализме в цифровую сферу и показывают, как эти представления усиливаются путем агрегирования цифровых изображений предметов культурного наследия. Существующие работы также объясняют результаты, показывающие преобладание художественных произведений двадцатого века в российских и французских коллекциях. «Художественные произведения, представленные в максимальном количестве, получают преимущество»⁴²⁹, и результаты этого преимущества особенно заметны, когда исследователи приступают к анализу больших данных. Диссертационное исследование показывает, как объяснения, представленные в работах в области колониальных теорий культуры, имеют отношение к нашим результатам с точки зрения 1) географического распределения и 2) распределения доминирующих авторов во французских и российских коллекциях.

⁴²⁸ Joconde Database. (2019). Présentation. URL: <http://www2.culture.gouv.fr/documentation/joconde/fr/apropos/presentation-joconde.htm> (accessed 10 March 2020)

⁴²⁹ Boast, R. (2011). Neocolonial collaboration: museum as contact zone revisited. *Mus Anthropol* 34:56–70, p. 64.

Результаты исследования демонстрируют отсутствие сбалансированности в географическом распределении изображении предметов культурного наследия для коллекций, опубликованных в GA&C, и отсутствие репрезентативности/сбалансированности в географическом распределении учреждений культуры, в которых хранятся культурные объекты. Это хорошо согласуется с исследованиями, показывающими, как цифровой колониализм и возникающие в результате структуры данных отражают «ожидания культуры, которая доминирует в репрезентации событий и явлений благодаря технологическим достижениям»⁴³⁰. Преобладание предметов культурного наследия из американских, британских и голландских культурных учреждений может быть следствием принципа прозрачности представления данных, который приводит к публикации и распространению данных⁴³¹.

Другая причина преобладания изображений, предоставленных западными учреждениями культуры, заключается в том, что для некоторых из них с данными можно взаимодействовать с помощью API. В результате большие данные и ранее оцифрованные изображения оказываются опубликованными и предоставляются широкой публике и коммерческим компаниям, которые становятся партнерами музеев и других учреждений культуры. Таким образом прозрачность данных, политика открытого доступа и другие стратегии позволяют строить партнерские отношения, используя и, возможно, монетизируя возможности изучения визуальных

⁴³⁰Loukissas, Y. A. 2019. *All Data Are Local: Thinking Critically in a Data-Driven Society*. Cambridge, Mass.: The MIT Press, p. 10; Risam R. (2019). *New Digital Worlds: Postcolonial Digital Humanities in Theory, Praxis, and Pedagogy*. Evanston, Illinois: Northwestern University Press.

⁴³¹Hansen, H. K., Flyverbom, M. (2014). The politics of transparency and the calibration of knowledge in the digital age. *Organization*, 22: 872-889; Schudson, M. (2015). *The Rise of the Right to Know*. Cambridge, MA: Belknap Publishing; Albu, O. B., Flyverbom, M. (2016) Organizational transparency: conceptualizations, conditions, and consequences. *Business & Society*, 58 (2): 268-297.

данных⁴³². В этом случае музейные изображения появляются на нескольких платформах, «при этом коллекцию можно посмотреть из нескольких мест (что само по себе помогает выполнить миссию музея)»⁴³³. Предыдущие исследования показывают, что в случае российских музеев предоставление открытого доступа и, следовательно, доступа к API, связано с ограничениями, связанными с российским законодательством⁴³⁴.

С другой стороны, большое количество данных, представленное рядом стран с максимальным количеством изображений в GA&C, может быть следствием колониальной практики работы с музейными коллекциями, когда именно в западных центрах проходила интерпретация и описание предметов из музейных коллекций⁴³⁵. Повсеместное распространение изображений ведущих учреждений культуры с максимальным количеством изображений в GA&C создает критическую массу концепций, взглядов и идеологий, которые вытесняют другие культуры и идеологии. Это означает, что постмодернистское разнообразие точек зрения⁴³⁶, которое могло входить в намерения редакторов GA&C⁴³⁷, было реализова-

⁴³² Valeonti F., Terras M., Hudson-Smith A., Zarkali Ch. Exploring New Business Models for Monetising Digitisation Beyond Image Licensing to Promote Adoption of OpenGlam. EVA Berlin 2018 Conference, 7 - 9 November 2018. URL: <https://discovery.ucl.ac.uk/id/eprint/10075085/1/Exploring%20new%20Business%20Models.pdf> (accessed 6 March 2020)

⁴³³ Tallon L. (2018). Scaling the Mission. The Met Collection API. Blogs/Now at the Met. The Metropolitan Museum of Art. 25 October, 2018. <https://www.metmuseum.org/blogs/now-at-the-met/2018/met-collection-api> (accessed 10 March 2020).

⁴³⁴ Kizhner I., Terras M., Romyantsev M., Sycheva K., Rudov I. (2019). Accessing Russian Culture Online: The scope of digitization in museums across Russia. *Digital Scholarship in the Humanities*, 34(2): 350–367.

⁴³⁵ Beaulieu J., Roberts M. (2002.) *Orientalism's Interlocutors: Painting, Architecture, Photography*, Durham: Duke University Press.

⁴³⁶ Connor S. (1991). *Postmodernist Culture: An Introduction to Theories of the Contemporary*. Oxford: Blackwell.

но таким образом, что это привело к усилению колониальных представлений, укреплению предрассудков и канонов печатной эпохи⁴³⁸, встроенных в колониальные представления прошлых эпох. Постколониальное и постмодернистское намерение изменить существующую систему⁴³⁹ должно было привести к мультикультурным и разнообразным коллекциям изображений, представляющих культуры, регионы и периоды в цифровом мире, значительно расширив доступ к информации через платформу GA&C. Это не привело к изменению системы. Впрочем, этот результат не наблюдался и при предполагаемых изменениях в сторону большего разнообразия культур и точек зрения в физическом музее⁴⁴⁰.

Такой вид дискриминации усиливается благодаря доминированию произведений искусства 20-го века в русских и французских коллекциях и, по-видимому, во всем агрегаторе. Кроме того, большое количество изображений из фотоархивов XX века в агрегаторе приводит к созданию представлений о воображаемых культурах и сообществах⁴⁴¹, а также созданию представлений о воображае-

⁴³⁷ Sood A. (2016). Every piece of art you've ever wanted to see - up, close and searchable. TED lecture. URL: https://www.ted.com/talks/amit_sood_every_piece_of_art_you_ve_ever_wanted_to_see_up_close_and_searchable (accessed 10 March 2020).

⁴³⁸ Earhart A.E. (2012). Can information be unfettered?: Race and the new digital humanities canon. In Gold M. K., (ed.), *Debates in the Digital Humanities*. Minneapolis: University of Minnesota Press, pp. 309-318; Zaagsma G. (2019). Digital History and the Politics of Digitization. Presented at Digital Humanities 2019 Conference, 9-12 July, 2019.

⁴³⁹ Harrison J. (1997). Museums as Agencies of Neocolonialism in a Postmodern World. *Studies in Cults, Organizations and Societies*, 3(1): 41–65.

⁴⁴⁰ Harrison J. (1997). Museums as Agencies of Neocolonialism in a Postmodern World. *Studies in Cults, Organizations and Societies*, 3(1): 41–65; Knell S. (2010). Introduction. In Knell S. J, Aronsson P., Amundsen, A.B. (eds), *National Museums: New Studies from Around the World*, London: Routledge; Knell, S. (2018). *The Contemporary Museum. Shaping Museum for the Global Now*. Routledge: London and New York.

⁴⁴¹ Anderson B. (2006). *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. Verso Books.

мых/конструируемых географических пространствах⁴⁴² на основе выборочных данных, представленных через призму фотографов и кураторов фотоархивов. К тому же, научные сотрудники музеев отбирают произведения, соответствующие их представлениям о том, как должны видеть культуру люди из других стран, что приводит к созданию этнографических историй, подготовленных сотрудниками учреждений культуры⁴⁴³, но сформированных западными ожиданиями. Однако большое количество произведений искусства XX века, представленное в GA&C, может объясняться тем, что это является реакцией музеев и учреждений культуры на прогнозируемые ожидания пользователей и высокий спрос на современное искусство со стороны современных потребителей.

Предпочтение, которое редакторы GA&C отдают современному искусству, может означать, что большие части музейных коллекций в России и Франции включают произведения искусства и фотографии с начала 20-го века до настоящего времени. Это явно не относится к французским музеям, если база данных «Joconde» отражает состояние цифровых коллекций. Впрочем, база данных «Joconde» может и не отражать стратификацию коллекций французских музеев. Распределение по периодам времени для оцифрованных произведений искусства из российских музеев невозможно узнать, поскольку Госкаталог не предоставляет этой возможности. Неизвестна также доля современного искусства в других коллекциях оцифрованных произведений по всему миру, хотя ряд крупных национальных музеев предоставляют функцию фильтрации коллекции по периодам времени с помощью поисковых системы своих веб-сайтов. Преобладание современного искусства во французских и российских коллекциях GA&C может зависеть от предпочтений тех научных сотрудников музея, которые взаимодействовали с редакторами портала GA&C и знаменует собой переход от экспертного мне-

⁴⁴² Said E.W. (2000). Invention, memory, and place. *Critical Inquiry*, 26: 175–192.

⁴⁴³ Boast R. (2011). Neocolonial collaboration: museum as contact zone revisited. *Mus Anthropol* 34:56–70.

ния по отношению к периодизации произведений искусства⁴⁴⁴ к акцентированию тех изменений в каноне истории искусства, которые произошли между концом XIX века до настоящего времени. Об этом свидетельствуют также те результаты диссертационного исследования, которые касаются авторов с максимальным количеством произведений в российской и французской коллекциях GA&C. Изменения канона на рубеже XIX и первых десятилетий XX века привели к широкому культурному обмену и взаимному влиянию культур. Эти изменения включали, среди прочего, фотографию как вид искусства⁴⁴⁵, фотографию как способ репродукции, пропагандирующий колониальную культуру и/или взаимное влияние культур⁴⁴⁶, изменения в дизайне модной одежды, связанные с необходимостью поиска источников вдохновения в другой культуре⁴⁴⁷, и изменения в стиле русской поэзии, о чем свидетельствуют изменения в поэтическом ритме и метре⁴⁴⁸. Все эти явления представлены авторами с максимальным количеством изображений во французской и российской коллекциях GA&C. Значительная часть авторов, представленных максимальным количеством изображений, имеют отношение к репрезентации колониальной культуры, когда «воображаемые сообщест-

⁴⁴⁴ Knell, S. (2018). *The Contemporary Museum. Shaping Museum for the Global Now*. Routledge: London and New York.

⁴⁴⁵ Beaulieu J., Roberts M. (2002.) *Orientalism's Interlocutors: Painting, Architecture, Photography*, Durham: Duke University Press.

⁴⁴⁶ Osborne P. D. (2000). *Travelling Light: Photography, Travel and Visual Culture*. Manchester: Manchester UP, 2000; Ryan J.R. (1997). *Picturing Empire: Photography and the Visualization of the British Empire*. Chicago: University of Chicago Press; Behdad A. and Gartlan L. (2013). *Photography's Orientalism: New Essays on Colonial Representation*, Los Angeles: The Getty Research Institute.

