

SOFIA INITIATIVE
„MINERAL DIVERSITY PRESERVATION”

IX International Symposium
**MINERAL DIVERSITY
RESEARCH AND PRESERVATION**

СОФИЙСКАЯ ИНИЦИАТИВА
„СОХРАНЕНИЕ МИНЕРАЛЬНОГО РАЗНООБРАЗИЯ”

IX Международнѳй симпозиум
**МИНЕРАЛЬНОЕ РАЗНООБРАЗИЕ
ИССЛЕДОВАНИЕ И СОХРАНЕНИЕ**

EARTH AND MAN NATIONAL MUSEUM
4, Cherny vrѳh Blvd., 1421 Sofia, Bulgaria
16-18 OCTOBER 2017

НАЦИОНАЛЬНЫЙ МУЗЕЙ „ЗЕМЛЯ И ЛЮДИ”
бул. „Черни връх” 4, София 1421, България
16-18 ОКТЯБРЬ 2017

© Национальный музей „Земля и люди”, София, 2018
ISSN - 1313-9231

ДМИТРИЙ ВАРЛАМОВ, Т. ДОКИНА, Н. ДРОЖЖИНА,
О. САМОХВАЛОВА

*Институт экспериментальной минералогии Российской Академии Наук Россия,
Московская обл., г. Черноголовка, ул. Акад. Осипьяна, 4 dima@iem.ac.ru*

Abstract. Article sums up some results of development in 20 years of existence of the WWW-MINCRYST information-calculation system (ICS) intended for work with crystalline structures of minerals and their synthetic analogs and assesses the prospects for further development. The main components of the ICS are the database (more than 10,200 records for more than 4280 unique mineral phases), equipped with a complex of search and information selection tools, multimedia presentation facilities (interactive structures and spectra), spectral information processing capabilities, cross-reference generation tools.

Резюме. Статья подводит некоторые итоги развития за 20 лет существования информационно-вычислительной системы (ИВС) WWW-MINCRYST, предназначенной для работы с кристаллическими структурами минералов и их синтетических аналогов, и оценивает перспективы дальнейшего развития. Основными компонентами ИВС являются база данных (более 10200 записей для более чем 4280 уникальных фаз), снабженная комплексом средств поиска и выбора информации, средствами мультимедийного представления информации (интерактивные структуры и спектры), возможностями обработки спектральной информации, средствами формирования перекрестных ссылок.

Введение. Информационно-вычислительная система (ИВС) WWW-MINCRYST (<http://mincryst.iem.ac.ru>, <http://database.iem.ac.ru/mincryst>) исторически создавалась на базе локальной базы данных по структурам минералов и была задумана как инструмент для максимально возможного расширения числа пользователей последней и всемерного развития ее функциональности. ИВС WWW-MINCRYST изначально опиралась на использование Интернет-технологий и стала одним из пионерских интерактивных научных Интернет-ресурсов как в России, так и в мире (в области наук о Земле). Первый полностью работоспособный вариант WWW интерфейса был создан в рамках гранта РФФИ № 96-07-89162 и представлен пользователям в декабре 1997 года. ИВС WWW-MINCRYST стала одним из первых интерактивных Интернет-ресурсов в области наук о Земле в России и в мире (Chichagov, 2001), и с тех пор успешно функционирует, развивается и активно используется пользователями уже почти 20 лет. Основы идеологии и технологии ИВС описаны здесь (Чичагов, 2007; Варламов, 2013).

Цель создания ИВС WWW-MINCRYST и сопровождающих ее интерактивных Интернет-инструментов – обеспечить максимально широкий круг пользователей во всех областях науки, оперирующих с кристаллическим веществом (геология, геохимия, кристаллография, физика твердого тела, физика поверхности и т.д.) наиболее достоверной и актуальной информацией в области структурной кристаллохимии минералов, их синтетических аналогов и элементов, а также дать дополнительный толчок в развитии методов обработки и визуализации кристаллоструктурных данных .

