

# ОБЩИЕ ВОПРОСЫ GENERAL QUESTIONS

УДК 004.67, 004.75

DOI: 10.17587/it.24.219-232

**А. В. Юрченко**, канд. физ.-мат. наук, первый зам. директора, e-mail: yurchenko@ict.nsc.ru, Институт вычислительных технологий Сибирского отделения Российской академии наук (ИВТ СО РАН), г. Новосибирск

## О сервисном подходе к формированию и оценке востребованности киберинфраструктуры науки

*Рассматриваются некоторые проблемы поддержки и развития информационно-вычислительной и телекоммуникационной инфраструктуры научных исследований как системы научных ИТ-сервисов. Определен ряд основных понятий научной киберинфраструктуры, научного ИТ-сервиса и их составляющих, устанавливаются взаимосвязи между ними, рассматриваются особенности классификации научных ИТ-сервисов и вопросы оценки их стоимости с примерами сервисов ключевых академических ИТ-провайдеров, обеспечивающих ИТ-ресурсами подведомственные ФАНО России научные учреждения на Дальнем Востоке, в Сибири и в г. Москва. Результаты работы могут быть использованы как для постановки и решения стратегических задач развития киберинфраструктуры российской науки, так и для уточнения технологий оценки востребованности и целесообразности поддержки отдельных компонентов такой инфраструктуры.*

**Ключевые слова:** киберинфраструктура науки, научный ИТ-сервис, наукоёмкость и ресурсоёмкость ИТ-сервиса, востребованность и полезность научного ИТ-сервиса, классификация научных ИТ-сервисов

### Введение

Технологии преобразуют мир, и наука не остается в стороне от этого преобразования. Являясь породителем новых технологий, сфера исследований и разработок является и главным их потребителем. Это верно и в отношении информационных технологий (ИТ). Новые возможности по сбору, хранению, обработке и анализу данных, обмену и распространению информации отразились, в первую очередь, на научном сообществе. Теперь, когда можно обмениваться любыми результатами исследований практически мгновенно и в любых количествах, остро стоят вопросы упорядочивания огромных массивов научных данных, обеспечения их целостности, сохранности и доступности, а самое важное — их максимально эффективного совместного использования для решения новых научных и практических задач. Актуальность этих вопросов только возросла в последнее время, когда практически вся измерительная техника, используемая для научных исследований, перешла в цифровой формат и генерирует большие объемы данных.

Передача, хранение и обработка данных — не бесплатный процесс. Его обеспечение требует высоких капиталовложений, аппаратных, про-

граммных и человеческих ресурсов. С одной стороны, это рутинная работа, в совершенстве освоенная современными провайдерами ИТ-услуг, такими как Microsoft, Amazon, Google и др. С другой стороны, научное сообщество предъявляет особые требования к такого рода услугам и не всегда готово оплачивать их, предпочитая покупать оборудование, генерирующее данные, и нередко просто забывая о проблемах, связанных с хранением этих данных и обеспечением доступа к ним и к полученным на их основе результатам.

Крупные игроки рынка облачных услуг, например Microsoft, осознавая, что научное сообщество, с одной стороны, не обладает большими финансовыми возможностями, с другой стороны, имеет достаточно умеренные, в сравнении с бизнес-клиентами, потребности в ИТ-ресурсах, при этом может обеспечить прорыв в решении имеющихся у самих компаний проблем научного и технологического характера, запускают программы, в рамках которых облачные ресурсы для исследований и разработок выделяются бесплатно на конкурсной основе [1]. В Российской Федерации такого пока не наблюдается, в первую очередь потому, что не сформировались игроки рынка облачных услуг, сопоставимые по масштабам с Microsoft, Google и Amazon. Но кро-

ме того, в то время как те же Microsoft, Google и Amazon сами разрабатывают технологии и являются крупнейшими заказчиками и потребителями научных разработок в области ИТ, российские игроки чаще всего являются пользователями готовых ИТ-продуктов в виде аппаратных и программных платформ, разработанных за рубежом, а когда занимаются самостоятельными разработками, то опираются преимущественно либо на зарубежные исследования, либо на исследования времен СССР.

Отсутствие внутри Российской Федерации поставщиков облачных услуг, готовых бесплатно (или условно бесплатно) отдать часть своих ресурсов для научного сообщества, заставляет российских ученых самостоятельно решать вопросы по обеспечению своих потребностей в передаче, хранении, обработке данных, распространении информации и ее оперативном обсуждении. В некотором плане большая часть этих потребностей может быть покрыта возможностями, предоставляемыми международными корпорациями. Так, программа по предоставлению ресурсов Microsoft Azure для научно-исследовательских групп [1] является открытой, и российские ученые также могут стать ее участниками, а облачные сервисы Google давно стали площадкой для обмена информацией и совместной работы с документами для многих российских ученых. Однако вопросы национальной безопасности и сохранения информационной независимости заставляют искать внутренние решения.

## **1. Проблема формирования киберинфраструктуры науки**

Хотя термин "киберинфраструктура" является далеко не новым (его происхождение связывают с пресс-конференцией Ричарда Кларка в 1998 г. [2]) и звучит несколько архаично, предлагается использовать его в сочетаниях "киберинфраструктура науки" и "научная киберинфраструктура" для обозначения комплекса инфраструктурных решений в области ИТ, обеспечивающих научное сообщество информационно-вычислительными и телекоммуникационными ресурсами, необходимыми для проведения исследований, сохранения, опубликования и обсуждения их результатов в цифровом виде. Интенсивное развитие такой инфраструктуры в США и Европе приходится на 2000-е годы [3, 4].

Задача построения киберинфраструктуры российской науки все еще не решена в достаточной степени, хотя соответствующая проблематика активно обсуждалась начиная со второй половины 2000-х годов, правда, ассоциируясь больше с суперкомпьютерными вычислениями [5, 6]. Попытки построить комплексную киберинфраструктуру

науки в России не могли увенчаться успехом в том числе потому, что задача является слишком масштабной и не имеет каких-то единых подходов к своему решению. Тем не менее, отдельные элементы такой инфраструктуры создаются, и вопрос создания общей киберинфраструктуры возможно перевести в плоскость интеграции существующих элементов, а ее развитие рассматривать как эволюцию этих элементов и инструментов совместного управления их использованием. В обозначенном виде проблема формирования целостной киберинфраструктуры научных исследований сводится к идейному оформлению единства тех разрозненных ее элементов, которые уже созданы, к созданию системы учета самих ИТ-ресурсов и их утилизации, оценке потребности в ИТ-ресурсах и своевременной модернизации элементов этой инфраструктуры.

Для того чтобы сделать еще один шаг в этом направлении, предлагается на примере трех крупных поставщиков ИТ-услуг научным организациям, подведомственным ФАНО России: Межведомственного суперкомпьютерного центра РАН, Института вычислительных технологий СО РАН, Вычислительного центра ДВО РАН, рассмотреть принципы организации доступа других учреждений к поддерживаемым этими организациями информационным, вычислительным и телекоммуникационным ресурсам с примерами таких ресурсов, оценить вклад этих ресурсов и различных сервисов на их основе в решение научных проблем. Также необходимо, опираясь на накопленный опыт предоставления ИТ-услуг, сформулировать ключевые понятия, которыми можно в дальнейшем оперировать при описании тех или иных решений, систематизировать ИТ-услуги и построить расширяемое описание иерархии научных ИТ-сервисов.