⁴⁴⁷ Martin R., Koda H. (1994). *Orientalism: Visions of the East in Western Dress*. New York: Metropolitan Museum of Art.

⁴⁴⁸ Ehrlich V. (1952). 'Russian Poets in Search of a Poetics'. *Comparative Literature*, 4 (1): 54-74.

ва»⁴⁴⁹ и экзотические, непохожие на западные страны были представлены Западу через призму западного восприятия, с помощью подходящих самопрезентаций. Создание представлений об отдаленных регионах часто сопровождалось распространением колониальных дискурсов, созданных авторами с максимальным количеством изображений французских коллекций, опубликованных на веб-сайте GA&C. К. Аддельман-Франкель⁴⁵⁰ показывает, как колониальная фотография XIX века, в частности, фотографии Пьера Тремо (второго автора с максимальным количеством произведений во французских коллекциях), демонстрирует развитие объективного взгляда на культуру колониальных стран и конкурирует с графическими изображениями, которые служили иллюстрациями в книгах, описывающих географическое положение и культуру колониальных стран. Несомненно, такие фотографии демонстрировали изменение культурного канона, и научные сотрудники музеев и других учреждений культуры могли считать их удачной находкой для международного агрегатора изображений.

Отсутствие сбалансированности в параметрах, связанных с периодом и типом экспонатов, согласуется с нашими выводами о количестве изображений, представляющих различные страны в агрегаторе, и доказывает смещение акцента в сторону постколониального представления культурных объектов. Такое явление может быть следствием намерения кураторов представить свои коллекции так, чтобы они были интересны международной аудитории. Это может быть следствием предвзятости по отношению к подобранным художественным произведениям, следствием изменения художественного канона и канона репрезентации, когда фотография начала заменять другие средства представления образов как способ более объективный и не менее художественный. Влияние изменений в каноне на подбор иллюстративных примеров для корпуса подтверждается литературой

⁴⁴⁹ Anderson B. (2006). *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. Verso Books.

⁴⁵⁰ Addleman-Frankel K. (2018). *The Experience of Elsewhere: Photography in the Travelogues of Pierre Trémaux*. *photographies*, 11: 31-56.

по сбору лингвистических или литературных примеров при составлении лингвистических корпусов. Г. С. Мортон⁴⁵¹ показывает, как создатели корпуса для Оксфордского словаря английского языка полагались на читателей-добровольцев, которых больше привлекали необычные слова, чем слова, которые часто встречаются в книгах, которые они читали⁴⁵². А. Грив-Смит показывает, что при составлении исторического корпуса литературного французского языка редакторы пользовались хрестоматиями, составленными влиятельными литературоведами, что привело к предвзятости и отсутствию сбалансированности в корпусе:

«Принцип авторитетности и влиятельности автора хрестоматии, использованный при выборе текстов для FRANTEXT, привел к другому источнику тенденциозности в корпусе: авторы и редакторы хрестоматий и важных литературоведческих трудов участвовали в его создании. Мало того, что эти авторы были предвзяты в силу своего социального происхождения и образования, самое важное заключалось в том, что цель этих литературоведческих работ заключалась в информировании публики об интересных событиях и изменениях в литературном каноне в рамках исследуемого периода. У литературоведов не было цели представить весь спектр произведений, созданных в то время».⁴⁵³

Тот факт, что коллекция фотожурнала LIFE составляет 76 % контента агрегатора и что фотография представляет значительную часть российских коллекций (и, возможно, коллекций некоторых других стран), превышая долю этого жанра в Госкаталоге РФ МФ, свидетельствует об обращении к этому жанру искусства как важному инструменту создания «воображаемых сообществ»⁴⁵⁴ и своего рода

⁴⁵¹ Morton H.C. *The Story of Webster's Third: Philip Gove's Controversial Dictionary and Its Critics*. Cambridge: Cambridge University Press, 1994, с. 95.

⁴⁵² Цитируется по Biber D. (1993). *Representativeness in Corpus Design*. *Literary and Linguistic Computing*, 8 (4): 243-257, с. 22.

⁴⁵³ Grieve-Smith A. (2019). *The importance of sampling frames in representative historical corpora: a case study of Parisian theater*. *CogniTextes*. <https://journals.openedition.org/cognitextes/1671>

⁴⁵⁴ Anderson B. (2006). *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. Verso Books.

способу оспорить достижения других жанров искусства⁴⁵⁵. Доминирование изображений из столиц и регионов, окружающих столицы, в российских и французских коллекциях также способствует нашему пониманию агрегатора как корпуса, который наследует тенденциозность эпохи печатных изданий. С другой стороны, мы не знаем, обращались ли редакторы GA&C к музеям из отдаленных регионов и связано ли их отсутствие в агрегаторе с намерением музеев контролировать появление коллекций на внешних платформах и держать их закрытыми. Эта гипотеза согласуется с нашими предыдущими результатами по российским цифровым коллекциям⁴⁵⁶ и выводами о том, что лишь немногие музеи в мире проводят политику открытого доступа⁴⁵⁷. Ряд коллекций приходится исключать из агрегатора по объективным причинам, связанным со сложностью получения изображений произведений искусства из мест, к которым трудно получить доступ по различным технологическим, этическим или социальным причинам⁴⁵⁸. Это может быть еще одной причиной тенденциозности и отсутствия сбалансированности в агрегаторах культурных данных. Каковы бы ни были намерения редакторов GA&C, и они могут иметь прямое отношение к доступности и разнообразию в представлении искусства и культуры⁴⁵⁹, тенденциозность и отсутствие репрезентативности неизбежно проникает в корпус изображений так же, как она неизбежно присутст-

⁴⁵⁵ Addleman-Frankel K. (2018). The Experience of Elsewhere: Photography in the Travelogues of Pierre Trémaux. *photographies*, 11: 31-56.

⁴⁵⁶ Kizhner I., Terras M., Rumyantsev M., Sycheva K., Rudov I. (2019). Accessing Russian Culture Online: The scope of digitization in museums across Russia. *Digital Scholarship in the Humanities*, 34(2): 350–367.

⁴⁵⁷ McCarthy D., Wallace A. (2020). Survey of GLAM open access policy and practice. Google Sheets. URL: https://docs.google.com/spreadsheets/d/1WPS-KJptUJ-o8SXtg00llcxq0IKJu8eO6Ege_GrLaNc/edit#gid=1216556120

⁴⁵⁸ Bode K. (2020). Why you can't model away bias, *Modern Language Quarterly* 81.1.

⁴⁵⁹ Sood A. (2016). Every piece of art you've ever wanted to see - up, close and searchable. TED lecture. URL: https://www.ted.com/talks/amit_sood_every_piece_of_art_you_ve_ever_wanted_to_see_up_close_and_searchable

вует в лингвистических и литературоведческих корпусах из-за предшествующего социального устройства, структур, предрассудков и канонов⁴⁶⁰. Однако, будучи частью коммерческого гиганта, GA&C может направлять и структурировать свои данные с тем, чтобы алгоритмы производили больше контента того типа, который востребован пользователем, и того типа, который необходим по каким-то иным причинам⁴⁶¹, усиливая существующие предубеждения, укрепляя тенденциозность и цифровой культурный колониализм. В отличие от лингвистического корпуса⁴⁶², научного издания или системы, созданной для работы с литературоведческими данными⁴⁶³, GA&C не предлагает обоснования выбора коллекций и изображений, не публикует ограничения корпуса и предупреждения о тенденциозности, которая может возникнуть в результате анализа больших данных.

Результаты диссертационного исследования, связанные с отсутствием репрезентативности и сбалансированности в агрегаторе Google Arts and Culture крайне важны еще и потому, что обнаруженная тенденциозность может повлиять на будущие платформы и их пользователей. В существующей литературе подчеркивается необходимость наличия изображений произведений искусства и культуры и систем описания этих произведений как наборов данных для обучения нейронных сетей эффективным действиям в поисковых системах. Это означает, что поисковые системы смогут отвечать на прямые визуальные запросы, определять предмет, не работая с текстовыми метаданными, и находить похожие объекты

⁴⁶⁰ Bode K. (2020). Why you can't model away bias, *Modern Language Quarterly* 81.1; Caliskan A., Bryson J. J., Narayanan A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science* 356 (6334): 183–86.

⁴⁶¹ Flyverbom M., Murray J. (2018). Datastructuring – Organizing and curating digital traces into action, *Big Data and Society* 5(2): 1–12; Zuboff S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: Profile Books.

⁴⁶² Biber D. (1993). Representativeness in Corpus Design. *Literary and Linguistic Computing*, 8 (4): 243-257.

⁴⁶³ Bode K. (2020). Why you can't model away bias, *Modern Language Quarterly* 81.1.

на изображениях⁴⁶⁴. Результаты подобной работы важны, поскольку они могут способствовать сравнительно-историческому анализу произведений искусства и культуры и развития искусствоведения, исторической науки и других гуманитарных дисциплин⁴⁶⁵ в ситуации, когда коллекции музеев и других культурных учреждений начинают восприниматься как данные⁴⁶⁶. Коммерческие компании, работающие с распознаванием тематики и объектов на изображениях различных стилей и жанров, разрабатывают подходы к кластеризации и поиску изображений с помощью машинного обучения⁴⁶⁷ и, используя данные агрегаторов, могут перенести накопленную тенденциозность на результаты своей работы. Существовавшее ранее отсутствие репрезентативности и сбалансированности, унаследованное от физических коллекций произведений искусства и культуры, может быть усилено в цифровых визуальных корпусах, что приведет к цифровому культурному колониализму, который будет унаследован стандартами и платформами машинного обучения, в связи с самой природой данных, на которых они обучаются. Поэтому выводы исследования касаются не только той инфраструктуры данных, которая доступна сегодня, но и тех технологий, которые будут пользоваться ими завтра.

⁴⁶⁴ Di Lenardo I., Seguin B., and Kaplan F. (2016). Visual patterns discovery in large databases of paintings. *Proceedings of Digital Humanities 2016*, Krakow. <https://dh2016.adho.org/abstracts/348>; Wevers M., Smits T. (2019). The Visual Digital Turn. Using Neural Networks to Study Historical Images. *Digital Scholarship in the Humanities*. URL: <https://academic.oup.com/dsh/advance-article/doi/10.1093/llc/fqy085/5296356>

⁴⁶⁵ Manovich L. (2015). Data Science and Digital Art History. *International Journal of Digital Art History*, 1. URL: http://www.dah-journal.org/issue_01.html

⁴⁶⁶ Allen L., Frost H., Padilla Th., Potvin S., Roke E.R., Varner S. (2019). The Collections as Data Framework: A Review from the Always Already Computational Project. URL: <https://tdl-ir.tdl.org/handle/2249.1/156364>

⁴⁶⁷ Lang, S. and Ommer, B. (2018). Attesting similarity: supporting the organization and study of art image collections with computer vision. *Digital Scholarship in the Humanities*, Volume 33, Issue 4, December 2018.