В начале проекта ресурс был призван обеспечить интерактивный доступ Интернет-пользователям к накапливавшимся с 1985 года в рамках локальной базы данных МИНКРИСТ данным по кристаллическим структурам (прежде всего литературным, а также авторским аналитическим — всего около 3500 структур на момент старта проекта). В ходе проекта в состав ИВС вводились программные разработки по обработке кристаллоструктурных и кристаллохимических данных, их анализу и мультимедийной визуализации. Главным инициатором и основным идеологом работ был заведующий группой рентгено-спектрального анализа ИЭМ РАН выдающийся кристаллограф Анатолий Чичагов (увы, безвременно ушедший от нас в 2010 году).

За это время ИВС WWW-MINCRYST нарастил свой информационный фонд почти в 2,5 раза (почти на 7000 объектов) и по праву вошел в первые ряды рентгеноструктурных и кристаллохимических баз данных, связанных с изучением минерального вещества, как по

оценкам пользователей, так и по мнению составителей отраслевых каталогов информационных ссылок.

Основным компонентом ИВС WWW-MINCRYST является собственно база данных или Информационный фонд (на декабрь 2017 года — более 10200 информационных объектов для почти 4280 уникальных фаз, включая 4000 природных минералов и не менее 300 синтетических фаз, не имеющих пока натуральных аналогов).

Информационный фонд содержит информацию о большинстве минеральных видов (более 4000 из почти 5200 официально признанных на декабрь 2017 года International Mineralogy Association – http://nrmima.nrm.se//IMA_Master_List_%282017-11%29.pdf или <http://truff.info/ima>), кристаллические структуры которых расшифрованы к настоящему времени. Помимо природных объектов, в базе данных представлены синтетические минералы – их структурные аналоги, отличающиеся по составу (например, с заменой одного из катионов), и неорганические соединения (силикаты, фосфаты, бораты и др.), близкие по свойствам к природным веществам. Информационный фонд содержит данные структурных работ из более 140 иностранных и российских журналов за период от 1930-х годов по настоящее время. Ежегодный прирост новых или заново переопределенных/уточненных кристаллических структур минералов и их аналогов достаточно значителен, чтобы требовалась постоянная актуализация информационного фонда (в среднем, до 350-400 структур в год + 150-200 структур, подвергнутых ревизии и изменениям). В последние годы основной акцент был сделан на новые минералы.

Информационная запись для индивидуального кристаллического вещества содержит информацию о названии (в соответствии с классификацией International Mineralogy Association или рекомендациями по наименованию неорганических веществ IUPAC), химическом составе, симметрии, параметрах элементарной ячейки, координатах атомных позиций с изотропными температурными факторами и заселенностями, информацию о межплоскостных расстояниях, HKL-индексах и интенсивностях сильнейших рефлексов рентгенодифракционной картины поликристалл-фазы, а также ссылки на соответствующие публикации по расшифровке или уточнению кристаллической структуры. Запись может быть специфицирована по полезным свойствам, особенностям химического состава и структуры, P-T условиям синтеза, принадлежности к условным минеральным группам. Каждая запись содержит "монокристалльные" и "поликристалльные" характеристики кристаллической фазы. Минералы классифицированы в соответствии с таксонами структурно-химической систематики минералов А.А. Годовикова, кристаллохимической классификации М. Чириотти, классификации по структурным типам минералов по Г.Б. Бокию. Для 2500 фаз сделаны экспресс-оценки потенциальной энергии кристаллической решетки. В настоящее время создается система расчета содержаний элементов в соответствии с идеальной и реальной формулами минерала. В последние годы в записи вводились также географические привязки (при их наличии в статье) и прочие дополнительные сведения.