## **2. Научные ИТ-сервисы: некоторые определения**

В первую очередь, необходимо отделить киберинфраструктуру науки от традиционных ИТ-услуг, компьютерной техники и телекоммуникационного оборудования общего назначения. Для этого потребуется определить понятие "научного ИТ-сервиса" как особого вида "услуги", которая, с одной стороны, опирается на стандартные программные и аппаратные компоненты, свойственные ИТ-инфраструктуре общего пользования, с другой стороны, обладает спецификой, которая не реализуется в бизнес-ориентированной ИТ-среде, но в которой остро нуждается научное сообщество для успешного решения своих научно-исследовательских задач.

Под *научным ИТ-сервисом* предлагается понимать реализованную в программном и аппарат-

ном виде технологию работы с научными данными. Отметим, что программный и аппаратный инструмент становится ИТ-сервисом, когда пользователи получают доступ к нему через специальные протоколы, либо API, либо приложения, в том числе веб-приложения. К научным данным отнесем результаты научных наблюдений, измерений и вычислений, в том числе преобразований и иных видов обработки данных, представленные в цифровой форме. Технология работы с научными данными — это последовательность (цепочка) действий с ними, приводящая к некоторому заданному результату, имеющему научную ценность. Основные действия с научными данными — это генерация, обработка и анализ, дополнительные — систематизация и представление (в том числе визуализация). Вспомогательные действия: базовые — передача, хранение и вычисления, высокоуровневые — сбор, обмен и публикация (распространение), поиск. Пример типичной цепочки работы с научными данными: генерация — сбор и предварительная обработка — систематизация и хранение — преобразование (обработка) и анализ — публикация.

Генерацией научных данных назовем процесс, в результате которого появляются новые данные, имеющие научную ценность. Примеры генерации научных данных — это научные измерения, например, в результате дистанционного зондирования Земли из космоса, мониторинга сейсмособытий и событий в морях и океанах, секвенирования геномов, компьютерной томографии, энцефало- и кардиографии в исследовательских целях, испытания материалов и конструкций, их спектрографии, проведения опытов на ускорителях частиц, наблюдения за дальним и ближним космосом и многого другого. Не менее важными примерами процессов генерации научных данных являются компьютерное моделирование и преобразование в результате обработки и анализа других данных. Пример генерации научных данных, о котором нельзя забывать, — это написание научных трудов, в том числе отчетов, статей и тезисов, монографий и т.п.

Обработка и анализ научных данных — процесс преобразования одних научных данных в другие научные данные, информацию и знания. Типичные примеры методов и механизмов обработки и анализа научных данных — фильтрация, кластеризация, классификация, статистический или интеллектуальный анализ и др. Систематизация — процесс поиска и установления порядка в наборах данных, построения системы связей между ними, например, в форме цифровых каталогов и электронных библиотек. Представление — процесс преобразования данных в визуальные объекты, от таблиц до диаграмм, графиков и иных изображений.

Передача и хранение научных данных мало отличаются от таковых для бизнеса и иных видов

цифровых данных, а вычисления — это нижний уровень обеспечения компьютерной обработки данных и моделирования. Эти действия опираются на базовые возможности аппаратной инфраструктуры центров обработки данных, вычислительных центров и телекоммуникационной среды и, в свою очередь, являются базовыми для выполнения остальных действий с данными. Сбор данных — автоматизированный или автоматический процесс агрегации данных из заданных источников, чаще всего — измерительных приборов, с частичной или полной их систематизацией перед дальнейшим хранением. Обмен и публикация — действия по организации доступа к научным данным и информации о них для исследователей и других заинтересованных лиц, отличных от получившего (сгенерировавшего, собравшего) их исследователя — владельца данных.

### **3. Научные ИТ-сервисы: примеры и варианты классификации**

Рассмотрим примеры научных ИТ-сервисов, предоставляемых несколькими ключевыми центрами ИТ-обеспечения подведомственных ФАНО России научных организаций, определим их состав и место в соответствии с введенными определениями и сопоставим установленное место с введенной и используемой ФАНО России классификацией.

Научные ИТ-сервисы, заявленные организациями, осуществляющими поддержку общей (коллективного пользования) ИТ-инфраструктуры научных учреждений ФАНО России в регионах и в центральной части, разнесены по следующим категориям (по материалам заседания рабочей группы Комиссии по информатизации ФАНО России по оценке заявок на поддержку ИТКИ для обеспечения функционирования и развития научных информационных сервисов, Москва, 28 апреля 2017).

1. Автоматизированные комплексы для биологии и медицины.
2. Вспомогательная узкоспециализированная информационная система.
3. Информационно-библиотечное сопровождение научной деятельности.
4. Информационные системы для коллективной работы ученых.
5. Системы компьютерной поддержки научных исследований.
6. Системы хранения научных данных общего назначения.
7. Специализированное программное обеспечение для моделирования процессов и явлений.
8. Специализированные научные экспертные информационные системы.
9. Суперкомпьютерные вычисления на специализированных устройствах.

10. Суперкомпьютерные вычисления общего назначения.

11. Тематические базы данных.

В *приложении* представлены списки научных ИТ-сервисов, работа которых обеспечивается ресурсами МСЦ РАН (филиал ФГУ ФНЦ НИИСИ РАН), ИВТ СО РАН и ВЦ ДВО РАН. Приведены наименования и краткие описания сервисов, указаны категории, к которым сервисы отнесены рабочей группой Комиссии по информатизации ФАНО России.

Ясно, что приведенный список категорий не является исчерпывающим и однозначным. Многие категории имеют частный характер, хотя их выделение обосновано наличием достаточного числа сервисов-представителей. При этом множество сервисов не удалось объединить в какие-либо специальные категории, и они были отнесены к категории иных научных ИТ-сервисов, которая названа "Системы компьютерной поддержки научных исследований". Несмотря на имеющиеся недостатки построенной классификации удалось оперативно решить задачу категоризации научных ИТ-сервисов в целях определения объемов финансовой поддержки, необходимой для обеспечения их работы. Для этого выделены наиболее часто встречаемые типы сервисов, выполнена оценка их ресурсоемкости и наукоемкости (табл. 1). Под наукоемкостью научного сервиса здесь и далее будем понимать необходимость использования труда научных работников, а под ресурсоемкостью — использования аппаратных и программных ресурсов, а также труда инженерного персонала.

Анализируя представленные списки сервисов и категорий сервисов, можно заметить, что

они существенно отличаются по наукоемкости, т.е. степени привлечения при их эксплуатации и создании научного труда (табл. 1). Наукоемкость ИТ-сервиса может возникать на этапе разработки, когда сама разработка и внедрение сервиса требуют решения новых научных задач, на этапе наполнения содержимым, как в случае тематических баз данных, когда добавляется содержимое, являющееся результатом научной деятельности, а также на этапе сопровождения, когда само использование сервиса требует научных консультаций, как, например, в случае с суперкомпьютерными вычислениями. Для упрощения последующей классификации в табл. 1 этапы "наполнение содержимым" и "сопровождение" объединены и названы просто "Сопровождение", что не должно привести к снижению качества классификации.

Рассмотрим подробнее вопрос ресурсоемкости ИТ-сервисов и для начала выделим набор сервисов, поддержка работоспособности которых требует скорее высококвалифицированной инженерной рабочей силы и особо не нуждается на этапе реализации и сопровождения в научных кадрах. Такие сервисы, как правило, являются и наиболее ресурсоемкими, так как требуют для своей работы дорогостоящего оборудования или внешних, приобретаемых услуг, а также существенных трудозатрат квалифицированных инженеров, как на этапе "создания", так и на этапе "поддержки".