Преобладание современной живописи и фотографии в российских и французских коллекциях, опубликованных платформой GA&C, также является важным результатом. Предварительные результаты, основанные на анализе изображений из российских и французских коллекций, показывают акцент на феномене «колониальной культуры» и представлений об экзотических странах, которые могут соответствовать ожиданиям западной публики. В то время как предвзятость, возникающая в результате оцифровки исторических документов и данных литературоведческих исследований, обсуждалась в литературе⁴⁶⁸, результаты данного диссертационного исследования значительно масштабнее того, что было представлено ранее в литературе, и показывают важность этой проблемы для будущих проектов оцифровки и агрегаторов культурного контента.

Таким образом, результаты исследования, представленного в данном параграфе, показывают, что данные о менее заметных культурах по-прежнему искажены и невидимы, а информация о культурах, которым уделяется меньше внимания, недостаточна. В литературе отмечается необходимость исправить несбалансированное представление данных в социальных и гуманитарных науках⁴⁶⁹. Несбалансированное представление данных, о котором идет речь в исследовании, может привести к тому, что знания о культуре останутся в той же этнической/экономической/социальной группе, что приведет к снижению социальной мобильности, отсутствию доверия и увеличению числа конфликтов⁴⁷⁰. Политика открытого доступа и готовность учреждений культуры видеть изображения из своих коллекций на внешних платформах могли бы исправить эту ситуацию.

⁴⁶⁸ Bode, K. (2018). *A World of Fiction: Digital Collections and the Future of Literary History*. Ann Arbor, MI: University of Michigan Press; Bode K. (2020). Why you can't model away bias, *Modern Language Quarterly* 81.1; Zaagsma G. (2019). *Digital History and the Politics of Digitization*. Presented at Digital Humanities 2019 Conference, 9-12 July, 2019.

⁴⁶⁹ Milan S., Treré E. (2019). Big Data from the south(s): Beyond data universalism. *Television and New Media*, 20(4): 319–335.

⁴⁷⁰ Bertrand M., Kamenica E. (2018). *Coming apart? Cultural distances in the United States over time*. University of Chicago mimeo. URL: <https://www.nber.org/papers/w24771>

Агрегаторы должны опубликовать методологию отбора произведений искусства и культуры. Этому может помочь знание литературы в области компьютерного литературоведения и корпусной лингвистики, связанной с решением проблемы отсутствия репрезентативности и сбалансированности лингвистического или литературоведческого корпуса.

Было бы наивно думать, что в настоящее время мы можем избежать социальных, этнических и технологических проблем, ведущих к исключению ряда художественных произведений из циркуляции, а следовательно, цифрового культурного колониализма. Это может произойти из-за проблем, связанных с географическим распределением данных, изменением культурного канона или мотивов, направленных на разработку алгоритмов, связанных с коммерческими интересами и ведущих к усилению предрассудков. Однако, если порталы, публикующие изображения предметов в области культурного наследия, не могут представить контент так, чтобы культуры были одинаково и пропорционально представлены в цифровых агрегаторах, администрируемых коммерческими компаниями и государственными учреждениями, мы можем, по крайней мере, стремиться понять природу тенденциозности, которая приводит к пробелам в знании культур и развитии предубеждений⁴⁷¹. В этом случае агрегаторы могут показать ограничения, в рамках которых результаты анализа их контента могут оказаться более объективными⁴⁷². Если цифровой контент агрегатора культурного наследия дает необъективные результаты, возможно, вся оцифровка является постколониальным способом исключения культур из публикации, распространения и исследований? Возможно, влияние на состав коллекций оказывают все заинтересованные стороны – социальные структуры, агрегаторы и учреждения культуры – и это явление имеет более глубокие эпистемологические и социальные причины и следствия. Редакторы GA&C должны сделать процедуры отбора и демонстрации произведе-

⁴⁷¹ Bode K. (2020). Why you can't model away bias, *Modern Language Quarterly* 81.1.

⁴⁷² Egbert J. (2019). Corpus design and representativeness. In Sardinha T. B. and Pinto V. (eds.) *Multidimensional Analysis. Research Methods and Current Issues*. Bloomsbury Academic, pp. 27-42.

ний искусства и культуры прозрачными. Важно, каким образом GA&C будет использовать свои ресурсы для отбора и распространения цифрового контента, представленного в агрегаторе.

Во второй главе данного диссертационного исследования были определены культурные закономерности формирования цифровых коллекций и функционирования агрегаторов цифровых материалов в области культурного наследия. Это было сделано с помощью исследования представленности российских цифровых коллекций, формирующих современное культурное наследие, и эмпирического количественного анализа цифровых коллекций Музейного фонда РФ и цифрового агрегатора Google Arts and Culture. Продемонстрирован значительный рост представленности российских цифровых коллекций, опубликованных на сайте Государственного каталога МФ РФ к моменту написания настоящего диссертационного исследования, по сравнению с 2015 годом, когда российские цифровые коллекции были представлены почти исключительно с помощью сайтов музеев, отправляющих статистическую отчетность Министерству культуры РФ. Несмотря на это, диссертационное исследование демонстрирует, что цифровые коллекции в области культурного наследия РФ созданы с целью учета и инвентаризации культурного наследия, а не с целью распространения и создания нового знания. Это связано прежде всего с невозможностью повторного использования изображений, опубликованных на сайте Государственного каталога, с помощью открытых лицензий. Это значит, что произведения из коллекций российских музеев не могут участвовать в создании нового цифрового канона и быть использованы для анализа больших данных в гуманитарных исследованиях. Второй параграф второй главы также проводит эмпирическое количественное исследование цифрового агрегатора Google Arts and Culture, показывая отсутствие сбалансированности в составе коллекций агрегатора. При этом ряд стран и учреждений культуры представлены в значительно большей степени, чем остальные страны. Результаты диссертационного исследования позволяют сделать вывод о новом цифровом колониализме, уже описанном в научной литературе. Результаты, представленные во второй главе диссертационного исследования, могут объясняться огра-

ничениями, связанными с распространением данных или недостаточным количеством оцифрованного контента в странах и регионах с недостаточной представленностью в цифровых коллекциях и агрегаторах. Вторая глава диссертационного исследования рекомендует редакторам цифровых коллекций и агрегаторов публиковать подробные описания принципов, которые стоят за отбором опубликованных материалов, с тем чтобы научное сообщество могло четко осознать особенности каждого собрания агрегированного цифрового контента.

ЗАКЛЮЧЕНИЕ

В данном диссертационном исследовании, посвященном выявлению складывающихся культурных закономерностей в современных процессах формирования цифровых коллекций и агрегаторов в области культурного наследия, сделаны следующие выводы.

В первом параграфе первой главы проведен обзор концепций и систематизированы знания о критическом анализе цифрового наследия, визуальном повороте в культуре, цифровой культуре, ремедиации визуальной культуры, цифровых коллекциях в области культурного наследия, критическом анализе цифровых инфраструктур в области культурного наследия. Это позволило проанализировать процессы, формы и способы создания цифровых коллекций и агрегаторов и их место в социальной, политической и культурной среде. Анализ подходов к критическому анализу культурного наследия, концепций цифровой культуры и ремедиации визуальной культуры продемонстрировал, что способы и формы применения новых медиа для фиксации и учета аналоговых материалов, анализа больших данных в области культурного наследия и публикации цифровых материалов с помощью механизмов трансляции, используемых цифровыми коллекциями и агрегаторами оцифрованного контента в области культурного наследия, определяют результаты исследования культуры и распространения/фиксации нового цифрового канона. Таким образом, концептуальные и методологические подходы к развитию цифровых технологий в области культурного наследия, во-первых, определяют цифровой поворот для всех культурных процессов, а во-вторых, влияют на способы и формы создания коллекций цифровых изображений и функционирования агрегаторов визуального контента в области культурного наследия. Это явление наблюдается во всех областях функционирования цифровых коллекций и агрегаторов: от рутинных процессов оцифровки, где отбор контента влияет на восприятие культурных закономерностей, до аннотирования изображений, ис-

пользования цифровых коллекций и создания агрегаторов цифровых данных в области культурного наследия.

Второй параграф посвящен анализу способов и форм создания коллекций цифровых изображений и функционирования агрегаторов визуального контента в области культурного наследия. Показано, что одной из самых важных функций цифровых коллекций и агрегаторов в области культурного наследия является создание и трансляция нового знания. Выявлена важность критического анализа цифровых инфраструктур в области культурного наследия, показана роль цифровых инфраструктур при анализе социальных, культурных и политических реалий настоящего времени. Уточнены понятия репрезентативности и сбалансированности в применении к цифровым коллекциям визуальных данных в области культурного наследия. Второй параграф демонстрирует, что в случае анализа больших данных для исследований в гуманитарных науках случайная выборка не приведет к тому, что получившийся набор данных окажется более репрезентативным и сбалансированным с точки зрения отражения эпохи создания произведений, поскольку принципы подбора и организации материалов физических и цифровых коллекций представляют собой социальные и культурные особенности организации материалов, отражающие те эпохи, когда формировались коллекции. Второй параграф рекомендует включать в цифровые коллекции и агрегаторы подробное описание методологических принципов, которыми руководствовались составители коллекций.

Первый параграф второй главы посвящен исследованию представленности российских цифровых коллекций, формирующих современное культурное наследие. Параграф вводит количественные методы анализа представленности коллекций и рассматривает проект ENUMERATE как пример исследования пропорционального отношения оцифрованных аналоговых коллекций и изображений предметов культурного наследия, опубликованных в информационном пространстве. Параграф рассматривает использование методики проекта ENUMERATE для оценки представленности российских цифровых коллекций. Показано, что использование методики демонстрирует невысокий уровень цифрового контента

(около 18 % аналоговых коллекций) и крайне низкий уровень представленности цифровых коллекций на сайтах музеев России (около 1 %) в 2014 году. В первом параграфе второй главы сделан вывод о том, что недостаточная представленность российских цифровых коллекций могла привести к тому, что цифровой контент, демонстрирующий культурное наследие, представленное в российских учреждениях культуры, не сможет сделать вклад в расширение нового цифрового канона художественных произведений и не сможет стать известно пользователю. Параграф демонстрирует, что российские цифровые коллекции, представленные в Государственном каталоге Музейного фонда Российской Федерации, включают в 2020 году около 30 % аналоговых коллекций основной части Музейного фонда РФ. Однако закрытость российских цифровых коллекций и невозможность повторного использования цифровых материалов с помощью открытых лицензий приводят к тому, что Государственный каталог носит учетный характер и не способствует возможности использования цифровых данных для исследований в гуманитарных науках.

Второй параграф второй главы демонстрирует результаты эмпирического количественного исследования цифровых коллекций Музейного фонда РФ и цифрового агрегатора материалов культурного наследия Google Arts and Culture. В первой части параграфа показаны результаты анализа статистической отчетности 2 367 российских музеев. Продемонстрировано процентное соотношение предметов с цифровыми изображениями из основной части МФ РФ по отношению к аналоговым материалам (14 %) и процентное соотношение цифровых изображений, опубликованных на сайтах музеев (1,5 %) в 2015 году. Показано, что задачи создания цифровых коллекций в РФ сводятся к учетным и инвентаризационным целям и только в редких случаях имеют отношение к предоставлению доступа к коллекциям или расширению цифрового канона произведений культуры и искусства. Во второй части параграфа представлены результаты количественного анализа оцифрованного контента, опубликованного с помощью цифровой платформы Google Arts and Culture. Диссертационное исследование показывает значительный рост контента цифровой платформы с 2011 года по настоящее время.

