

УДК 622.3:004

ПРИНЦИПЫ ПОСТРОЕНИЯ ЦИФРОВОЙ ПЛАТФОРМЫ ДЛЯ НАУЧНО-ТЕХНИЧЕСКОГО ЦЕНТРА

М. М. Хасанов, Р. М. Галеев, А. М. Маргарит, Ф. В. Краснов
*Научно-технический центр ООО «Газпром нефть», г. Санкт-Петербург,
ntc_odo@gazpromneft-ntc.ru*

Цифровые платформы признаны успешным инструментом для создания прорывных стратегий как в цифровых, так и в нецифровых отраслях. Авторы данного исследования делятся результатами разработки цифровой платформы для научно-технического центра в нефтегазовой отрасли. В исследовании сделан анализ известных подходов к созданию цифровых платформ, определены особенности научно-технической деятельности, показано, как была учтена специфика нефтегазовой отрасли в цифровой платформе «ЭРА:ГРАД».

Ключевые слова: цифровая трансформация, цифровые платформы, архитектура программного обеспечения, процесс разработки программного обеспечения, экосистема инноваций.

DIGITAL PLATFORM BUILDUP PRINCIPLES FOR SCIENCE AND TECHNOLOGY CENTER

M. M. Khasanov, R. M. Galeev, A. M. Margarit, F. V. Krasnov
*Gazprom Neft's Science and Technology Center, Saint Petersburg,
ntc_odo@gazpromneft-ntc.ru*

Digital platforms are recognized as a successful tool for the development of breakthrough strategies in both digital and non-digital industries. The authors of this study share the results of a digital platform project implemented at a science and technology center of the petroleum industry. The study analyzes known approaches to the establishment of digital platforms, identifies specific features related to scientific and research activities in the petroleum industry and shows how these features have been taken into account in the ERA:GRAD digital platform.

Keywords: digital transformation, digital platforms, software architecture, software development process, innovation ecosystem.

Введение. Цифровые платформы являются для компаний не просто модой, а отражением глубинных изменений в экономике. Появляются новые гигантские компании, а построенные «на века» компании исчезают. Цифровые платформы создают сетевой эффект, что приводит к изменениям в основных бизнес-процессах. Происходит создание новых источников ценностей, а не перераспределение существующих.

Общеизвестны примеры цифровых платформ в Интернет-отраслях – это Яндекс, МТС, Билайн, и др. Каждый октябрь InterBrand выпускает свой список самых известных компаний. И здесь есть европейские (Ikea и BMW), американские (Coca-Cola и Apple) и азиатские (Honda и Toyota) бренды. Если мы разделим эти фирмы по признаку наличия платформы, то 13 из 30 крупнейших корпораций являются платформенными компаниями. Примером компании, которая обладала 49 %-й долей на рынке смартфонов в 2009 г. и потеряла ее (2 % в 2015 г.), проигнорировав создание платформы, стала Канадская компания Research In Motion (бренд BlackBerry). Также показательно сравнение развития Apple и Microsoft в 80-е и 90-е годы прошлого века. Рыночная стоимость Microsoft, которая в это время активно развивала экосистему, росла по экспоненте, в то же время Apple не создавала свою платформу и росла линейно (рис. 1).



Рис. 1. Сравнение доли рынка Microsoft и Apple

Есть распространенное заблуждение, что цифровые платформы существуют только для высокотехнологичных фирм. Это не так. Например, General Electric (бытовая техника) уверенно движется в направлении платформ для интернета вещей и мониторинга удаленных устройств. Honda и BMW движутся в направлении цифровых платформ, Nike движется в направлении платформ, добавив датчики к обуви и мобильное приложение, чтобы вы могли сравнить себя с другими спортсменами. Daimler развивает экосистемы на своей Цифровой платформе, внедряя сервисы совместного использования поездок. Есть множество примеров цифровых платформ, создаваемых в сельском хозяйстве, энергетике и науке.

Подход к цифровизации НТЦ. Качественный скачок в структуре и динамике развития производительных сил обеспечивается деятельностью отраслевых научно-технических центров (далее – НТЦ). Количество НТЦ в энергетической отрасли растет из года в год, а по мере исчерпания запасов легкодобываемой нефти роль научной составляющей в ее добыче возрастает. Поэтому эффективность деятельности НТЦ является ключевой характеристикой, нуждающейся в современных методиках развития, таких как платформенность.