В первую очередь, это базовые сервисы поддержки телекоммуникационной и вычислительной среды, такие как "Базовая информационно-телекоммуникационная инфраструктура для поддержки научных исследований", "Опорные

Таблица 1

Оценка наукоемкости и ресурсоемкости ИТ-сервисов по категориям

Наименование категории ИТ-сервисов	Наукоемкость		Ресурсоемкость	
	Разработка и развитие	Сопровождение	Создание и развитие	Поддержка работы
Автоматизированные комплексы для биологии и медицины	Высокая	Средняя	Средняя	Средняя
Вспомогательная узкоспециализированная информационная система	Средняя/ Высокая	Средняя/ Высокая	Средняя/ Высокая	Низкая/ Средняя
Информационно-библиотечное сопровождение научной деятельности	Низкая	Низкая	Средняя	Низкая/ Средняя
Информационные системы для коллективной работы ученых	Низкая	Низкая	Средняя/ Высокая	Средняя
Системы компьютерной поддержки научных исследований	Низкая/ Средняя	Низкая/ Средняя	Средняя/ Высокая	Средняя/ Высокая
Системы хранения научных данных общего назначения	Средняя	Низкая	Высокая	Высокая
Специализированное программное обеспечение для моделирования процессов и явлений	Высокая	Средняя	Средняя/ Низкая	Средняя/ Высокая
Специализированные научные экспертные информационные системы	Высокая	Высокая	Средняя	Низкая/ Средняя
Суперкомпьютерные вычисления на специализированных устройствах	Средняя	Средняя	Высокая	Высокая
Суперкомпьютерные вычисления общего назначения	Средняя	Высокая	Высокая	Высокая
Тематические базы данных	Низкая	Высокая	Средняя	Низкая

системы и компоненты информационно-телекоммуникационной инфраструктуры поддержки научных исследований" или "Система гарантированной передачи научных данных ("Федерация")". С одной стороны, такие сервисы можно полностью передать на аутсерсинг соответствующим провайдерам, с другой стороны, практика показывает, что научные задачи требуют консолидации таких ресурсов, и более эффективно их можно использовать, в том числе закупать и поддерживать, при определенном уровне централизации. Кроме того, большая часть физической инфраструктуры, составляющей основу функционирования этих сервисов внутри научных центров, исторически формировалась и создавалась поддерживающими эти сервисы организациями во времена, когда само понятие "глобальные компьютерные сети" лишь начинало выходить за пределы научных лабораторий. Созданные в результате этой деятельности корпоративные академические компьютерные сети стали в свое время основной для организации междисциплинарного научного взаимодействия и развития научных исследований в целом. Это, в частности, подтверждается получением коллективом ученых из Сибири под руководством академика Ю. И. Шокина премии Правительства Российской Федерации в области науки и техники 2012 г. за создание информационно-телекоммуникационной инфраструктуры междисциплинарных научных исследований как основы экономического и социального развития восточных регионов России [7], а в 2014 г. — получением аналогичной премии коллективом под руководством академика Г. И. Савина за комплекс научных, технических, технологических и организационных решений для обеспечения суперкомпьютерными технологиями и информационными ресурсами научных исследований и инновационных разработок [8]. Таким образом, наукоемкость этих сервисов является историческим фактом, а их значимость для выполнения современных научных исследований нельзя умалить.

Более высокоуровневые ИТ-сервисы, связанные с выделением вычислительных ресурсов и пространства для долговременного и краткосрочного хранения научных данных, могут быть реализованы различными способами, от простейших: путем предоставления прямого доступа к серверам и системам хранения данных (модель "инфраструктура как сервис" — IaaS), до более сложных: путем предоставления доступа к системам управления заданиями на вычислительных кластерах и grid-системах, к различным блочным, файловым и более высокоуровневым протоколам и API систем хранения, облачным системам управления ресурсами ЦОД и суперкомпьютерных центров, вплоть до систем виртуализации, СУБД и т.д. (модели "платформа как сервис" — PaaS и ее

производные: "базы данных как сервис" — DBaaS, "промежуточное программное обеспечение как сервис" — MWaaS и др.). Примерами таких сервисов являются "Сервис хранения научных данных" и "Суперкомпьютерные вычисления с массивным параллелизмом и хранение данных", "Система поддержки научных информационных сервисов".

В целом вокруг задач выделения вычислительных ресурсов и предоставления пространства для хранения и обмена данными можно построить целые иерархии сервисов, направленных на решение задач различных уровней. Так, в ИВТ СО РАН строится набор сервисов хранения, обмена и совместной работы с документами [9]. Нижним уровнем является сервис выделения пространства хранения научных данных через блочный либо файловый протоколы. Этот сервис ориентирован на монопольное пользование данными в рамках одного научного подразделения (коллектива) и долговременное хранение, включающее резервное копирование больших объемов данных. Следующий уровень — система обмена данными и совместной работы с документами на базе платформы Nextcloud, являющаяся аналогом Dropbox и Google Drive. Этот сервис предназначен для среднесрочного размещения не очень больших по объему данных и документов в целях обмена ими, а также для совместной работы с документами с поддержкой их версионности и встроенными веб-редакторами форматов Open Document. Верхним уровнем в этой иерархии должна стать интегрированная система хранения и анализа научных данных, которая объединит функционал электронной библиотеки с возможностями систем автоматического сбора научных данных и систем обработки и анализа данных.

Необходимо выделить ИТ-сервисы, которые способствуют выполнению научных исследований опосредовано, за счет решения организационных задач. Это сервисы обмена информацией и данными, в том числе доступа к библиотекам научных документов, совместной работы с документами, поддержки проведения научных мероприятий, организации проведения научных конкурсов и сбора информации о результатах научной деятельности, поддержки проектной, научно-издательской деятельности, в том числе научных журналов и др., такие как "Информационный сервис управления издательской деятельностью научной организации", "Информационная система управления научными проектами (ИС Грант)", "Система поддержки организации и проведения научных конференций". К ним относятся и сервисы поддержки передачи электронных сообщений и коллективных коммуникаций, такие как "Корпоративная коммуникационная система", построенная на платформе с открытым исходным кодом Zimbra и на платформе Microsoft и позволяющая организовывать общие про-

странства для обмена текстовой информацией, организовывать встречи и собрания, осуществлять оперативную аудио- и видеосвязь, многосторонние аудио- и видеоконференции. Такие сервисы часто не являются уникальными, но, тем не менее, играют важную роль в организации научной деятельности и взаимодействия научных групп и отдельных ученых.

Большая часть высокоуровневых научных ИТ-сервисов опирается на возможности перечисленных выше инструментов. Например, тематические базы данных задействуют сложный стэк ИТ-сервисов нижнего уровня. Они базируются на веб-платформах, роль которых обычно играют различные CMS (Content Management System — система управления содержимым веб-сайта), используя для размещения данных системы управления базами данных (СУБД). В свою очередь, CMS и СУБД размещаются на платформах виртуализации, использующих серверные ресурсы центров обработки данных (ЦОД или датацентр), а для размещения самих данных используются ресурсы систем хранения данных. При этом доступ к тематическим базам данных как для их пользователей, так и для поставщиков содержимого осуществляется через компьютерную сеть, при задействовании телекоммуникационных ресурсов инфраструктуры. Аналогичные стеки могут быть описаны для любого высокоуровневого научного ИТ-сервиса. Таким образом, вспомогательные и ненаучные ИТ-сервисы позволяют существенно упростить и ускорить процессы разработки новых научных ИТ-сервисов, а также их последующее сопровождение и поддержку. Отсюда вытекают и некоторые сложности в оценке стоимости научного ИТ-сервиса, так как их ресурсоемкость состоит и из ресурсоемкости используемых сервисов нижнего уровня.