мя. В то же время второй параграф отмечает значимое отсутствие сбалансированности в агрегаторе, когда произведения из ряда стран и учреждений преобладают в контенте со значительным перевесом. Более того, во втором параграфе представлены результаты подробного исследования разделов агрегатора, включающих цифровые коллекции российских и французских учреждений культуры. Показано, что коллекции столичных музеев, архивов и библиотек представлены в значительно большей степени, чем провинциальные коллекции. Диссертационное исследование отмечает преобладание искусства двадцатого века в российских и французских коллекциях агрегатора. Второй параграф демонстрирует, что идеи культурного колониализма, ранее представленные в литературе в применении к аналоговым коллекциям, могут быть применены к цифровым коллекциям. Эта часть диссертационного исследования также предлагает интерпретацию результатов исследования и вводит ряд объяснений найденным явлениям. Объяснения включают отсутствие политики открытости данных в учреждениях культуры тех стран, которые представили наименьшее количество оцифрованного контента, возможное отсутствие оцифрованного контента и постколониальные культурные практики отбора и интерпретации собраний произведений культурного наследия, свойственные большим городам в западной культуре. Диссертационное исследование предлагает рекомендации редакторам и кураторам цифровых коллекций и агрегаторов по разработке методологических руководств, сопровождающих цифровые коллекции и отмечающих ограничения, связанные с особенностями представления цифровых коллекций в области культурного наследия. Только в этом случае у научного сообщества появится возможность проводить исследования визуального контента и больших данных в гуманитарных науках и делать широкие обобщения с учетом особенностей представленных данных.

СПИСОК ЛИТЕРАТУРЫ

1. Асеев, Ю. А. Вступительная статья редакторов перевода / Ю. А. Асеев, Я. А. Шер // Р. Чинхолл. Музейная каталогизация и ЭВМ / Р. Чинхолл. – М.: Мир, 1983. – С. 5–17.
2. Браккер, Н. В. Российские учреждения культуры в проекте ATHENA / Н. В. Браккер // 12-ая ежегодная международная конференция EVA 2009 Москва. – Москва, 30 ноября – 2 декабря 2009.
3. Браккер, Н. В. Европейская цифровая библиотека Europeana и российские учреждения культуры / Н. В. Браккер, Л. А. Куйбышев // ИФЛА. – 2013. – №1.
4. Пиков, Н. О. Время разворачивать камни / Н. О. Пиков, И. Н. Рудов, А. А. Ковалев и др. // Виртуальная археология (с воздуха, на земле, под водой и в музее). – Издательство «Государственный Эрмитаж». – 2018. – С. 179–189.
5. Государственный каталог МФ РФ [Электронный ресурс]: сайт Министерства культуры Российской Федерации. – Режим доступа: <https://goskatalog.ru/portal/#/collections> (дата обращения: 07. 06. 2020).
6. Есалиева, С. Правовое сопровождение цифровых коллекций произведений искусства. Выпускная квалификационная работа / С. Есалиева. – Сибирский федеральный университет, Красноярск, 2017.
7. Корчагин, К. М. Поэзия XX века в поэтическом подкорпусе Национального корпуса русского языка: Проблема репрезентативности / К. М. Корчагин // Труды Института русского языка им. В. В. Виноградова. – 2015. – Т. 6. – С. 235–256.
8. Миклас, Х. Мультиспектральная съемка для цифровой реставрации древних рукописей: устройства, методы и практические аспекты / Х. Миклас, С. Бреннер, Р. Саблатниг // Историческая информатика. – 2017. – № 3. – С. 116–134.

9. Шер, Я. А. Первые шаги отдела музейной информатики в Эрмитаже (1975–1985 гг) / Я. А. Шер // Информационные технологии в музее. – № 2. – Государственный Эрмитаж, Санкт-Петербург, 2006.

10. A Convergence of Methodologies: Notes on Data-Intensive Humanities Research / N. Tahmasebi, N. Hagen, D. Brodén, M. Malm // In: Digital Humanities in the Nordic Countries 4th Conference. – Helsinki, 2019.

11. Kizhner, I. Accessing Russian Culture Online: The scope of digitization in museums across Russia / I. Kizhner, M. Terras, M. Romyantsev // Digital Scholarship in the Humanities. – 2019. – Vol. 34. – Issue 2 – P. 350–367.

12. Addleman-Frankel, K. The Experience of Elsewhere: Photography in the Travelogues of Pierre Trémaux / K. Addleman-Frankel // photographies. – 2018. – Vol. 11. – P. 31–56.

13. Freire, N. Aggregation of Linked Data in the Cultural Heritage Domain: A Case Study in the Europeana Network / N. Freire, R. Voorburg, R. Cornelissen et al. // Information. – 2019. – Vol. 10. – 252 p.

14. Albu O. B. Organizational transparency: conceptualizations, conditions, and consequences / O.B. Albu, M. Flyverbom // Business and Society. – 2016. – Vol. 58. – Issue 2. – P. 268–297.

15. American Alliance of Museums [Электронный ресурс]. – URL: <http://ww2.aam-us.org/resources/assessment-programs/accreditation/accredited-museums> (дата обращения: 06.07.2020).

16. Anderson, B. Imagined Communities: Reflections on the Origin and Spread of Nationalism / B. Anderson. – Verso Books, 2006.

17. Arnold, M. Culture and Anarchy / M. Arnold. – Cambridge: Cambridge University Press, 1969.

18. Ashley, S. First Nations on View: Canadian Museums and Hybrid Representations of Culture / S. Ashley // In: Hybrid Entities, Annual Graduate Conference Hosted by the York / Ryerson Programme in Communication and Culture. – Canada: Rogers Communication Centre, York University, 2005. – P. 31–40.

19. Atkins, S. *Corpus Design Criteria* / S. Atkins, J. Clear, N. Ostler // *Journal of Literary and Linguistic Computing*. – 1992. – Vol. 7. – Issue 1. – P. 1–16.
20. Aufderheide, P. *The impact of copyright permissions culture on the US visual arts community: the consequences of fear of fair use* / P. Aufderheide, T. Milosevic, B. Bello // *New Media and Society*. – 2016. – Vol. 8. – Issue 9. – P. 2012–27.
21. Babeu, A. *Rome wasn't digitized in a day: building a cyberinfrastructure for digital Classics* / A. Babeu // *Council on Library and Information Resources*. Washington DC. – 2011.
22. Baca, M. (ed.). *Introduction to metadata* / M. Baca // *The Getty Research Institute*, Los Angeles, 2008.
23. Beaulieu, J. *Orientalism's Interlocutors: Painting, Architecture, Photography* / J. Beaulieu, M. Roberts. – Durham: Duke University Press, 2002.
24. Behdad, A. *Photography's Orientalism: New Essays on Colonial Representation* / A. Behdad, L. Gartlan. – Los Angeles: The Getty Research Institute, 2013.
25. Bennett, T. *The Birth of the Museum: History, theory, politics* / T. Bennett. – London: Routledge. 1995.
26. Bennett, T. *Collecting, Ordering, Governing: Anthropology, Museums and Liberal Government* / T. Bennett, F. Cameron, N. Dias, B. Dibley, I. Jacknis, R. Harrison and C. McCarthy. – Durham: Duke University Press, 2017.
27. Bertrand, M. *Coming apart? Cultural distances in the United States over time* / M. Bertrand, E. Kamenica // *University of Chicago*, 2018.
28. Besser, H. *Visual access to visual images: The UC Berkeley image database project*. *Library Trends*. – 1990. – Vol. 38, Issue 4 – P. 787–798.
29. Besser, H. *The transformation of the museum and the way it's perceived* / H. Besser // *In The Wired Museum: Emerging Technology and Changing Paradigms*; K. Jones-Garmil, (ed.). – American Association of Museums: Washington, D.C., 1997. – P. 153–170.
30. Besser, H. *The changing role of photographic collections with the advent of digitization* / H. Besser // *In The Wired Museum: Emerging Technology and Changing*

Paradigms, K. Jones-Garmil, (ed.). – American Association of Museums: Washington, D.C., 1997. – P. 115–127.

31. Biber, D. Representativeness in Corpus Design / D. Biber // *Literary and Linguistic Computing*. – 1993. – Vol. 8. – Issue 4. – P. 243–257.

32. Biber, D. *Corpus Linguistics: Investigating Language Structure and Use* / D. Biber, S. Conrad, R. Reppen. – New York: Cambridge University Press, 1998.

33. Bielsten, S. *Permissions, A Survival Guide: Blunt Talk About Art as Intellectual Property* / S. Bielstein. – Chicago: University of Chicago Press, 2006.

34. Boast, R. Neocolonial collaboration: museum as contact zone revisited / R. Boast // *Mus Anthropol*. – 2011. – Vol. 34. – P. 56–70.

35. Bode, K. *Methods and Canons: An Interdisciplinary Excursion* / K. Bode, T. Murphy // *Advancing Digital Humanities: Research, Methods, Theories*. Eds. Paul Longley Arthur and Katherine Bode. – London: Palgrave Macmillan, 2014. – P. 175–93.

36. Bode, K. *A World of Fiction* // K. Bode. – en. Ann Arbor, MI: University of Michigan Press, 2018.

37. Bode, K. Why you can't model away bias / K. Bode // *Modern Language Quarterly*. – Volume 81, Issue 1, 2020.

38. Boock, M. *Organizing for Digitization: A Survey* / M. Boock, R. Vondracek // *portal: Libraries and the Academy*. – 2006. – Vol. 6. Issue 2. – P. 197–217.

39. Borgman, C. *Big Data, Little Data, No Data* / C. Borgman. – Cambridge, MA: MIT Press, 2015.

40. Boyle, J. *The Public Domain, Enclosing the Commons of the Mind* / J. Boyle. – New Haven, CT: Yale University Press, 2010.

41. Caliskan, A. Semantics derived automatically from language corpora contain human-like biases / A. Caliskan, J.J. Bryson, A. Narayanan // *Science*. – 2017. – Vol. 356. – Issue 6334. – P. 183–86.

42. Cameron, F. Digital Futures I: Museum collections, digital technologies, and the cultural construction of knowledge / F. Cameron // *Curator* – 2003. – Vol. 46. – P. 325–340.

43. Cameron, F. Theorising heritage collection digitisations in global computational infrastructures / F. Cameron // In Hannah Lewi, Wally Smith, Dirk vom Lehn, Steven Cooke (eds) *The Routledge International Handbook of New Digital Practices in Galleries, Libraries, Archives, Museums and Heritage Sites*, Routledge, 2019.

44. Castells, M. *The Rise of the Network Society, the Information Age: Economy, Society and Culture* / M. Castells. – Oxford: Oxford University Press, 1996.