Традиционно НТЦ создавались по лекалам российских научно-проектных институтов, которые оценивали запасы месторождений нефти и газа, ставили их на государственный баланс и формировали проектные документы на разработку месторождений. В задачи таких институтов также входила разработка и внедрение новых технологий и материалов, но часто проявлялась их главная уязвимость – обособленность от бизнеса.

Изначально после получения разрешительных документов в Центральной комиссии по разработке месторождений НТЦ отходил в сторону, и в дело вступали производственники. Современные НТЦ представляют собой научно-проектную структуру, которая полностью интегрирована в производство.

Основные сквозные бизнес-процессы НТЦ отображены на рис. 2.

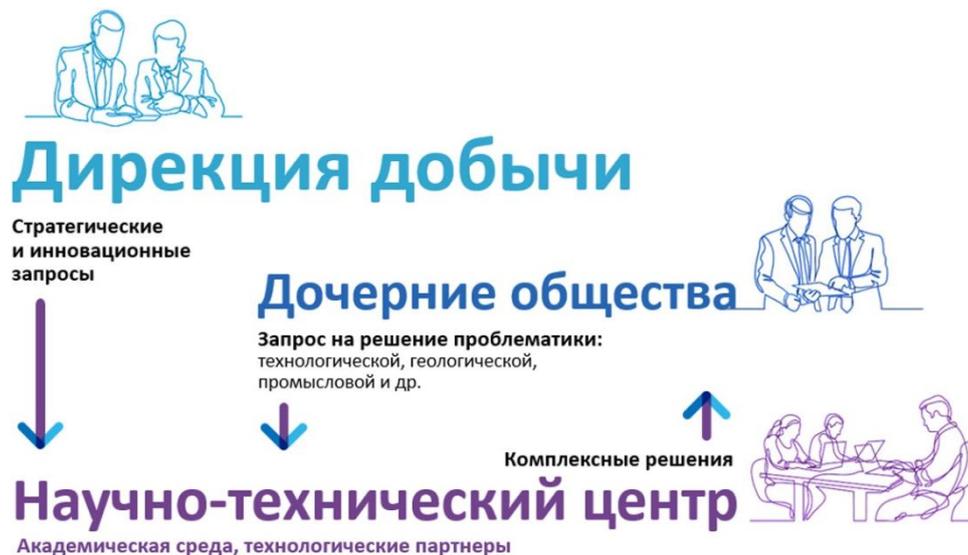


Рис. 2. Сквозные бизнес-процессы, связывающие НТЦ с Корпоративным Центром (КЦ) и Дочерними Обществами

Необходимо отметить, что рынок услуг НТЦ в РФ является высококонкурентным согласно исследованию компании Deloitte, поэтому от НТЦ требуется непрерывно повышать эффективность бизнес-процессов.

В книге [1] рассматривается создание в РФ рынка знаний, поэтому правомочно рассмотреть как частный случай рынок услуг НТЦ. Согласно экономической теории на этом рынке существуют две разные модели: «от спроса» и «от предложения». Поясним необходимость повышения эффективности бизнес-процессов НТЦ с использованием этих моделей. В случае варианта «от спроса», то есть при обращении к НТЦ для выполнения научно-исследовательских и опытно-конструкторских работ (далее – НИОКР), затраты НТЦ будут расти быстрее, а количество НИОКР, которые будет способен выполнить НТЦ, – уменьшаться. При работе «от предложения» стоимость НИОКР будет уменьшаться, а количество – увеличиваться. Оба варианта будут искать равновесный вариант и не возможны в абсолютной форме, а только в сочетании.

Цифровая платформа является одним из способов повышения эффективности бизнес-процессов [2], так как производит их глубинную трансформацию и позволяет найти новую точку равновесия на фазовой диаграмме «спрос-предложение».

Цифровая платформа «ЭРА:ГРАД». «Электронная разработка активов (ЭРА)» – это компонент цифровой трансформации «Газпром нефти» в сфере разведки и добычи, которая охватывает все основные направления деятельности: геологоразведку, геологию, обустройство месторождений, бурение, разработку, добычу. Ее реализация началась в 2012 году, а в 2014 году она была включена в Технологическую стратегию блока разведки и добычи «Газпром нефти» в качестве одного из ключевых направлений.