#### 4. "Цена" научного ИТ-сервиса

Научный ИТ-сервис, как и любой другой ИТ-сервис, необходимо разработать, внедрить, сопровождать и развивать, тогда его можно будет эффективно использовать для решения актуальных научных задач. Соответственно, и "себестоимость" научного ИТ-сервиса будет складываться из стоимости каждого из этих этапов. Вопросы разработки и внедрения новых научных ИТ-сервисов целесообразно рассматривать в разрезе выполнения научно-исследовательских проектов, в частности, в рамках государственного задания на проведение фундаментальных научных исследований. Дальнейший жизненный цикл этих продуктов научного труда традиционно выходит за рамки фундаментальных и даже прикладных научных исследований, становясь часто инженерной, вспомогательной задачей, которая обычно остается "за кадром" основной деятельности научных органи-

заций, из-за чего, собственно, пользователи быстро теряют интерес к большинству создаваемых научным сообществом ИТ-продуктов. Речь идет об этапах сопровождения работы и развития научных ИТ-сервисов.

Для поддержки работы научных ИТ-сервисов, безусловно, нужна аппаратная инфраструктура (серверы, их кластеры, системы хранения данных, компьютерные сети, инженерная часть центров обработки данных, включая системы энергоснабжения, охлаждения и пожаротушения) и телематические сервисы. С точки зрения финансовых затрат это наиболее ресурсоемкая составляющая стоимости сервисов. Однако потребности сервисов из разных категорий в таких ресурсах очень разнообразны. Сбалансировать их частично позволяет использование виртуальных машин вместо отдельных серверов и систем динамического выделения ресурсов. Наиболее эффективно такие инструменты начинают работать при определенном уровне централизации ресурсов. Нужно отметить, что некоторые виды ИТ-сервисов, например из категории "Специализированное программное обеспечение для моделирования процессов и явлений", могут потребовать ресурсов, которых нет ни у одного центра обеспечения (вычислительного центра, центра хранения и обработки данных), тогда необходимо "виртуально" объединять территориально распределенные ресурсы, что возможно только при наличии развитых сетей и телекоммуникационных ресурсов. Три ключевых типа ресурсов: телекоммуникационные, вычислительные и хранения, составляют основную часть "ресурсоемкости" сервиса как на этапе разработки и развития, так и на этапе поддержки работы. Необходимость поддержки работы этих базовых ресурсов квалифицированными ИТ-инженерами, стоимость труда которых сопоставима со стоимостью труда научных работников, а нередко и превышает ее, только увеличивает "себестоимость" ИТ-сервисов. Однако, как показывает анализ рынка подобных услуг (табл. 2), стоимость их аутсерсинга может быть существенно выше, чем в случае рациональной централизации таких ресурсов.

Научное сопровождение ИТ-сервисов, которое заключается в обучении и консультировании пользователей либо в совместном с ними решении их научных задач, является отдельной проблемой. Традиционно в российских научных организациях считается, что научный работник должен выполнять самостоятельные исследования, публиковать их результаты и продумывать вопросы их внедрения. В итоге труд научных работников по сопровождению научных сервисов, к которым относятся не только ИТ-сервисы, но и другие виды "услуг", например, по проведению экспериментов и испытаний, различных измерений, созданию научных приборов и систем,

## Оценка стоимости коммерческой аренды базовых информационно-вычислительных и телекоммуникационных ресурсов ИВТ СО РАН

Описание сервиса, тип	Единица измерения ресурса и количество	Стоимость поддержки*, млн руб. в год	Стоимость аренды аналога
Высокоскоростной обмен данными (специализированный 10 Gbit сегмент)	10 Gbit линии связи, 6 шт	—	» 3,5
Размещение оборудования в серверных залах/залах ЦОД	Серверные стойки, 40 шт.	—	> 12,0
Виртуализация информационно-вычислительных систем и инфраструктура как сервис (IaaS)	Серверы и кластеры, 20 шт.	—	> 3,5
Отказоустойчивое хранение данных**	Петабайт хранилища, 1,0/2,5	—	> 12,0/30,0
<b>ИТОГО</b>		13,8/22,3	> 31,0/49,0**
* Без учета/с учетом амортизации оборудования			
** 2017/2018 гг.			

в том числе программных, далеко не всегда получает должную оценку. В качестве возможного "поощрения" такого труда можно рассматривать включение выполняющих его научных работников в число соавторов публикаций, в которых используются результаты измерений, испытаний, экспериментов. Однако, не являясь основными авторами таких публикаций, они не всегда могут и претендовать на включение представленных в них результатов в свои квалификационные работы, диссертации. Это еще одна сторона "цены" научного ИТ-сервиса — трудовой ресурс высококвалифицированных научных работников.

По аналогии с любым другим продуктом, кроме себестоимости необходимо учитывать и востребованность научных ИТ-сервисов, которые, в совокупности, формируют некоторый аналог "рыночной цены" сервиса, а также уметь оценивать их "результативность". Востребованность сервиса можно оценивать по числу результативных обращений. Однако определение этого числа является одной из ключевых проблем ИТ-индустрии в целом, особенно ярко проявившей себя в области интернет-рекламы, а теперь и в области продвижения не только товаров, но и информации, идей, мнений, когда множество так называемых фэйковых (от англ. fake) действий осуществляется программными инструментами либо людьми, для которых генерация "фальшивых" действий (так называемых накруток) — это профессия.

Так, для ИТ-сервисов с авторизацией пользователя число зарегистрированных пользователей — лишь нулевое приближение к оценке эффективности. К нему долгое время апеллировали операторы мобильной связи, позже — владельцы информационных ресурсов для демонстрации собственной значимости, в том числе, чтобы заинтересовать рекламодателей. Но качество этой характеристики нивелировано современными технологиями генерации фальшивых пользователей и процессами "отмирания пользователей".

Достаточно быстро было осознано, что зарегистрированный пользователь совсем не обязательно пользуется ресурсом. Появилось понятие "активный пользователь" — пользователь, который за определенный период хотя бы однажды авторизуется в системе или производит хоть одно действие, например телефонный звонок или отправку сообщения. Со временем активность зарегистрированных пользователей стала детально отслеживаться и изучаться владельцами информационных ресурсов и сервисов, выделившись в отдельную индустрию, изначально ориентированную именно на рекламный рынок.

В случае с научными ИТ-сервисами верификация пользователей все еще может быть организована достаточно эффективно для того, чтобы можно было использовать характеристики "число зарегистрированных пользователей" и "число активных пользователей" для оценки востребованности того или иного ИТ-сервиса. Однако и этот подход, не говоря уже о более сложных методах учета (аккаунтинга) действий пользователей, может стать для научного ИТ-сервиса отдельной сложной проблемой, решение которой, особенно с учетом отсутствия монетизации (так как она не является главной целью разработки и внедрения научного ИТ-сервиса, главной целью является применение таких сервисов при решении научных задач для генерации нового научного знания), выходит за рамки интересов разработчиков и даже обеспечивающих работоспособность структур. Тем не менее, есть пример научных ИТ-сервисов, где аккаунтинг является достаточно проработанным вопросом — это суперкомпьютерные вычисления. Так как традиционно суперкомпьютерные ресурсы являются "дефицитным" продуктом, т. е. желающих воспользоваться им больше, чем имеется возможностей, а также потому, что очень быстро возникла необходимость sharing'a — одновременного совместного использования этих ресурсов, воз-

никло и желание повысить эффективность освоения суперкомпьютеров путем введения автоматизированных или полностью автоматических систем управления заданиями и, соответственно, учета использования ресурсов. Аналогичная проблема была решена для телекоммуникационных ресурсов по заказу операторов связи, для которых это уже был вопрос, напрямую связанный с монетизацией их услуг. В настоящее время можно считать решенной и проблему учета использования основных "облачных" ресурсов (пространств для хранения данных, серверных ресурсов) как исчислимых (Терабайты хранилища \* часы/сутки/месяцы, Гигагерцы \* ядра CPU/GPGPU \* часы, Гигабайты RAM \* часы и др.). Конечно, качественная система учета использования компьютерных или телекоммуникационных ресурсов и сейчас является очень дорогостоящим продуктом, приобретение которого имеет смысл только тогда, когда речь идет о монетизации этих ресурсов. Соответственно, закладывая определенный уровень рентабельности, стоимость владения, скорость амортизации, владелец переводит эти характеристики в "рубли". Но может ли позволить себе сделать то же самое наука, когда речь идет о ресурсах для научных исследований, результаты которых далеко не всегда измеряются рублями?