45. Chabries, D. M., Imaging the past: recent applications of multispectral imaging technology to deciphering manuscripts / D. M. Chabries, D. M. Booras, G. H. Bearman // *Antiquity*. – 2003. – Vol. 77. – Issue 296. – P. 359–72.

46. Chenhall, R. G. *Museum cataloguing in the computer age* / R.G. Chenhall. – Nashville, TN: American Association for State and Local History, 1975.

47. Chenhall, R. The world of (almost) unique objects / R. Chenhall, D. Vance // In: Parry R. (ed.), *Museums in a Digital Age*. – London; New York: Routledge, 2010. – P. 39–48.

48. Chomsky, N. *Syntactic structures* / N. Chomsky. – The Hague: Mouton, 1957.

49. Clear, J. Corpus sampling / J. Clear // In Leitner G. (ed.) *New Directions in English Language Corpora*. Berlin: Mouton-de-Gruyter, 1992. – P. 21–31.

50. Junginger, P. Close-Up Cloud: Gaining A Sense Of Overview From Many Details / P. Junginger, D. Ostendorf, B.A. Vissirini et al. // Presented at Digital Humanities 2019 Conference, 9–12 July, 2019.

51. Concordia, C. Not just another portal, not just another digital library: A portrait of Europeana as an application program interface / C. Concordia, S. Gradmann, S. Siebinga // In: *International Federation of Library Associations and Institutions*. – 2010. – Vol. 36. – Issue 1. – P. 61–69.

52. Szekely, P.A. Connecting the smithsonian american art museum to the linked data cloud / P. A. Szekely, C. A. Knoblock, F. Yang et al. // In P. Cimiano, O. Corcho,

V. Presutti, L. Hollink, S. Rudolph (eds.), Proc. of ESWC 2013. – Springer, 2013. – P. 593–607.

53. Connor, S. Postmodernist Culture: An Introduction to Theories of the Contemporary / S. Connor. – Oxford: Blackwell, 1991.

54. Cordell, R. ‘Q i-jtb the Raven’: Taking Dirty OCR Seriously / R. Cordell // Book History. – 2017. – Vol. 20. – P. 188–225.

55. Couldry, N. Mediatization or mediation: alternative understandings of the emergent space of digital storytelling / N. Couldry // New Media & Society. – 2008. – Vol. 10. – Issue 3. – P. 373–391.

56. Couldry, N, Hepp, A. The Mediated Construction of Reality / N. Couldry, A. Hepp. – Cambridge: Polity Press, 2017.

57. Crane, G. Towards a cultural heritage digital library / G. Crane, C. Wulfman // In L. Decambre, G. Henry, & C.C. Marshall (Eds.), Proceedings of the 3th ACM/IEEE-CS Joint Conference on Digital Libraries. – New York, NY: ACM Press, 2003. – P.75–86.

58. Daly, L. Digital Book History [Электронный ресурс] / Daly, L., Whitney T. (eds). – URL: <https://digitalbookhistory.com/> (дата обращения: 06.07.2020).

59. De Haan, J. Museum statistics and cultural policy / J. de Haan // EGMUS conference: Oslo, 2010.

60. Deuze, M. Participation, Remediation, Bricolage: Considering Principal Components of a Digital Culture / M. Deuze // The Information Society. – Volume 22, Issue 2, 2006 – P. 63–75.

61. Deuze, M. Media life / M. Deuze // Media Culture and Society. – Volume 33, Issue 1, 2011. – P. 137–148.

62. Hedrick, B. Digitization and the future of natural history collections / B. Hedrick, M. Heberling, E. Meinecke et al. // BioScience. – 2020. – Vol. 70. – Issue 3. – P. 243–251.

63. Kizhner, I. Digital Cultural Colonialism: measuring bias in aggregated digitized content held in Google Arts and Culture / I. Kizhner, M. Terras, M. Rummyantsev // Digital Scholarship in the Humanities. – 2020 (в печати).

64. Digital Hermeneutics: Agora and the Online Understanding of Cultural Heritage / Ch. van den Akker, S. Legêne, M. Van Erp et al. // In: Proceedings of the 3rd International Conference on Web Science (WebSci 2011). – 2011. – P. 1–7.

65. Di Lenardo, I. Visual patterns discovery in large databases of paintings / I. Di Lenardo, B. Seguin, F. Kaplan // Presented at Digital Humanities 2016 Conference, 9–12 July, 2016.

66. Donadio, R. Revisiting the Canon Wars / R. Donadio // The New York Times, September 16, 2007.

67. Drucker, J. Is there a ‘digital’ art history? / J. Drucker // Visual Resources. – 2013. – Vol. 29. – Issue 1–2. – P. 5–13.

68. Drucker, J. Graphesis: Visual Forms of Knowledge Production / J. Drucker. – Cambridge, MA: Harvard University Press, 2014.

69. Dunst, A. The Graphic Narrative Corpus (GNC): Design, Annotation, and Analysis for the Digital Humanities / A. Dunst, R. Hartel, J. Laubrock // In: Proceedings of the 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR). – Kyoto, Japan, 9–15 November 2017. – P. 15–20.

70. Earhart, A.E. Can information be unfettered? Race and the new digital humanities canon / A. E. Earhart // In M. K. Gold, editor, Debates in the Digital Humanities. – University of Minnesota Press, Minneapolis, London, 2012. – P. 309–318.

71. Willems, G. Easy and cost-effective cuneiform digitizing / G. Willems, F. Verbiest, W. Moreau et al. // In Short and Project Papers Proceedings of 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST2005), Mudge M., Ryan N., Scopigno R., (Eds.), Eurographics Association, 2005. – P. 73-80.

72. Eder, M. Visualization in Stylometry: Cluster Analysis Using Networks / M. Eder // Digital Scholarship in the Humanities. – 2015. – Vol. 30.

73. Egbert, J. Corpus design and representativeness / J. Egbert // In Sardinha T. B. and Pinto V. (eds.) Multidimensional Analysis. Research Methods and Current Issues. Bloomsbury Academic, 2019. – P. 27–42.

74. Ehrlich, V. Russian Poets in Search of a Poetics / V. Ehrlich // Comparative Literature. – 1952. – Vol. 4. – Issue 1. – P. 54–74.

75. Ellis, D. A. Behavioural Model for Information Retrieval System Design / D. Ellis // *Journal of Information Science*. – 1989. – Vol. 15. – Issue 4–5. – P. 237–247.

76. Elkins, J. *The Domain of Images* / J. Elkins. – Ithaca, NY: Cornell University Press, 1999.

77. Daquino, M. Enhancing semantic expressivity in the cultural heritage domain: Exposing the Zeri Photo Archive as Linked Open Data / M. Daquino, F. Mambelli, S. Peroni et al. // *Comput. Cult. Herit.* – 2017. – Vol. 10. – P. 21–42.

78. Enriching and Publishing Cultural Heritage as Linked Open Data / N. Simou, A. Chortaras, G. Stamou, S. Kollias // In: Ioannides M., Magnenat-Thalmann N., Papagiannakis G. (eds) *Mixed Reality and Gamification for Cultural Heritage*. – Springer, Cham, 2017.

79. European Group on Museum Statistics [Электронный ресурс]. – URL: <https://www.egmus.eu/> (дата обращения: 06.07.2020).

80. Europeana – a search engine for digitized cultural heritage material / V. Petras, T. Hill, J. Stiller, and M. Gäde, *Datenbank Spektrum*. – 2017. – Vol. 17. – Issue 1. – P. 41- 46.

81. Evans, J. *Visual culture: The reader* / J. Evans, S. Hall. – London: Sage and Open University Press, 1999.

82. Examination of historical paintings by state-of-the-art hyperspectral imaging methods: from scanning infra-red spectroscopy to computed X-ray laminography / S. Legrand, F. Vanmeert, G.V.d. Snickt et al. // *Herit. Sci.* – 2014. – Vol. 2. – P. 13–24.

83. Exploring New Business Models for Monetising Digitisation Beyond Image Licensing to Promote Adoption of OpenGlam / F. Valeonti, M. Terras, A. Hudson-Smith, Ch. Zarkali // *EVA Berlin 2018 Conference*, 7–9 November 2018.

84. Feldman, M. H. Rethinking the Canon of Near Eastern Art in the Internet Age / M.H. Feldman // *Journal of Ancient Near Eastern History*. – 2017. – Vol. 3. – P. 57–79.

85. Fenlon, K. Modeling Digital Humanities Collections as Research Objects / K. Fenlon // *Proceedings of ACM Joint Conference on Digital Libraries*. – 2019.

86. Flanders, J. Data Modeling / J. Flanders, F. Jannidis // In: Susan Schreibman, Ray Siemens, and John Unsworth (eds) *A New Companion to Digital Humanities*. – New York: Wiley and Sons, 2016. – P. 229–37.

87. Flanders, J., Siemens, R. Considering the scholarly edition in the digital age: an engagement by the modern language association's committee on scholarly editions / J. Flanders, R. Siemens // *Int J Digit Humanities*. – 2019.

88. Flyverbom, M. Data structuring – Organizing and curating digital traces into action / M. Flyverbom, J. Murray // *Big Data and Society*. – 2018. – Vol. 5. – Issue 2. – P. 1–12.

89. Fraistat, N. The Canon, The Web, and the Digitization of Romanticism / N. Fraistat, S.E. Jones, C. Stahmer // *Romanticism On the Net*. – 1998. – Vol. 10.

90. From digital resources to historical scholarship with the British library 19th century newspaper collection / I. Gregory, P. Atkinson, A. Hardie et al. // *Journal of Siberian Federal University. Humanities and Social Sciences*. – 2016. – Vol. 9. – Issue 4. – P. 994–1006.

91. Fyfe, P. An Archaeology of Victorian Newspapers / P. Fyfe // *Victorian Periodicals Review*. – 2016. – Vol. 49. – Issue 4. – P. 546–77.

92. Gentry, K., Smith, L. Critical heritage studies and the legacies of the late-twentieth century heritage canon / K. Gentry, L. Smith // *International Journal of Heritage Studies*. – 2019. – P. 1–21.

93. Glinka, K. Visualising the «Un-seen»: Towards Critical Approaches and Strategies of Inclusion in Digital Cultural Heritage Interfaces / K. Glinka, S. Meier, M. Dörk // In: Busch, C, Sieck, J (eds.) *Kultur und Informatik*. – Berlin: Cross Media, Werner Hülsbusch. – 2015. – P. 105–18.

94. Gooding, P. The myth of the new: mass digitization, distant reading and the future of the book / P. Gooding, M. Terras, C. Warwick // *Literary and Linguistic Computing*. – 2013. – Vol. 28. – Issue 4. – P. 629–39.

95. Gold, M. K. *Debates in the Digital Humanities* / M. K. Gold (ed). – Minneapolis: University of Minnesota Press, 2012.

96. Govindan, V. K. Character recognition – a review / V. K. Govindan, A. P. Shivaprasad // *Pattern Recognition*. – 1990. – Vol. 23. – No. 7. – P. 671–683.

97. Gref, M. Improved transcription and indexing of oral history interviews for digital humanities research / M. Gref, J. Köhler, A. Leh // *International Conference on Language Resources and Evaluation (LREC)*, 2018.