В основу платформенной философии «ЭРА:ГРАД» были положены три следующих принципа: «Автономность», «Магазин алгоритмов» и «Все – это таблица». Остановимся на них подробнее.

Автономность. Принцип автономности продиктован квази-статичным поведением источников данных во времени и поступательным характером работы с данными. В «ЭРА:ГРАД» задействовано более 30 различных источников данных, каждый из которых можно образно представлять в виде отдельной ИТ-системы со своими пользователями, регламентами и структурами данных. Подстраиваться под регламент обновления данных каждого источника даже для выполнения простого кросс-функционального расчета не представляется возможным, а в расчетах средней сложности бывают задействованы до 10 различных источников.

Для получения согласованной версии данных в «ЭРА:ГРАД» используется механизм локального кеширования в in-memoy базе данных. Следовательно агент (пользователь) платформы может работать со своей версией согласованных данных. Важно отметить, что некоторые данные имеют специальный режим доступа и не могут быть сохранены локально. Таким образом, в полученных при расчете результатах содержится метка кэша, на котором выполнены расчеты. Такой уровень прозрачности похож на один из элементов технологии Blockchain, когда в результате «защита» возможность проверить всю цепочку вычислений до источника [3].

Магазин алгоритмов. Вспоминая классический учебник тьюринговского лауреата Н. Вирта «Алгоритмы и структуры данных» [4], можно отметить, что алгоритмы живут своей отдельной жизнью от данных. Со времени написания этого учебника отношение к алгоритмам изменилось, наиболее точно оно выражено в теореме «no free lunch», доказанной Д. Волпертом в 1997 году [5]. Суть этой теоремы в сжатом и упрощенном виде состоит в том, что если вы не знаете ничего о данных, то не важно, какой алгоритм применять для их обработки. Подчеркнем еще раз, что точная математическая суть этой теоремы гораздо глубже, но для настоящего изложения достаточно упрощенного понимания. На рис. 3 изображена схема работы принципа «Магазин алгоритмов». Магазин в этой схеме представляет возможность выбора продукта на витрине, содержащей все доступные агенту (пользователю) алгоритмы. В доминирующей в настоящее время поведенческой модели общества потребления товаров и услуг такая образность является наиболее удобной и емкой. Аллюзия с магазинами приложений в популярных экосистемах (Apple Store, Google PlayMarket) также имеет смысл, подразумевая легкость применения приложений и возможность попробовать приложение на синтетических данных.

Приведем цитату визионера цифровизации Esko Kilpi: «Приложения сейчас выполняют то, что раньше делали менеджеры» («Apps can do now what managers used to do») [6], которая достаточно точно отражает то, как приложения изменяют ролевую модель бизнес-процессов.

На рис. 3 показана схема «Магазина алгоритмов», примененная в цифровой платформе «ЭРА:ГРАД».