Другая сторона организации учета использования научных ИТ-сервисов на основе авторизации связана с необходимостью построения единой системы для всего стека ИТ-ресурсов, задействуемого этими сервисами. Это прямое следствие того, что ИТ-сервисы могут быть взаимозависимы, и совершенно однозначно, что любой ИТ-сервис использует ресурсы базовых сервисов, при этом с различной интенсивностью. Таким образом, для того, чтобы корректно оценивать востребованность тех или иных ИТ-сервисов и распределять на основе этой оценки ресурсы поддержки, необходимо учесть использование всех задействованных при обращении к конкретному ИТ-сервису ресурсов, представленных также в форме ИТ-сервисов. Нужно помнить и тот факт, что различные базовые ресурсы могут принадлежать/управляться различными поставщиками. Это делает систему учета еще более сложной и дорогостоящей как в разработке, так и в эксплуатации, в том числе за счет повышения трудоемкости ее внедрения и сопровождения.

Еще один важный момент заключается в том, что аутентификация и авторизация снижает удобство использования, а иногда даже противоречит условиям опубликования результатов исследований. Научное сообщество постепенно движется к реализации парадигмы открытых данных [11], что позволит не только верифицировать результаты и выводы, которые делают другие научные коллективы, но и, формируя сложные пулы и иерархии данных, получать из них новые знания,

которые невозможно получить, пользуясь данными из одного источника. Этот тренд в некотором плане дополнительный к тому, что наблюдается в физике элементарных частиц, при котором множество групп исследователей, опираясь на одни и те же данные, ставшие "общим достоянием" (полученные с сенсоров БАК и аналогичных коллаборативных проектов), исследуют различные аспекты одного явления или даже один аспект, распараллеливая решение высокоресурсоемких задач анализа огромных объемов данных. Имея возможность создавать сложные композиции источников данных, исследователь сможет творчески перерабатывать их, извлекая из них новую информацию, новые знания. В этом случае любое ограничение на доступ к данным и механизмам их обработки отрицательно повлияет на научный прогресс, замедляя его. Именно поэтому ряд значимых научных изданий, например PLOS, уже начали требовать открытия данных, на которых основаны публикуемые результаты исследований, предоставляя для этого и специализированные ресурсы [12, 13], а в ряде случаев прямо требуется наличие возможности не ограниченного даже простой аутентификацией свободного доступа к сервисам и ресурсам, как для системы "GTRD — Gene Transcription Regulation Database" [13].

## Выводы и заключение

Поддержка и развитие инфраструктуры научных исследований является одной из важнейших проблем современной науки и научного сообщества. Наличие современной аппаратной и программной базы — необходимое условие проведения исследований на передовом крае мировой науки. Цифровые данные — ключевой драйвер современной науки, а научные коммуникации, без которых невозможно представить научное сообщество, существенно опираются на возможности, предоставленные новыми ИТ. Поэтому информационно-телекоммуникационные и вычислительные ресурсы как основной элемент киберинфраструктуры науки и ИТ-сервисы как форма предоставления доступа к этим ресурсам стали неотъемлемой частью инфраструктуры научных исследований.

Построенная и использованная ФАНО России в 2016 и 2017 гг. система оценки полезности информационно-телекоммуникационной инфраструктуры на основе востребованности научных ИТ-сервисов позволила сделать еще один важный шаг в направлении развития киберинфраструктуры науки. В ходе работы с этой системой выявлен ряд значительных недостатков, среди которых, в частности, отсутствие учета ресурсоемкости ИТ-сервисов, их наукоемкости и результативности, под которой можно понимать

результативность исследований, выполняемых на основе этих ИТ-сервисов. Построенная классификация сервисов также несовершенна и требует значительной доработки. В частности, она все еще позволяет отнести один и тот же сервис в разные категории, а в рамках одной категории есть сервисы, существенно отличающиеся по таким важным характеристикам, как наукоемкость и ресурсоемкость. Тем не менее, эта система стала определенной вехой, позволившей сохранить накопленный научно-технический потенциал в области ИТ-сопровождения научных исследований, утрата которого может стать невосполнимой для всей российской науки.

Предлагаемые специализированными научными организациями научные ИТ-сервисы покрывают широкий спектр потребностей российской науки в информационных, вычислительных и телекоммуникационных ресурсах. Их востребованность растет по мере осознания российским научным сообществом важности цифровых научных данных, обеспечения их сохранности, необходимости создания и развития методов, программных и программно-аппаратных систем для их обработки и анализа. Уже решенные этими своеобразными "сервис-провайдерами" задачи во многом предвосхитили и опередили потребности российского научного сообщества, совершая своеобразное "принуждение к инновациям", способствуя ускоренному включению в общемировое исследовательское поле в его самых современных формах.

Тем не менее, вопрос о формах поддержки развития и содержания киберинфраструктуры науки в России, о ценности тех или иных решений, которые она предоставляет, остается открытым. Ценность научных ИТ-сервисов, безусловно, должна выражаться в тех научных результатах, которые получены с их помощью, а необходимость поддержки должна оцениваться на основе соотношения стоимости содержания и развития к ценности, определяемой востребованностью и результативностью. Однако нельзя забывать и об эффекте "малых величин", в частности ИТ-сервисов, прямые эффекты от которых не столь заметны, но выявить которые, оценить их значимость и указать на необходимость учета которых — важнейшая задача научных экспертов, так как пренебрежение такими "малыми величинами" может иметь огромное влияние, в нашем случае — на возможность и эффективность решения наиболее актуальных научно-технических проблем.