98. Gregory, I. Historical GIS: structuring, mapping and analysing geographies of the past / I. Gregory, R. Healey // *Progress in Human Geography*. – 2007. – Vol. 31. – P. 638–53.

99. Gries, S.T. What is corpus linguistics? / S.T. Gries // *Language and Linguistics Compass*. – 2009. – Vol. 3. – Issue 5. – P. 1225–1241.

100. Grieve-Smith, A. The importance of sampling frames in representative historical corpora: a case study of Parisian theater / A. Grieve-Smith // *CogniTextes*. – 2019.

101. Hall, S. Memories of style: Digitizing the Yorkshire Fashion Archive / S. Hall // *Art Libraries Journal*. – 2017. – Vol. 42. – Issue 3. – P. 157–161.

102. Hansen, H. K. The politics of transparency and the calibration of knowledge in the digital age / H.K. Hansen, M. Flyverbom // *Organization*. – 2014. – Vol. 22. – P. 872–889.

103. Harrison, J. Museums as Agencies of Neocolonialism in a Postmodern World / J. Harrison // *Studies in Cults, Organizations and Societies*. – 1997. – Vol. 3. – Issue 1. – P. 41–65.

104. Harrison, R. *Understanding the Politics of Heritage* / R. Harrison. – Manchester: Manchester University Press, with Open University, 2010.

105. Harrison, R. *Heritage: Critical Approaches* / R. Harrison. – Abingdon and New York: Routledge,, 2013.

106. Harvey, D. *The Condition of Postmodernity: An Enquiry into the Origins of Cultural Change* / D. Harvey. – Cambridge: Blackwell, 1990.

107. Harvey, D. Heritage Pasts and Heritage Presents: Temporality, Meaning and the Scope of Heritage Studies / D. Harvey // *International Journal of Heritage Studies*. – Volume 7, Issue 4, 2001 – P. 319–338.

108. Hayles, N.K. *How We Think: Digital Media and Technogenesis* / N.K. Hayles. – Chicago, IL: University of Chicago Press, 2012.

109. Hayles, N. K. *Comparative Textual Media: Transforming the Humanities in the Postprint Era* / N.K. Hayles, J. Pressman (eds). – Minneapolis: University of Minnesota Press, 2013.

110. Hewison, R. *The Heritage Industry: Britain in a Climate of Decline* / R. Hewison. – London: Methuen, 1987.

111. Hindmarch, J. *On virtual auras. The cultural heritage object in the age of 3D digital reproduction* / J. Hindmarch, M. Terras, S. Robson // In Hannah Lewi, Wally Smith, Dirk vom Lehn, Steven Cooke (eds) *The Routledge International Handbook of New Digital Practices in Galleries, Libraries, Archives, Museums and Heritage Sites*, Routledge, 2020.

112. Hitchcock, T. *Confronting the digital or how academic history writing lost the plot* / T. Hitchcock // *Cultural and Social History*. – 2013. – Vol.10. – P. 9–23.

113. Hooper-Greenhill, E. *Museums and the Shaping of Knowledge* / E. Hooper-Greenhill. – Routledge, London, 1992.

114. Hughes, L. *Digitizing collections: strategic issues for the information manager* / L. Hughes. – London: Facet Publishing, 2004.

115. Huster, A. C. *Assessing Systematic Bias in Museum Collections: A Case Study of Spindle Whorls* / A. C. Huster // *Advances In Archaeological Practice*. – 2013. – Vol. 1. – Issue 2. – P. 77–90.

116. Hyvonen, E. *Using the Semantic Web in Digital Humanities: Shift from data publishing to data-analysis and serendipitous knowledge discovery* / E. Hyvonen // *Semantic Web – Interoperability, Usability, Applicability*. – 2020. – Vol. 11. – Issue 1. – P. 187–193.

117. IBM. *Hermitage museum project* [Электронный ресурс]. – URL: <https://www.research.ibm.com/haifa/projects/software/hermitage> (дата обращения: 07.07.2020).

118. Imbert, C. *Transfert(s): Western fashion meets the East* / C. Imbert // *Google Arts and Culture* [Электронный ресурс]. – URL:

<https://artsandculture.google.com/exhibit/EwISenSdmLJ8IQ> (дата обращения: 13.07.2020).

119. Jackson, C. Using social network analysis to reveal unseen relationship in medieval Scotland / C. Jackson // *Digital Scholarship in the Humanities*. – 2017. – Vol. 32, Issue 2 – P. 336-343.

120. Jarlbrink, J. Cultural Heritage as Digital Noise: Nineteenth Century Newspapers in the Digital Archive / J. Jarlbrink, S. Pelle // *Journal of Documentation*. – 2017. – Vol. 73. – Issue 6. – P. 1228–43.

121. Jockers, M.L. *Macroanalysis: Digital Methods & Literary History* / M.L. Jockers. – University of Illinois Press, Champaign, Illinois, 2013.

122. Joconde Database. Présentation [Электронный ресурс]. – URL: <http://www2.culture.gouv.fr/documentation/joconde/fr/apropos/presentation-joconde.htm> (дата обращения: 13.07.2020).

123. Jones, A. L. Exploding Canons: the Anthropology of Museums / A. L. Jones // *Annual Review of Anthropology*. – 1993. – Vol. 22. – Issue 1. – P. 201–20.

124. Jørgensen, C. *Image Retrieval: Theory and Research* / C. Jørgensen. – Lanham, MD: Scarecrow Press, 2003.

125. Karp, I. *Exhibiting Cultures: The Poetics and Politics of Museum Display* / I. Karp, S. D. Lavine. – Washington, DC: Smithsonian Institution Press, 1991.

126. Kapsalis, E. The Impact of Open Access on Galleries, Libraries, Museums, & Archives / E. Kapsalis // *Statewide Agricultural Land Use Baseline*. – 2015.

127. Kapsalis, E. 21st-Century Diffusion with Smithsonian Open Access / E. Kapsalis . – Smithsonian, 25 February 2020.

128. Kelly, K. *Images of Works of Art in Museum Collections: The Experience of Open Access* / K. Kelly. – Council on Library and Information Resources, 2013.

129. Kizhner, I. Museum digitization practices across Russia: Survey and web site exploration results / I. Kizhner, M. Terras, M. Rummyantsev // In: *Digital Humanities 2016: Conference Abstracts*. – Jagiellonian University and Pedagogical University, Kraków. – 2016. – P. 600–2.

130. Knell, S. J. *National Museums: New Studies from Around the World* / S. J. Knell, P. Aronsson, A. B. Amundsen. – London: Routledge, 2010.
131. Knell, S. Preface / S. Knell // In S. J. Knell, P. Aronsson & A. B. Amundsen (eds), *National Museums: New Studies from Around the World*. – London: Routledge, 2010.
132. Koller, D. Research challenges for digital archives of 3D cultural heritage models / D. Koller, B. Frischer, and G. Humphreys // *ACM Journal on Computing and Cultural Heritage*. – 2009. – Volume 2. – Issue 3.
133. Knappett, C. *Network analysis in archaeology: New approaches to regional interaction* / C. Knappett. – Oxford: Oxford University Press, 2013.
134. Lang, S. Attesting similarity: supporting the organization and study of art image collections with computer vision / S. Lang, B. Ommer // *Digital Scholarship in the Humanities*. – 2018. – Vol. 33. – Issue 4.
135. Leech, G. New resources, or just better old ones? The Holy Grail of representativeness / G. Leech // In: Hundt, M., Nesselhauf, N., Biewer, C. (eds.), *Corpus Linguistics and the Web*. – Rodopi, Amsterdam, 2007. – P. 133-150.
136. Lessig, L. *Remix: Making Art and Commerce Thrive in the Hybrid Economy* / L. Lessig. – New York: Penguin Press, 2008.
137. Liang, H. Advances in multispectral and hyperspectral imaging for archaeology and art conservation / H. Liang // *Appl Phys A Mater Sci Process*. – 2012. – Vol. 106. – P. 309–323.
138. Licensing images from Russian museums for an academic project / I. Kizhner, J. Stankevich, M. Rummyantsev, I. Makarchuk // *Journal of Siberian Federal University. Humanities and Social Sciences*. – 2016. – Vol. 7. – Issue 9. – P. 1618–30.
139. Light, R. B. *Museum Documentation Systems: developments and applications* / R. B. Light, D. A. Roberts, J. D. Stewart. – London, 1986.
140. Limb P. The politics of digital reform and revolution: towards mainstreaming and African control of African digitization / P. Limb // *Innovation*. – 2007. – Vol. 34.

141. Linked Open Data for cultural heritage: Evolution of an information technology / J. Marden, C. Li-Madeo, N. Whysel, J. Edelstein // In Proceedings of the 31st ACM International Conference on Design of Communication. ACM, New York, NY, 2013. – P. 107–112.

142. Liu, A. Toward Critical Infrastructure Studies: Digital Humanities, New Media Studies, and the Culture of Infrastructure / A. Liu. – University of Connecticut, 23 February, 2017.

143. Liu, A. Toward a Diversity Stack: Digital Humanities and Diversity as Technical Problem / A. Liu // PMLA. – 2020. – Vol. 135. – Issue 1. – P. 130–151.

144. Lopatovska, I. Exploring requirements for online art collections / I. Lopatovska, I. Bierlein, H. Lember, E. Meyer // Proceedings of the American Society for Information Science and Technology. – 2013. – Vol. 50. – Issue 1. – P.1–4.

145. Loukissas, Y. A. All Data Are Local: Thinking Critically in a Data-Driven Society / Y.A. Loukissas. – Cambridge, Mass.: The MIT Press, 2019.

146. Lowenthal, D. The Past is a Foreign Country / D. Lowenthal. – Cambridge: CUP, 1985.

147. Lumley, R. The Museum Time-machine: Putting Cultures on Display / R. Lumley. – London: Routledge, 1988.

148. Lynch, C. A. Digital collections, digital libraries and the digitization of cultural heritage information / C.A. Lynch // First Monday – Volume 7, Issue 5, 2002.

149. Mak, B. Archaeology of a Digitization / B. Mak // Journal of the Association for Information Science and Technology. – Volume 65, Issue 8, 2014. – P. 1515–26.

150. Manoff, M. Theories of the Archive from Across the Disciplines / M. Manoff // portal: Libraries and the Academy. – 2004. – Vol. 4. – Issue 1. – P. 9–25.

151. Manovich, L. The Language of New Media / L. Manovich. – Cambridge, MA: MIT Press, 2001.

152. Manovich, L. Software takes command / L. Manovich. – New York and London: Bloomsbury Academic, 2013.

153. Manovich, L. Data science and digital art history / L. Manovich // International Journal for Digital Art History. – 2015. – Vol. 1. – P. 12–35.

154. Manovich, L. Cultural analytics, social computing and digital humanities. / L. Manovich // In Schafer M. T., van Es K. (eds), *The Datafied Society: Studying Culture Through Data*. Amsterdam: Amsterdam University Press, 2016. – P. 55–69.

155. Mansfield, E. C. Google Art Project and Digital Scholarship in the Visual Arts / E. C. Mansfield // *Visual Resources: An International Journal of Documentation*. – 2014. – Vol. 30. – Issue 1. – P. 110-117.