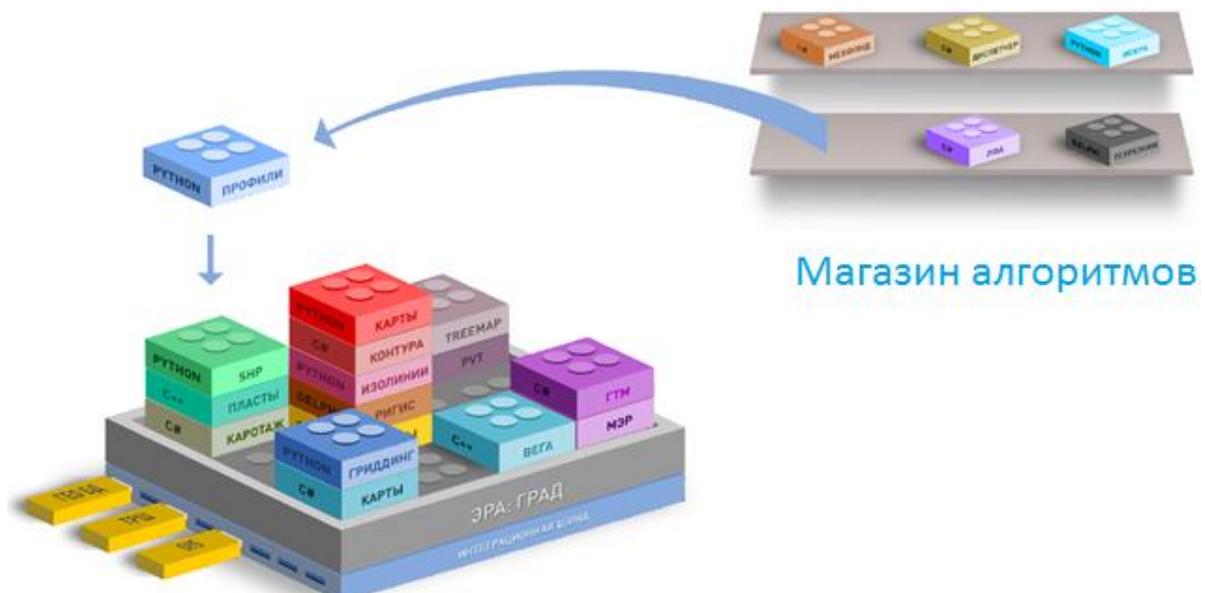


Рис. 3. Схема «Магазина приложений (алгоритмов)»

«Все – это таблица». Корни такого принципа тесно связаны с философией операционной системы Unix «Все – это файл», которая описывает одну из определяющих функций Unix

и его производных: широкий спектр ресурсов ввода/вывода, таких как документы, каталоги, жесткие диски, модемы, клавиатуры, принтеры и даже некоторые межпроцессные и сетевые коммуникации – это простые потоки байтов, представленные в пространстве имен файловой системы. В «ЭРА:ГРАД» лист MS Excel является началом и концом любого расчета, в процессе которого могут быть задействованы различные источники данных, информационные ресурсы и другие таблицы. Сравним этот принцип с распространенной среди Data Scientists оболочкой Python Notebooks. «ЭРА:ГРАД» также оперирует одной вычислительной ячейкой, которая может содержать код, комментарий и результат, но в отличие от Python Notebooks [7] выигрывает по доступности и позволяет строить более гибкие макеты.

Экосистема. Устойчивость работы цифровой платформы «ЭРА:ГРАД» обусловлена изложенными выше принципами. Агент (пользователь), проводящий расчеты на платформе «ЭРА:ГРАД», не зависит от пропускной способности каналов связи, производительности корпоративных информационных систем (источников) и технических регламентов обслуживания информационных систем. В расчетах используется гармонизированная версия данных и алгоритмов, которые будут отражены в результате. При таком построении неважно, кто работает в качестве агента: сотрудник в производственном процессе ДО, студент, выполняющий лабораторную работу, разработчик нового алгоритма или инфобот, занимающийся контролем качества данных.

Заложенные в «ЭРА:ГРАД» принципы позволили создать экосистему, называемую «Экосистема Научного Инжиниринга». Рассмотрим подробнее модель «Экосистемы Научного Инжиниринга», изображенную на рис. 4.



Рис. 4. Модель «Экосистемы Научного Инжиниринга»

Блок Научного Инжиниринга (далее – БНИ) НТЦ в этой схеме выступает в качестве центра компетенций по кросс-функциональным исследованиям и владельцем продукта «ЭРА:ГРАД». Основываясь на бизнес-процессах НТЦ (рис. 2), БНИ инициирует различные формы сотрудничества на платформе «ЭРА:ГРАД» с целью выполнения текущих НИР и создания заделов на будущее через обучение студентов, создание новых алгоритмов отраслевыми консорциумами и выполнение работ по государственным целевым программам и в рамках национальных проектов.

Дискуссия. Мы отдаем себе отчет в том, что изложенные выше принципы могут иметь право на существование только как плод проверки определенных гипотез в определенных условиях. Таковы требования научного метода. Поэтому далее будут сформулированы наши основные гипотезы.

Гипотеза о поступательном характере исследований. Под исследованием в утилитарном смысле будем понимать человеко-машинный процесс по преобразованию входных данных в выходные. Использование результатов («выходов») одного исследования как входных условий для другого представляется важным принципом для построения непрерывных процессов и уменьшения повторов одинаковых действий («каждый раз с нуля»).