**Благодарности.** Автор благодарит к. т. н. А. А. Сорокина, к. т. н. Б. М. Шабанова, М. М. Харций и других членов Комиссии по информатизации ФАНО России за полезные дискуссии, способствовавшие подготовке материалов статьи, формулировке проблематики и задач, а также формированию представленных решений. Особую признатель-

*ность автор выражает академику Ю. И. Шокину за всестороннюю поддержку проводимой работы, критическую оценку результатов, конструктивные замечания и предложения, которые позволили существенно улучшить представленный материал.*

#### Список литературы

1. Microsoft Azure for Research. URL: <https://www.microsoft.com/en-us/research/academic-program/microsoft-azure-for-research/> (дата доступа 27.12.2017).
2. Press Briefing by Richard Clarke, National Coordinator for Security, Infrastructure Protection and Counter-terrorism; and Jeffrey Hunker, Director of the Critical Infrastructure Assurance Office. News release. The White House Office of the Press Secretary. May 22, 1998. URL: <https://fas.org/irp/news/1998/05/980522-wh3.htm> (дата доступа 27.12.2017).
3. Bottum J. R., Davis J. F., Siegel P. M., Wheeler B., Oblinger D. G. Cyberinfrastructure: In Tune for the Future // *Educare Review*. 43 (4). Jul/Aug. 2008. URL: <https://web.archive.org/web/20080907155037/http://connect.educause.edu/Library/EDUCAUSE%2BReview/CyberinfrastructureInTune/46966> (дата доступа 27.12.2017).
4. Newman H. B., Ellisman M. H., Orcutt J. A. Data-Intensive e-Science Frontier Research in the Coming Decade // *Communications. Association for Computing Machinery*. Nov. 2003. 46 (11). CiteSeerX 10.1.1.72.5841. Freely accessible. DOI: 10.1145/948383.948411.
5. Абрамов С. М., Заднепровский В. Ф., Московский А. А. Отечественные Супер-ЭВМ и грид-системы. Проблемы развития национальной киберинфраструктуры в России // *Тр. XII Научно-практической конференции Университета города Переславля. Переславль-Залесский*, 2008. С. 1—27. URL: <http://www.botik.ru/PSI/RCMS/publications/publ-texts-2008/supercomputers-grid-systems.pdf> (дата доступа 27.12.2017).
6. Шокин Ю. И., Федорук М. П., Чубаров Д. Л., Юрченко А. В. О развитии инфраструктуры суперкомпьютерных и распределенных вычислений в СО РАН // *Информационные технологии и вычислительные системы*. 2011. № 3. С. 9—19.
7. Распоряжение Правительства РФ от 27 февраля 2013 г. № 254-р "О присуждении премий Правительства Российской Федерации 2012 года в области науки и техники". URL: <http://government.ru/docs/643/> (дата доступа 27.12.2017).
8. Распоряжение Правительства РФ от 26 февраля 2015 г. № 303-р "О присуждении премий Правительства Российской Федерации 2014 года в области науки и техники". URL: <http://government.ru/docs/all/95036/> (дата доступа 27.12.2017).
9. Юрченко А. В. К концепции информационно-аналитической системы поддержки научных исследований, основанных на интенсивном использовании цифровых данных // *Вычислительные технологии*. 2017. Т. 22, № 4. С. 105—120.
10. Nathan L. Yo., Stephen F. S., Pardis C. S. Data sharing: Make outbreak research open access // *Nature*. 2015. Vol. 518, Iss. 7540. P. 477—479.
11. Lowenberg D., Ross A., Ganley E. Introducing the PLOS Open Data Collection // *PLOS Collections in Open Data*. 2016. Nov. 10.
12. Haendel M., Vasilevsky N. Data sharing in a modern world; well, maybe not so modern, redux // *PLOS Collections in Open Data*. 2017. Nov. 13.
13. Yevshin I., Sharipov R., Valeev T., Kel A., Kolpakov F. GTRD: a database of transcription factor binding sites identified by ChIP-seq experiments // *NUCLEIC ACIDS RESEARCH*. 2017. Vol. 45, D1. P. D61—D67. DOI: 10.1093/nar/gkw951

# Приложение.

## Списки научных ИТ-сервисов, предоставляемых МСЦ РАН, ИВТ СО РАН, ВЦ ДВО РАН

Таблица 3

ИТ-сервисы МСЦ РАН — филиала ФГУ ФНЦ НИИСИ РАН (Москва)

Наименование	Краткое описание	Кат.
Сервис доступа к электронным каталогам журналов и книг БЕН РАН	Доступ онлайн к электронным каталогам БЕН РАН	3
Распределенный читальный зал БЕН РАН	Доступ к платформам зарубежных издательств, доступных в институтах РАН через распределенный читальный зал РАН	3
Сервис доступа к информационной системе "Волоконная оптика"	Информационная система "Волоконная оптика" включает электронную библиотеку научной литературы по фундаментальным и прикладным проблемам волоконной оптики; труды сотрудников Научного центра волоконной оптики РАН; информационные материалы по применению волоконной оптики и связанным с нею производствам	3
Сервисы "Наблюдательные данные ПРАО" и "Рабочая среда радиоастронома"	Доступ к данным наблюдений Пушинской радиообсерватории	11
Сервисы "Геофизические данные ИЗМИРАН"	Доступ к данным о вариациях магнитного поля Земли, метеоданным, данным зондирования ионосферы и прогнозу солнечной активности	11
Сервис "База данных двойных звезд"	Сервис онлайн-доступа к астрономическим и астрофизическим базам данных	11
Сервис доступа к базе данных VALD3	Сервис онлайн-доступа к базе данных "Vienna Atomic Lines Database"	11
Сервис доступа к базе данных SAO/NASA Astrophysics Data Systems	Сервис онлайн-доступа к базе данных по физике и астрономии SAO/NASA Astrophysics Data Systems	11
Информационная система сопровождения и предоставления цифровых информационных ресурсов	Телекоммуникационный сервис доступа к информационным ресурсам и информационной системе поддержки научных исследований и образовательных программ	5
Сервис формирования и описания баз данных цифровых объектов	Формирование и сопровождение электронных информационных ресурсов	5
Сервис формирования тематических электронных коллекций и виртуальных выставок	Формирование баз данных и тематических коллекций из научных фондов организаций ФАНО России	5
Сервис интеграции цифровых фондов организаций ФАНО России	Сбор, обработка и интеграция цифровых объектов, предоставляемых организациями, подведомственными ФАНО России	5
Сервис распределенного долговременного хранения данных цифровых фондов и тематических коллекций	Обеспечение хранения, сохранности и доступности распределенных цифровых фондов и тематических коллекций	5
Предоставление высокоскоростного доступа к суперкомпьютерным ресурсам МСЦ РАН	Предоставление высокоскоростного доступа к суперкомпьютерным ресурсам МСЦ РАН через сетевую инфраструктуру МСЦ РАН	5
Предоставление высокоскоростного доступа к распределенной инфраструктуре обработки данных БАК	Предоставление высокоскоростного доступа через волоконно-оптические каналы связи и виртуальные локальные сети к распределенной инфраструктуре обработки данных БАК	5
Удостоверяющий сервис EDUROAM	Удостоверяющая инфраструктура для взаимной авторизации пользователей беспроводных сетей, часть международной инфраструктуры EDUROAM	5
Телематический сервис научных организаций	Предоставление службы сервисов доменных имен, адресов	5
Хостинг информационных сервисов и систем научных организаций	Предоставление услуги размещения и поддержки оборудования, предоставление каналов связи для научных организаций, предоставляющих информационные сервисы	5
Сервис организации вычислений с массивным параллелизмом на основе программно-обеспечения для моделирования процессов и явлений	Организация вычислений с использованием прикладных программ моделирования процессов и явлений на универсальных высокопроизводительных ресурсах с использованием решающего поля на основе процессоров общего назначения	7
Суперкомпьютерные вычисления с массивным параллелизмом и хранение данных	Предоставление вычислительных ресурсов для универсальных высокопроизводительных вычислений с использованием решающего поля на основе процессоров общего назначения	10