156. Martin R. *Orientalism: Visions of the East in Western Dress* / R. Martin, H. Koda. – New York: Metropolitan Museum of Art, 1994.

157. Martinez, M. ‘Not adopted’: The UK orphan works licensing scheme and how the crisis of copyright in the cultural heritage sector restricts access to digital content / M. Martinez, M. Terras // *Open Library of Humanities*. – 2019. – Vol. 5. – Issue 1 – P. 1–51.

158. Marty, P. F. *Museum Informatics and Information Infrastructures: Supporting Collaboration across Intra-Museum Boundaries* / P. F. Marty // *Archives and Museum Informatics*. – 1999. – Vol. 13. – P. 169–85.

159. McCarthy, D. *Open Access Scope in Open GLAM* / D. McCarthy // *Medium*, 2 May 2019.

160. McCarty, W. *Humanities Computing* / W. McCarty – London: Palgrave Macmillan, 2005.

161. McGann, J. *A New Republic of Letters: Memory and Scholarship in the Age of Digital Reproduction* / J. McGann. – Chicago: University of Chicago Press, 2014.

162. *Metadata, Digital Infrastructure, and the Data Ideologies of Cultural Anthropology* / L. Poirier, K. Fortun, B. Costelloe-Kuehn, M. Fortun // In: Crowder J., Fortun M., Besara R., Poirier L. (eds) *Anthropological Data in the Digital Age*. Palgrave Macmillan, Cham, 2020.

163. Milan, S. *Big Data from the south(s): Beyond data universalism* / S. Milan, E. Treré // *Television and New Media*. – 2019. – Vol. 20. – Issue 4. – P. 319–335.

164. Mirzoeff, N. *What is visual culture?* / N. Mirzoeff // In N. Mirzoeff (Ed.). *The Visual Culture Reader*. – New York: Routledge, 1998. – P. 1-13.

165. Morton, H.C. *The Story of Webster's Third: Philip Gove's Controversial Dictionary and Its Critics* / H.C. Morton // Cambridge: Cambridge University Press, 1994.
166. Müller, M. *Fundamentals of Music Processing* / M. Müller. – Springer International Publishing: Cham, Switzerland, 2015.
167. Nauta, J. G. *Survey report on digitization in European Cultural Institutions 2015* / J. G. Nauta, W. van den Heuvel. – Europeana Pro, 2015.
168. Navarette, T. *A History of Digitization: Dutch Museums* / T. Navarette. – University of Amsterdam, 2014.
169. *New reflection transformation imaging methods for rock art and multiple-viewpoint display* / M. Mudge, T. Malzbender, C. Schroer, M. Lum,. In: Ioannides, M., Arnold, D., Niccolucci, F. (Eds.), *Proceedings of the 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST2006)*. Eurographics Association, 2006. – P. 195–200.
170. *No specimen left behind: industrial scale digitization of natural history collections* / V. Blagoderov, I. Kitching, L. Livermore et al. // *ZooKeys*. – 2012. – Vol. 209. – P. 133–146.
171. Noble, S.U. *Algorithms of oppression: How search engines reinforce racism* / S.U. Noble. – New York: NYU Press, 2018.
172. Nora, P. *Between Memory and History: Les Lieux De Mémoire*. / P. Nora // *Representations*. – 1989. – Volume 26. – P. 7–25.
173. Nora, P. *Reasons for the Current Upsurge in Memory* / P. Nora // *Transit* – 2002. – Volume 22. – Issue 1. – P. 4–8.
174. *On Close and Distant Reading in Digital Humanities: A Survey and Future Challenges* / S. Janicke, G. Franzini, M. Cheema, G. Scheuermann // *In Proc. of EuroVis–STARs*, 2015. – P. 83–103.
175. Oomen, J. *Crowdsourcing in the cultural heritage domain: opportunities and challenges* / J. Oomen, L. Aroyo // *Communities and Technologies*. – 2011. – P. 138–149.

176. Ortega, É. Archives, Libraries, Collections, and Databases: A First Look at Digital Literary Studies in Mexico / É. Ortega // *Hispanic Review*. – 2018. – Vol. 86. – Issue 2. – P. 229–47.
177. Osborne, P. D. Travelling Light: Photography, Travel and Visual Culture / P.D. Osborne. – Manchester: Manchester UP, 2000.
178. Pachali, D. Open Access: How museums are opening their digital archives / D. Pachali.– Goethe-Institut, 2014.
179. Pagenstecher, C. Curating and Analyzing Oral History Collections / C. Pagenstecher // Selected papers from the CLARIN Annual Conference. Linköping Electronic Conference Proceedings, 2019. – P. 144–151.
180. Parry, R. Recoding the Museum: Digital Heritage and the Technologies of Change / R. Parry. – Routledge, 2007.
181. Parry, R. Museums in a Digital Age, Leicester Readers in Museum Studies / R. Parry. – Routledge, 2007.
182. Paterson, L.L. Geographical Information Systems and Textual Sources / L.L. Paterson, I.N. Gregory // In: *Representations of Poverty and Place*. Palgrave Macmillan, Cham, 2019.
183. Peacock, D. The information revolution in museums / D. Peacock // In: Marty, P.F. and Burton Jones, K. (eds) *Museum informatics: People, information, and technology in museums*. – London and New York: Routledge, 2007. – P. 59–76.
184. Petras, V. Decade of Evaluating Europeana – Constructs, Contexts, Methods & Criteria / V. Petras, J. Stiller // In *Research and Advanced Technology for Digital Libraries*, ed. by J. Kamps, G. Tsakonas, Y. Manolopoulos et al. – Berlin: Springer, 2017. – P. 233–245.
185. Petri, G. The Public Domain vs. the Museum: The Limits of Copyright and Reproductions of Two-Dimensional Works of Art / G. Petri // *Journal of Conservation and Museum Studies*. – 2014. – Vol. 12. – Issue 1.
186. Pfisterer, U. Big Bang Art History / U. Pfisterer // *Digital Art History*. – 2018. – Issue 3.

187. PHAROS: A digital research space for photo archives / C. Caraffa, E. Pugh, T. Stuber, L. Ruby // *Art Libraries Journal*. – 2020. – Vol. 45. – Issue 1. – P. 2-11.
188. Piper, A. *Enumerations: Data and Literary Study* / A. Piper. – Chicago: Chicago University Press, 2018.
189. Poll, R. NUMERIC: statistics for the digitization of European cultural heritage, *Electronic Library and Information Systems*. – 2010. – Vol. 44. – Issue 2. – P. 122–31.
190. Polm, M. Museum representations of Roman Britain and Roman London: a postcolonial perspective / M. Polm // *Britannia*. – 2016. – Vol. 47. – P. 209-241.
191. Price, K. M. Digital Scholarship, Economics, and the American Literary Canon / K. M. Price // *Literature Compass*. – 2009. – Vol. 6. – Issue 2. – P. 274–90.
192. Raineri, S. Corpora and Representativeness: Where to go from now? / S. Raineri, C. Debras // *CogniTextes*. – 2019. – Vol. 2019.
193. Ridge, M. (ed.). *Crowdsourcing our cultural heritage* / M. Ridge. – London: Ashgate Publishing, Ltd, 2014.
194. Risam, R. *New Digital Worlds: Postcolonial Digital Humanities in Theory, Praxis, and Pedagogy* / R. Risam. – Evanston, Illinois: Northwestern University Press, 2019.
195. Robinson, P. Towards a Theory of Digital Editions / P. Robinson // *Variants: The Journal of the European Society for Textual Scholarship*. – 2013. – Vol. 10. – P. 105–31.
196. Robinson, P. Five desiderata for scholarly editions in digital form / P. Robinson // *In Digital Humanities 2013: Conference Abstracts*. University of Nebraska-Lincoln. – 2013.
197. Rogers, N. Museum drawers go digital / N. Rogers // *Science*. – 2016. – Vol. 352. – P. 762–765.
198. Ryan J.R. *Picturing Empire: Photography and the Visualization of the British Empire* / J.R. Ryan. – Chicago: University of Chicago Press, 1997.
199. Said, E. W. *Culture and Imperialism* / E.W. Said. – New York: Random House, 1993.

200. Said E.W. Invention, memory, and place / E.W. Said // *Critical Inquiry*. – 2000. – Vol. 26. – P. 175–192.
201. Salamon-Cindori B. Digitization: challenges for Croatian museums / B. Salamon-Cindori, M. Tot, D. Zivkovic // *Qualitative and Quantitative Methods in Libraries*. – 2014. – Vol.1. – P. 213–220.
202. Santos, B. *The End of the Cognitive Empire. The Coming of Age of Epistemologies of the South* / B. Santos. – Durham: Duke University Press, 2018.
203. Saussy, H. Exquisite Cadavers Stitched from Fresh Nightmares: Of Memes, Hives, and Selfish Genes. / H. Saussy // In H. Saussy (ed.) *Comparative Literature in an Age of Globalization*. – Baltimore: Johns Hopkins University Press, 2004.
204. Schreibman, S. *A companion to Digital Humanities*. S. Schreibman, R. Siemens, J. Unsworth (eds). – Blackwell, 2004.
205. Schreibman, S. Textuality in 3D. Three-Dimensional (re)constructions as digital scholarly editions / S. Schreibman, C. Papadopoulos // *International Journal of Digital Humanities*. – 2019.
206. Schudson, M. *The Rise of the Right to Know* / M. Schudson. – Cambridge, MA: Belknap Publishing, 2015.
207. Schwartz, J.M. Negotiating the visual turn: new perspectives on images and archives / J.M. Schwartz // *Am Arch*. – 2004. – Vol. 67. – P. 107–122.
208. Sher, J. Museum documentation systems and computers: USSR experience. / J. Sher // In: Light, R.B., Roberts, D.A., and Stewart, J.D. (eds) *Museum Documentation Systems. Developments and Applications*. – Butterworths: London, Boston, 1986. – P. 287–292.
209. Sinn, D. Impact of digital archival collections on historical research / D. Sinn // *Journal of the American Society for Information Science & Technology*. – Vol. 63. – Issue 8. – 2012. – P. 1521–1537.
210. Smith, L. *The Uses of Heritage* / L. Smith. – London: Routledge, 2006.
211. Smith, L. *Intangible Heritage* / L. Smith and N. Akagawa. – London: Routledge, 2009.

212. Smithsonian team Flickr: A library, archives, and museums collaboration in Web 2.0 space / M. R. Kalfatovic, E. Kapsalis, K. P. Spiess et al. // *Archival Science*. – 2008. – Vol. 8. – Issue 4. – P. 267–277.

213. Sood, A. Every piece of art you've ever wanted to see - up, close and searchable / A. Sood. – TED lecture, 2016.

214. Spontaneous speech corpus of Japanese / K. Maekawa, H. Koiso, S. Furui, H. Ishara // *Proceedings of the Second International Conference of Language Resources and Evaluation (LREC)*, 2000. – P. 947-952.

215. Stiller, J. Learning from digital library evaluations: The Europeana case / J. Stiller, V. Petras // *ABI Technik*. – 2018. – Vol. 38. – Issue 1. – P. 37–45.