Веб-интерфейс ИВС WWW-МИНКРИСТ представляет собой многоуровневую информационную двуязычную (русский и английский, включая содержание и языки интерфейса) систему. В нее входят следующие компоненты:

а) средства поиска по названиям (и их фрагментам) минералов и содержанию их спецификаций, химическому составу (комбинации присутствующих/отсутствующих элементов, анионные и силикатные группы, устойчивые сочетания элементов), кристаллографическим параметрам, литературным источникам, вспомогательной информации, а также некоторым их сочетаниям. По ряду параметров поиска у ИВС пока нет аналогов;

б) мультимедийные интерактивные формы визуализации структур и спектров. Модуль **WWW-Crystpic** формирует динамические интерактивные изображения моделей кристаллических структур в шарах-сферах и в полиэдрических проекциях (до 138 позиций и до 1500(!) атомов на структуру). Модуль позволяет проводить всевозможные манипуляции с моделью структуры, включая масштабирование, непрерывное и/или автоматическое дискретное вращение вокруг "экранных" осей X,Y,Z, ориентацию по кристаллографическим

осям, hkl-фрагментацию структуры (на hkl-ориентированные фрагменты толщиной $d(hkl)$), наращивание элементарных ячеек вдоль любых выбранных направлений для формирования "сверхструктур" и мотивов, а также прямой "ручной" и автоматизированный для малых полиэдров (тетраэдров и октаэдров) расчет любых межатомных расстояний и углов в структуре. Программа изображает любые полиэдры, включая "дефектные" с необычно малыми ("плохими") межатомными расстояниями. Модуль **WWW-Mixipol** предназначен для графического представления полных расчетных спектральных профилей поликристалл-рентгенограмм с возможностями манипулирования спектрами для разных источников излучения и разных типов спектральных шкал. Также модуль способен формировать рентгенограммы *смесей* фаз (до 6 фаз) при возможности варьирования содержаниями компонентов смеси. Предусмотрены упрощенные варианты представления информации в виде традиционных шаровых структур и линейчатых спектров;

в) классификационные схемы (Годовиков, Бокий, Chiriotti);

г) системы динамически формируемых WWW-ссылок на внешние информационные ресурсы (Mindat, Webmineral и др., поисковые системы, Wiki-системы);

д) инструментарий разработчика (включая систему импорта, проверки и редакции данных, а также архивации и резервного копирования).

В теоретической основе ИВС WWW-MINCRYST был положен ряд принципиальных на тот момент "идеологических" моментов. Синтез информации о кристаллической фазе, рассматриваемой как "монокристалл" и/или "поликристалл" с заменой экспериментальных поликристалл-стандартов расчетными – базовая идея WWW-MINCRYST, существенно расширяющая информационную базу сравнительно с чисто аналитической. На базе введенной информации программный пакет эксперта-разработчика позволяет автоматически сформировать вторую, производную от первой (аналитической), базу *расчетных* поликристалл-стандартов, проводя синтез двух типов информации о кристаллической фазе. Связка "Кристаллическая структура фазы и ее *расчетная* поликристалл-рентгенограмма" является информационной основой ИВС WWW-MINCRYST и служит важнейшим инструментарием в руках пользователя. Стержневая идея WWW-MINCRYST о синтезе двух типов кристаллоструктурной информации о кристаллической фазе реализует принципиально новый подход к формированию всей кристаллоструктурной информации о веществе и организации доступа к ней.

Основным источником информации для ИВС служат журнальные статьи. Извлеченная кристаллоструктурная информация помещается (в соответствии со специальным форматом записи) в ASCII-файл с последующей программной экспертизой (проводимой на локальных ПК) по результатам расчета межатомных расстояний и других кристаллоструктурных характеристик и, в случае положительного решения, импортируется в ИВС специальными средствами одиночного или группового импорта.

В ИВС WWW-MINCRYST реализован импорт одобренной экспертами информации через web-интерфейс как в виде единичных записей, так и пакетов записей (до нескольких сот) с входным контролем данных, что позволяет проводить постоянную актуализацию информационного фонда. Также реализована возможность редактирования записей в on-line режиме, их удаления, замены служебных файлов записей, архивации и восстановления базы данных через web-интерфейс.