Гипотеза об автоматизации рутинных интеллектуальных операций. Исследования состоят из рутинных и творческих (изобретательских) интеллектуальных операций. Важно понимать разницу и не пытаться автоматизировать с помощью ИТ-систем творческие операции, что может привести к выхолащиванию творческого подхода и отторжению цифровой платформы пользователями-исследователями. Наоборот, выявление и автоматизация рутинных интеллектуальных операций сделают цифровую платформу более привлекательной для исследователей, которые смогут полнее выразить свои интенции. Процесс выявления рутинных операций не может проводиться «сверху вниз», так как любой регламент, построенный на основании гипотез о производственных процессах, является лишь их упрощенной моделью, подменяющей задачи автоматизации рутинных операций на задачи декларации собственной уместности. Основой для автоматизации рутинных интеллектуальных операций является системное понимание операционной деятельности исследователей и создание лучших практик по использованию «ЭРА:ГРАД» в непосредственном диалоге с пользователями-исследователями.

Гипотеза о внедрении. ИТ-ландшафт НТЦ в нефтегазовой отрасли представляет симбиоз собственных разработок и коммерческих программных продуктов (Petrel, Eclipse, tNavigator, и т. п.). Создание отдельно стоящих (не взаимодействующих с другими программными продуктами) корпоративных приложений не может быть целью для цифровой платформы, так же как и приобретение коммерческих программных продуктов должно рассматриваться в контексте уже имеющихся собственных разработок и цифровой платформы.

Между коммерческими программными продуктами и собственными разработками в глазах пользователей-исследователей происходит конкурентная борьба. Нельзя не учитывать законы конкуренции (например, «Пять сил Портера») при планировании цифровой платформы [8].

Импортозамещение при внедрении следует понимать как двустороннее вытеснение дорогостоящих и малоиспользуемых опций, а также простейших операций из коммерческих продуктов в собственные разработки. Очевидные преимущества собственных разработок, состоящие в большей гибкости процесса, лучшей поддержке, прозрачности процессов расчетов, скорости и аутентичности, необходимо явно доказывать, а не только подразумевать при разворачивании.

Под количественной оценкой внедрения цифровой платформы следует понимать долю пользователей (целевой аудитории), использующих функции цифровой платформы в своих производственных процессах. Профиль разворачивания платформы цифровой платформы неоднороден: некоторые алгоритмы более востребованы, а некоторые менее.

Гипотеза многовариантности данных о предмете исследования. Многовариантность данных означает возможность одновременного существования различных версий одних и тех же данных. Под разными версиями данных будем понимать не только разную точность, разные масштабы и изменение данных о недрах во времени, но и то, что данные могут быть получены на основании различных методик разными людьми (в разных программных продуктах). Тем не менее данные должны быть сохранены и доступны для сравнения в течение длительного времени как принадлежность процесса, в результате которого они были получены. Цифровая платформа должна предоставлять возможность отслеживания полной истории от момента получения данных и их версии через применение алгоритмов и до конечного результата.

Гипотеза об онтологии. Несмотря на огромное количество алгоритмов, используемых в нефтегазовой отрасли, можно построить их онтологию. В такой онтологии будут присутствовать все характерные признаки для предметной области: таксономия терминов, их описания и правила вывода (рис. 5).

Геомоделирование на этапе разработки



Рис. 2. Фрагмент онтологии алгоритмов разработки месторождений

В основе платформы «ЭРА:ГРАД» заложена онтология алгоритмов. Существует преемственность научного развития, любой новый алгоритм обязан найти свое место в существующей структуре и стать ее частью. Важно отметить, что речь идет именно об онтологии, а не иерархии информационных технологий, как, например, в исследовании [9].

Гипотеза о технологиях. Существует много различных технологий для разработки цифровых платформ. Наша гипотеза состоит в том, что в стратегической перспективе не так важен выбор какой-то одной технологии и «религиозное» почитание ее преимуществ и выискивание недостатков перед другими. Основой для этой гипотезы служат исторические примеры:

- крупнейшая цифровая платформа Facebook создана на PHP, хотя PHP используется в основном начинающими веб-программистами;
- в определенный момент времени компания Apple признала нецелесообразность собственных процессоров и операционной системы и перешла на процессоры Intel и сделала клон операционной системы FreeBSD;
- платформа Яндекс.Маркет изначально имела пользовательский (бекенд) интерфейс на MS Access.

И этот список можно продолжать. Наличие пользователей и их удовлетворенность гораздо важнее используемых разработчиками технологий разработки программного обеспечения.

Заключение. В настоящей работе рассмотрены принципы построения цифровой платформы для научно-технического центра в нефтегазовой отрасли. Авторы привели результаты практического применения рассмотренных принципов для создания цифровой платформы «ЭРА:ГРАД» для НТЦ «Газпром нефть».