216. Supporting linked data production for cultural heritage institutes: The amsterdam museum case study / V. de Boer, J. Wielemaker, J. van Gent et al. // *In The Semantic Web: Research and Applications*. – 2012. – P. 733–747.

217. Tallon, L. Sparking Global Connections to Art Through Open Data and Artificial Intelligence / L. Tallon // *Blogs/Now at the Met*. – The Metropolitan Museum of Art, 4 February 2019.

218. Tanner S. Digitising the Dead Sea Scrolls: Archiving 2009 / S. Tanner, G. Bearman. – Arlington, VA: The Society for Imaging Science and Technology. – 2009. – P. 119–23.

219. Taylor J. Digitization, digital interaction and social media: embedded barriers to democratic heritage / J. Taylor, L.K. Gibson // *International Journal of Heritage Studies*. – Vol. 23. – Issue 5. – P. 1–13.

220. Terras, M. Reading the Readers: Modelling Complex Humanities Processes to Build Cognitive Systems / M. Terras // *Literary and Linguistic Computing*. – 2005. – Vol. 20, Issue 1. – P. 41-59.

221. Terras, M. Image to Interpretation: Intelligent Systems to Aid Historians in the Reading of the Vindolanda Texts. *Oxford Studies in Ancient Documents* / M. Terras. – Oxford University Press, Oxford, 2006.

222. Terras, M. Digital Images for the Information Professional / M. Terras. – Ashgate, Aldershot, England, 2008.

223. Terras, M. The rise of digitization / M. Terras // In: Rikowski R. (ed.), *Digitization Perspectives*. – Rotterdam, Boston, Taipei: Sense Publishers, 2011. – P. 3–20.

224. Terras, M. Opening Access to collections: the making and using of open digitised cultural content / M. Terras // *Online Information Review*. – 2015. – Vol. 39. – Issue 5. – P. 733 – 752.

225. Terras, M. Crowdsourcing in the Digital Humanities. / M. Terras // In: S. Schreibman, R. Siemens, and J. Unsworth (eds). *A New Companion to Digital Humanities*. – Oxford: Blackwell, 2016. – P. 420–438.

226. The Art Newspaper [Электронный ресурс]. How Google became a major producer of cultural content. – 19 January, 2018. – URL: <https://www.theartnewspaper.com/feature/how-google-became-a-major-producer-of-cultural-content> (дата обращения: 06.07.2020).

227. The Collections as Data Framework: A Review from the Always Already Computational Project / L. Allen, H. Frost, Th. Padilla et al. – *Collections as Data*, 2019.

228. The Collector of Success: Leon Bakst / I. Shumanova, Z. Shergina, T. Nikolaeva et al. // *Google Arts and Culture* [Электронный ресурс]. – URL: <https://artsandculture.google.com/exhibit/KgICSXEITyGILA> (дата обращения: 13.07.2020).

229. The extended specimen network: a strategy to enhance US biodiversity collections, promote research and education / J. Lendemer, B. Thiers, A.K. Monfils // *BioScience* – 2020. – Vol. 70, Issue 1. – P. 23–30.

230. The Europeana Data Model (EDM) / M. Doerr, S. Gradmann, S. Hennicke et al. // Paper presented at the World Library and Information Congress: 76th IFLA General Conference and Assembly. – Gothenburg, Sweden, 2012.

231. The European Digital Library-Factor for Long-life learning in Arts and Cultural Studies / I. Yankova, E. Velez, K. Nusheva et al. // *Qualitative & Quantitative Methods in Libraries*. – 2015. – Vol. 4. – Issue 4. – P. 965–971.

232. The FAIR Guiding Principles for Scientific Data Management and Stewardship / M. D. Wilkinson, M. Dumontier, I. Aalbersberg, et al. *Scientific Data*. – 2016. – Vol. 3. – P. 1–9.

233. The Rijksmuseum Collection as Linked Data // C. Dijkshoorn, L. Jongma, L. Aroyo // *Semantic Web*. – 2018. – Vol. 9. – Issue 2. – P. 221–230.

234. The value of critical destruction: Evaluating multispectral image processing methods for the analysis of primary historical texts / A. Giacometti, A. Campagnolo, L. MacDonald et. al. // *Journal of Digital Scholarship in the Humanities*. – 2017. – Vol. 32. – Issue 1. – P. 101–122.

235. Thiers, B.M. Digitization of the new york botanical garden herbarium / B. M. Thiers, M.C. Tulig,, K.A. Watson // *Brittonia*. – 2016. – Vol. 68. – Issue 3. – P. 324–33.

236. Thylstrup, N.B. The politics of mass digitization / N.B. Thylstrup. – The MIT Press, Cambridge 2018.

237. Towards a digital infrastructure for illustrated handwritten archives / A. Weber, M. Ameryan, K. Wolstencorft et al. // In: Loannides, M. (ed.) *Digital Cultural Heritage. Information Systems and Applications, incl. Internet/Web, and HCI*. – Vol. 10605. – Springer International Publishing, 2018. – P.155–166.

238. Towards a three-dimensional cost-effective registration of the archaeological heritage / J. de Reu, G. Plets, G.Verhoeven // *J. Archaeol. Sci.* – 2013. – Vol. 40. – Issue 2. – P. 1108-1121.

239. Trumpener, K. Critical Response I. Paratext and Genre System: A Response to Franco Moretti / K. Trumpener // *Critical Inquiry*. – 2009. – Vol. 36. – Issue 1. – P. 159–71.

240. Transforming scholarship in the archives through handwritten text recognition: Transkribus as a case study / G. Muehlberger, L. Seaward, M. Terras // *Journal of Documentation*. – 2019. – Vol. 75, No. 5. – P. 954–976.

241. Turner, B. The Erosion of Citizenship / B. Turner // *British Journal of Sociology*. – 2001. – Vol. 52. – Issue 2. – P. 189–209.

242. Tylor, E. B. *Primitive culture* / E.B. Tylor. – New York: Harper and Row, 1958.

243. Ul-Hasan, A. OCRoRACT: a sequence learning OCR system trained on isolated characters / A. Ul-Hasan, S.S. Bukhari, A. Dengel // 12th IAPR Workshop on Document Analysis Systems (DAS), IEEE, 2016. – P. 174–79.

244. Underwood, T. *Distant Horizons: Digital Evidence and Literary Change* / T. Underwood. – University of Chicago Press, 2019.

245. Van Dijck, J. *The Platform Society* / J. Van Dijck, T. Poell, M. de Waal. – New York: Oxford University Press. 2018.

246. Villaespesa, E. *Museum Collections and Online Users: Development of a Segmentation Model for the Metropolitan Museum of Art* / E. Villaespesa // *Visitor Studies*. – 2019. – Vol. 22. – Issue 2. – P. 233–252.

247. Valeonti, F, Terras, M, Hudson-Smith, A. How open is OpenGLAM? Identifying barriers to commercial and non-commercial reuse of digitised art images / F. Valeonti, M. Terras, A. Hudson-Smith // *Journal of Documentation*. – 2020. – Vol. 76 . – Issue 1. – P. 1–26.

248. Valtysson, B. *Europeana – The digital construction of Europe's collective memory* / B. Valtysson // *Information, Communication & Society*. – 2012. – Vol. 15. – Issue 2. – P. 151–170.

249. Van Dijck, J. *The Platform Society* / J. Van Dijck, T. Poell and M. de Waal. – New York: Oxford University Press, 2018.

250. Van Riel, C. *Why people love art museums: a reputation study about the 18 most famous museums among visitors in 10 countries* / C. Van Riel, P. Heijndijk. – Rotterdam School of Management, Erasmus University, 2017.

251. Vergo, P. 1989. *The New Museology* / P. Vergo. – London: Reaktion Books, 1989.

252. *Virtual reality for cultural heritage monuments – from 3D data recording to immersive visualization* / T. P. Kersten, F. Tschirschwitz, S. Deggim, M. Lindstaedt // *In Euro-Mediterranean Conference*. – Springer, Cham, 2018. – P. 74–83.

253. Von Haller Groesbaek, M. GLAMorous remix: Openess and sharing for cultural institutions / M. von Haller Groesbaek // Sanderhoff, M. (ed.) Sharing is caring: Openness and sharing in the cultural heritage sector. – 2014. – P. 141–153.

254. Wallace, A., Deazley, R. Display at your own risk: an experimental exhibition of digital cultural heritage [Электронный ресурс]. – URL: <http://displayatyourownrisk.org/publications/> (дата обращения: 08.07.2020).

255. Walsh, D. Characterising online museum users: a study of the national museums liverpool museum website / D. Walsh, M.M. Hall, P. Clough, J. Foster. //International Journal on Digital Libraries. – 2020. – Vol. 21. – P. 75–87.

256. Wang, M.C. Digitization and the Depoliticization of Museum Access. The Situation at the National Palace Museum / M.C. Wang // Advances in Computer Science: an International Journal. – 2016. – Vol. 5. – Issue 3. – No.21.

257. Warren, C. N. Historiography's Two Voices: Data Infrastructure and History at Scale in the Oxford Dictionary of National Biography (ODNB) / C.N. Warren //Journal of Cultural Analytics. – 2018.

258. Warwick, C. Digital Humanities in Practice / C. Warwick, M. Terras, J. Nyhan. – London: Facet Publishing in association with UCL Digital Humanities Centre, 2012.

259. Westerby, G. Digital Art History and the Museum: The Online Scholarly Collection Catalogues at the Art Institute of Chicago / G. Westerby, K. Keegan // Visual Resources. – 2019. – P. 1–14.

260. Wevers, M. The visual digital turn: Using neural networks to study historical images / M. Wevers, T. Smits // Digital Scholarship in the Humanities. – 2020. –Vol. 35. – Issue 1. – P. 194–207.

261. Willems, G. Easy and cost-effective cuneiform digitizing / G. Willems, F. Verbiest, W. Moreau et al. In Short and Project Papers Proceedings of 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST2005), Mudge M., Ryan N., Scopigno R., (Eds.). – Eurographics Association. – 2005. – P. 73–80.

262. Williams, D. A brief history of museum computerization / D. Williams // In: Parry R. (ed.), *Museums in a Digital Age*. – London, New York: Routledge, 2010. – P. 15–22.

263. Whalen M. What's wrong with this picture? An examination of art historians' attitudes about electronic publishing opportunities and the consequences of their continuing love affair with print / M. Whalen // *Art Documentation: Bulletin of the Art Libraries Society of North America*. – 2009. – Vol. 28. – Issue 2. – P. 13–22.

264. Whitelaw, M. Generous interfaces for digital cultural collections / M. Whitelaw // *Digital Humanities Quarterly*. – 2015. – Vol. 9, Issue 1.

265. Woods, A. *Statistics in language studies* / A. Woods, P. Fletcher, A. Hughes. – Cambridge: Cambridge University Press, 1986.

266. Zaagsma, G. *Digital History and the Politics of Digitization* / G. Zaagsma // Presented at Digital Humanities 2019 Conference, 9-12 July, 2019. – Utrecht, 2019.

267. Zuboff S. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* / S. Zuboff. – New York: Profile Books, 2019.