## ИТ-сервисы ИВТ СО РАН (Новосибирск)

Наименование	Краткое описание	Кат.
Активный архив информационных продуктов на основе данных сенсора MODIS	Предоставление доступа к архиву информационных продуктов, генерируемых в оперативном и архивном режимах на основе данных сенсоров, установленных на спутниках Terra, Aqua, NPP Suomi, поступающих с нескольких приемных комплексов	11
Архив спутниковых данных SPOT 4	Предоставление доступа к архиву данных дистанционного зондирования системы SPOT 4	11
Web-трансляция научных мероприятий	Организация Web-трансляции различных научных мероприятий: семинаров, конференций, совещаний, диссертационных советов	4
Система видеоконференцсвязи СО РАН	Трансляция и многосторонняя связь для научных мероприятий, конференций, заседаний Президиума СО РАН, совещаний ОУС	4
Информационная система CRIS	Учет научных публикаций, конференций, проектов, персон и организаций. Генерация отчетов	3
Корпоративная коммуникационная система	Поддержка индивидуальных и коллективных коммуникаций между учеными и научными коллективами: электронная почта, аудио- и видеосвязь, видеоконференции, обмен сообщениями	5
Распределенная система ZooSPACE	Доступ к распределенным научно-библиотечным информационным ресурсам и базам данных (Новосибирск, Иркутск, Томск, Красноярск)	3
Сервер интерактивных вычислений Jupyter Notebook	Выполнение научных расчетов с использованием специализированных библиотек языка Python в интерактивном режиме с поддержкой возможностей совместной работы	10
Система виртуализации научных информационных систем	Размещение научных информационных сервисов и систем в высоконадежной кластерной системе виртуализации с выделением дискового пространства	6
Система поддержки организации и проведения научных конференций	Информационная система "Конференции" предназначена для комплексной поддержки организации и проведения научных конференций	3
Система хостинга научных информационных ресурсов	Размещение научных информационных веб-ресурсов и систем на основе открытых программных платформ	5
Опорные системы и компоненты информационно-телекоммуникационной инфраструктуры поддержки научных исследований	Ключевые компоненты системы обмена данными и информацией через локальные и глобальные компьютерные сети	5
Высокоскоростной обмен научными данными (специализированный 10 Gbit сегмент)	Обмен научными данными через высокоскоростную телекоммуникационную инфраструктуру — 10 Gbit технологический сегмент ККС (доступно в Новосибирском, Красноярском, Томском, Кемеровском и Иркутском научных центрах)	5
Информационная система поддержки деятельности научного журнала	Сервис поддержки деятельности научных журналов на основе open source платформ (OJS или др.) с электронной редакцией (прием рукописей, организация их рецензирования и отслеживания прохождения, публикация материалов в соответствии с международными стандартами)	3
Сервис хранения научных данных	Хранение научных данных на высоконадежной СХД с доступом по файловым или блочным протоколам	6
Сервис обмена и совместной работы с научными документами	Web-сервис для размещения научных документов и управления доступом к ним на основе open source платформы (аналог DropBox)	3
GTRD — Gene Transcription Regulation Database	База данных результатов транскрипции (ChIP-seq) генов мыши и человека	11
Платформа для анализа биомедицинских данных BioUML	Поиск и визуализация данных из биологических баз данных, статистический анализ биологических данных, анализ биологических данных через Galaxy, моделирование биологических систем	11
Цифровые репозитории результатов научной деятельности	Цифровые архивы полнотекстовых научных документов, публикаций и пр.	3

## ИТ-сервисы ВЦ ДВО РАН (Хабаровск)

Наименование	Краткое описание	Кат.
Базовая информационно-телекоммуникационная инфраструктура для поддержки научных исследований	Доступ к научным информационным системам и сервисам, ресурсам ЦКП, а также обмен научными данными в рамках Дальнего Востока России	5
Система гарантированной передачи научных данных ("Федерация")	Взаимное использование ресурсов крупнейших суперкомпьютерных центров, ЦКП, информационных систем и хранилищ данных (Центральный регион, Сибирь, Дальний Восток)	5
Система поддержки научных информационных сервисов	Обеспечение функционирования научных информационных сервисов в интересах организаций, подведомственных ФАНО России (виртуальные машины, Web-серверы и т.п.)	5
Информационная система управления научными проектами (ИС Грант)	Система предназначена для управления научными проектами в ДВО РАН	8
Информационный сервис "Дистанционный мониторинг активности вулканов Камчатки и Курил" VolSatView	VolSatView обеспечивает специалистов-вулканологов оперативными спутниковыми данными среднего разрешения и различными информационными продуктами, получаемыми на основе их обработки, для мониторинга вулканической активности Камчатки и Курил. Кроме этого, в ИС создан и постоянно пополняется архив спутниковых данных высокого разрешения, позволяющий анализировать различные продукты извержений вулканов (отложения лавовых и пирокластических потоков и др.). Система позволяет работать как с оперативными, так и с архивными данными, накапливаемыми в VolSatView, а также с данными ЦКП "ИКИ-Мониторинг", Объединенной системы работы с данными центров НИЦ "Планета", геопортала ГК "Роскосмос", АИС "Сигнал"	5
Информационный сервис "ВЕГА — Дальний Восток"	Сервис создан для обеспечения доступа к данным распределенной информационной системы коллективного использования данных космического дистанционного зондирования Земли (ИСКИ ДЗЗ) для проведения научной, образовательной и инновационной деятельности в области исследования и контроля состояния окружающей среды в регионах Дальнего Востока России. В настоящее время система ориентирована на предоставление пользователям, ведущим научные исследования на территории Дальневосточного региона, как архивных, так и оперативных спутниковых данных и результатов их обработки. Пользователи Вега ДВ имеют возможности работать с архивами информации, полученной на основе данных спутников NOAA, Метеор М № 1, Terra, Aqua, Landsat и др., а также метоинформацией. Система обеспечивает возможность работы с более чем 10-летними архивами данных	5
Автоматизированная информационная система "Сигнал"	АИС обеспечивает единую информационную среду для работы с данными различных сетей инструментальных наблюдений (видео, геодинамическая т.п.) ДВО РАН и управления их ресурсами. Система осуществляет работу в реальном времени с 30 пунктами наблюдений на территории Дальнего Востока. В состав АИС входят сервисы специализированной обработки данных и математического моделирования	7
Сахалинский сейсмологический сервис	Сейсмический мониторинг о-ва Сахалин и прибрежного шельфа, исследования и получение научной информации о землетрясениях	5
Информационный сервис для исследования и мониторинга вулканической активности Центральных и Южных Курил	Мониторинг и исследование вулканической активности Курильских островов	5
Геоинформационная система "Природные ресурсы Магаданской области"	Целью проекта является публикация геопространственных баз данных по геологии, минеральным ресурсам и биоразнообразию Магаданской области; организация доступа к просмотру и редактированию баз данных; также создание инструментов пространственного анализа геоданных через интерактивный графический Web-интерфейс посредством геоинформационных технологий	5
Геопортал ИВиС ДВО РАН	Тематический Web-портал, обеспечивающий единую точку доступа к вулканологическим и сейсмологическим пространственным данным и сервисам Института вулканологии и сейсмологии ДВО РАН. Включает каталог метаданных для поиска данных, картографический сервер для публикации пространственных данных, картографические WMS и WFS сервисы визуализации пространственных данных. Предоставляет доступ к коллекциям данных — Web-ориентированным базам данных и информационным системам	5