Разработана процессная модель деятельности НТЦ верхнего уровня. Отметим, что возникающие в процессной модели положительные обратные связи нашли отражение в цифровой платформе «ЭРА:ГРАД», которая позволяет развивать экосистему и использовать ее сетевой эффект для перехода от модели экономики спроса на услуги НТЦ к модели экономики предложения услуг НТЦ. В корпоративном дискурсе это изменение соответствует переходу от запроса «Исследуйте возможности увеличения КИН с помощью технологии X» к предложению «Исследованы все возможные решения и наиболее оптимальным по NPV для месторождения X в перспективе до XXXX года является решение XX».

Описанные в статье принципы построения цифровой платформы для научно-технического центра обладают высокой степенью генерализации, так как не привязаны к одному проблемному домену (финансовый учет, геология, инжиниринг). Наоборот, цифровая платформа «ЭРА:ГРАД» является отражением того уровня синергии дисциплин, в котором нуждается научно-производственная организация в данный момент и в стратегической перспективе.

Важной особенностью также является низкий порог вхождения для агента в «ЭРА:ГРАД». Пользовательский интерфейс имитирует под привычную MS Excel страницу, но имеет

расширения для работы по методикам Machine Learning с помощью библиотек TensorFlow [10], Scikit-learn [11], Keras [12] и PyTorch [13], позволяет использовать программно-аппаратную архитектуру параллельных вычислений CUDA, которая существенно увеличивает вычислительную производительность при помощи графических процессоров. Программные агенты (инфоботы) используют API платформы ЭРА:ГРАД.

В заключение отметим, что контроль за методиками расчетов не менее важен, чем контроль за используемыми данными. Авторы неоднократно сталкивались в своей практике с «гениальными открытиями», основанием для которых послужило неграмотное использование алгоритмов и непонимание физических процессов, которые они моделируют. Возможности отследить смещенные экспертные оценки, ставшие результатом таких «открытий», через отпечаток (fingerprint) задействованного алгоритма на цифровой платформе «ЭРА:ГРАД» представляется уникальным рычагом методического контроля за научной честностью.

Литература

1. Велихов Е. П., Бетелин В. Б., Кушниренко А. Г. Промышленность, инновации, образование и наука в России. М. : Наука, 2009. 141 с.
2. Куприяновский В. П., Соколов И. А., Талашкин Г. Н., Дунаев О. Н., Зажигалкин А. В., Распопов В. В., Намиот Д. Е., Покусаев О. Н. Цифровая совместная экономика: технологии, платформы и библиотеки в промышленности, строительстве, транспорте и логистике // International Journal of Open Information Technologies. 2017. Т. 5, № 6. С. 65–75.
3. Chaum D. Security Without Identification: Transaction Systems to Make Big Brother Obsolete // Communications of the ACM. 1985. Vol. 28, No. 10. P. 1030–1044.
4. Никлаус В. Алгоритмы и структуры данных. Саратов : Профобразование, 2017. 274 с.
5. Wolpert D. H., Macready W. G. No Free Lunch Theorems for Optimization // IEEE Transactions on Evolutionary Computation. 1997. Vol. 1, No. 1. P. 67–82.
6. Flinn K. Leadership Development: A Complexity Approach. London : Routledge, 2018. 40 p.
7. Kluyver T., Ragan-Kelley B., Pérez F., Granger B. E., Bussonnier M., Frederic J., Ivanov P. Jupyter Notebooks – a Publishing Format for Reproducible Computational Workflows // Conference ELPUB. 2016. P. 87–90.
8. Портер М. Международная конкуренция: конкурентные преимущества стран. М. : Альпина Паблшер, 1993. 896 с.
9. Гимранов Р. Д. Стратификация информационных систем // Вестник кибернетики. 2016. № 1. С. 57–62.
10. Abadi M., Barham P., Chen J., Chen Z., Davis A., et al. TensorFlow: a System for Large-Scale Machine Learning // OSDI. 2016. Vol. 16. P. 265–283.
11. Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B. et al. Scikit-learn: Machine Learning in Python // Journal of Machine Learning Research. 2011. Vol. 12. P. 2825–2830.
12. Gulli A., Pal S. Deep Learning with Keras. : Packt Publishing Ltd, 2017. 318 p.
13. Paszke A., Gross S., Chintala S., Chanan G., Yang E. et al. Automatic Differentiation in PyTorch // 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA. 4 p.