Наименование	Краткое описание	Кат.
Репозиторий ИВиС ДВО РАН	Электронный архив документов научного и образовательного назначения, произведенных сотрудниками учреждений Камчатского научного центра ДВО РАН	3
Информационная система "Вулканы Курило-Камчатской островной дуги" (VOKKIA)	VOKKIA предназначена для систематизации и интеграции геологических, геофизических, геохимических и других научных данных по наземным и подводным вулканам Курило-Камчатской островной дуги и омывающих ее морей	5
Автоматизированная аналитическая система Группы реагирования на вулканические извержения KVERT	Исследования и мониторинг активности вулканов Камчатки и Северных Курил	5
Информационная издательская система ИВиС ДВО РАН	Предназначена для подготовки и публикации научных журналов ИВиС ДВО РАН: "Вулканология и сейсмология"; "Вестник КРАУНЦ. Серия: Науки о Земле" и материалов ежегодных конференций ИВиС ДВО РАН	3
Информационный сервис управления издательской деятельностью научной организации	Информационный сервис для управления издательской деятельностью научной организации	3
Автоматизированная система для ведения архивов научной организации	Информационная система для ведения архива научной организации	3
Медиаархив ДВО РАН	Система предназначена для трансляции и публикации медийных научных материалов в интересах организаций, подведомственных ФАНО России	4
Распределенные информационные сервисы для работы с данными дистанционного зондирования Земли (Центр коллективного пользования "ИКИ-Мониторинг" и сопряженные информационные системы) (инфраструктура на Дальнем Востоке и в Сибири)	Обеспечение доступа к большим, многолетним архивам спутниковых данных, различным информационным продуктам, получаемым на их основе, а также к вычислительным ресурсам, позволяющим проводить их анализ и обработку. Доступ обеспечивается через ИС "Вега-Science", а также посредством различных программных интерфейсов	5
АСОМБИ	Сбор и статистический анализ медико-биологической информации	1
Электронные каталоги научной библиотеки ИВиС ДВО РАН	Библиотечное сопровождение научной деятельности	3
Система видеоконференцсвязи ДВО РАН	Обеспечение коллективной работы ученых для проведения научных исследований посредством различных информационных сервисов (видео, презентации и т.п.)	4

**A. V. Yurchenko**, PhD, first deputy director, e-mail: yurchenko@ict.nsc.ru  
Institute of Computational Technologies of the Siberian Branch of the Russian Academy of Sciences  
(Institute of Computational Technologies SB RAS), Novosibirsk, Russia

## On the Approach Considering Scientific IT-Service as a Base Unit for Cyberinfrastructure of Science

*Purpose of the study is to consider some problems of developing the information, computing and telecom infrastructure for scientific research. Significance of such infrastructure is well understood, but in Russian Federation there is still no common approach and no institutions for its support. We are basing on the scientific IT-service as a key unit for accounting of usage of such infrastructure. We are describing notions for scientific cyberinfrastructure, scientific IT-service and their components and stating relations between them. We are considering peculiarities of such services classification and estimation of their cost. Examples of scientific IT-services provided by main academic IT-providers and available for FASO Russia's organizations are listed and parsed. Being decomposed any of these services leads to basic ones and that makes us to search for ways to assess their science intensity and resource intensity. Corresponding estimation for different service categories is done. Also, the cost of support for the part of considered cyberinfrastructure is compared to some commercial analogues. Finally, after showing the importance of the problem of scientific cyberinfrastructure support we are stating the problem of developing a common approach to that support using the IT-service as a basic unit and taking into account the experience gathered by the workgroup of the FASO's Commission on information for the assessment of applications for support of ITCI. Obtained results will be useful for developing the strategy of building the cyberinfrastructure of Russian science and for advancing the accounting technologies of relevance and usefulness of that infrastructure and its components.*

**Keywords:** cyberinfrastructure of science, scientific IT-service, science intensity and resource intensity of IT-service, relevance and usefulness of scientific IT-service, scientific IT-services classification

### References

1. **Microsoft** Azure for Research, available at: <https://www.microsoft.com/en-us/research/academic-program/microsoft-azure-for-research/> (accessed December 29, 2017).
2. **Press** Briefing by Richard Clarke, National Coordinator for Security, Infrastructure Protection and Counter-terrorism; and Jeffrey Hunker, Director of the Critical Infrastructure Assurance Office. News release. The White House Office of the Press Secretary, 1998, May 22, available at: <https://fas.org/irp/news/1998/05/980522-wh3.htm> (accessed December 29, 2017).
3. **Bottum J. R., Davis J. F., Siegel P. M., Wheeler B., Oblinger D. G.** Cyberinfrastructure: In Tune for the Future, *Educause Review*, 43 (4), Jul/Aug. 2008, available at: <https://web.archive.org/web/20080907155037/http://connect.educause.edu/Library/EDUCAUSE%2BReview/CyberinfrastructureInTune/46966> (accessed December 29, 2017).
4. **Newman H. B., Ellisman M. H., Orcutt J. A.** Data-Intensive e-Science Frontier Research in the Coming Decade, *Communications. Association for Computing Machinery*, 46 (11), Nov. 2003, CiteSeerX 10.1.1.72.5841, Freely accessible, DOI: 10.1145/948383.948411.
5. **Abramov S. M., Zadneprovskii V. F., Moskovskii A. A.** Otechestvennye Super-EVM i grid-sistemy. Problemy razvitiya natsional'noj kiberinfrastruktury v Rossii (Russian supercomputers and grid-systems. Problems of the national cyberinfrastructure in Russia), *Trudy XII-oy Nauchno-prakticheskoy konferentsii Universiteta goroda Pereslavlja* (Proc. of XII-th Scientific-practical conference at the Pereslavl' University), Pereslavl'-Zalesskii, 2008, pp. 1—27, available at: <http://www.botik.ru/PSI/RCMS/publications/publ-texts-2008/supercomputers-grid-systems.pdf> (accessed December 27, 2017) (in Russian).
6. **Shokin Yu. I., Fedoruk M. P., Chubarov D. L., Yurchenko A. V.** O razvitiy infrastrukturnykh superkomp'yuternykh i raspredelennykh vychisl'nykh v SO RAN (Development of the supercomputing and distributed computing infrastructure in the Siberian Branch of the Russian Academy of Sciences), *Informatsionnye Tekhnologii i Vychislitel'nye Sistemy* (Information Technologies and Computing Systems), 2011, no. 3, pp. 9—19.
7. **Order** of February 27, 2013 number 254-r of the Government of the Russian Federation "On awarding the prizes of the Government of the Russian Federation in 2012 in the field of science and technology", available at <http://government.ru/docs/643/> (accessed December 27, 2017, in Russian).
8. **Order** of February 26, 2015 number 303-r of the Government of the Russian Federation "On awarding the prizes of the Government of the Russian Federation in 2014 in the field of science and technology", available at <http://government.ru/docs/all/95036/> (accessed December 27, 2017, in Russian).
9. **Yurchenko A. V.** K kontseptsii informatsionno-analiticheskoy sistemy podderzhki nauchnykh issledovaniy, osnovannykh na intensivnom ispol'zovanii tsifrovyykh dannykh (On the concept of information-analytical system for supporting data intensive science), *Vychislitel'nye Tekhnologii* (Computational Technologies), 2017, vol. 22, no. 4, pp. 105—120.
10. **Nathan L. Yo., Stephen F. S., Pardis C. S.** Data sharing: Make outbreak research open access, *Nature*, 2015, vol. 518, iss. 7540, pp. 477—479.
11. **Lowenberg, D., Ross, A., Ganley, E.** Introducing the PLOS Open Data Collection, *PLOS Collections in Open Data*, 2016, Nov. 10.
12. **Haendel M., Vasilevsky N.** Data sharing in a modern world; well, maybe not so modern, redux, *PLOS Collections in Open Data*, 2017, Nov. 13.
13. **Yevshin I., Sharipov R., Valeev T., Kel A., Kolpakov F.** GTRD: a database of transcription factor binding sites identified by ChIP-seq experiments, *NUCLEIC ACIDS RESEARCH*, vol. 45, D1, 2017, pp. D61—D67, DOI: 10.1093/nar/gkw951