В ИВС реализована система поиска фаз по отдельным наборам или комплексу критериев: название минерала (полное или частичное, по спецификации, в составе общеупотребительных групп типа цеолитов, алфавитные списки), химический или элементный состав в различных комбинациях (присутствие/отсутствие элементов и их комбинации), кристаллоструктурные характеристики (симметрия, пространственные группы, параметры элементарной ячейки, межплоскостные расстояния $d(hkl)$), что в сочетании с химическим (элементным) составом дает возможность прямого интерактивного качественного рентгенофазового анализа. В систему поиска добавлен поиск по классификационным параметрам нескольких кристаллохимических и структурных классификаций (Годовиков,

Чириотти, Бокий). Система поиска по этим параметрам обеспечивает поиск (и группировку) минералов по указанным низшим и промежуточным таксонам классификаций. По ряду параметров поиска ИВС WWW-MINCRYST до сих пор не имеет аналогов среди минералого-кристаллографических баз данных.

В 2017 году в систему поиска по химическому составу сделаны важные дополнения для поиска по устойчивым сочетаниям элементов (типа анионов/катионов: SO_4 , CO_3 ; нейтральных групп – H_2O , NH_3 ; структурных силикатных мотивов – SiO_3 , Si_2O_7 , Si_3O_8 и т.п.), реализована возможность поиска по примесным изоморфным атомам, не отраженным в идеализированных формулах минералов (что важно для рассеянных элементов); улучшены схемы поиска по взаимным критериям «отсутствие-присутствие» элементов. Закончен переход на унифицированную кодировку utf-8 (для внутреннего представления текстовой информации в базе данных и выдачи пользователю), что упрощает работу с современными браузерами и позволяет использовать диакритические знаки в написании оригинальных имен минералов. В модуле WWW-Crystpic исправлен ряд неточностей и замечаний по работе с размножением сверхструктурных элементов, расширены возможности работы с большими (по количеству атомов, до 200, например, в сложных структурах цеолитах) структурами.

По результатам поиска пользователь получает ссылку на список фаз, отвечающих поисковым критериям, с указанием ID записи, названия фазы, формулы и пространственной группы с возможностью последующего перехода по гиперссылке. В случае единичного успешного ответа пользователь сразу автоматически переводится на найденный объект, если найдена группа объектов – предлагается список выбора.

В «карте» найденного объекта представлены основные данные (как хранимые в БД, так и динамически рассчитываемые при формировании «карты») – название, формула, параметры ячейки, различные кристаллографические данные (кол-во рефлексов, рентгеновские плотности, коэффициенты поглощения); CPDS карта – 20 максимальных рефлексов, их hkl позиции и интенсивности; базовые атомные позиции и их заселенности; полная информационно-расчетная карта (рентгеновские плотности, все рефлексы и т.п.); рассчитанные энергии решетки; принадлежность к различным классификационным схемам; источники данных (ссылки на публикации, поправки и т.п.); автоматически формируемые ссылки на внешние ресурсы.

Для всех записей через апплет WWW-Crystpic доступны динамически создаваемые интерактивные изображения моделей кристаллических структур в шарах-сферах и в полиэдрических проекциях (до 138 позиций и до 1500(!) атомов на структуру). Программа позволяет делать всевозможные манипуляции с моделью структуры, включая масштабирование, непрерывное и/или автоматическое дискретное вращение вокруг "экранных" осей X,Y,Z; ориентацию по кристаллографическим осям, hkl-фрагментацию структуры (на hkl-ориентированные фрагменты толщиной $d(hkl)$), наращивание элементарных ячеек вдоль любых выбранных направлений для формирования "сверхструктур" и мотивов, а также прямой "ручной" и автоматизированный для малых полиэдров (тетраэдров и октаэдров) расчет любых межатомных расстояний и углов (плоских и телесных) в структуре.

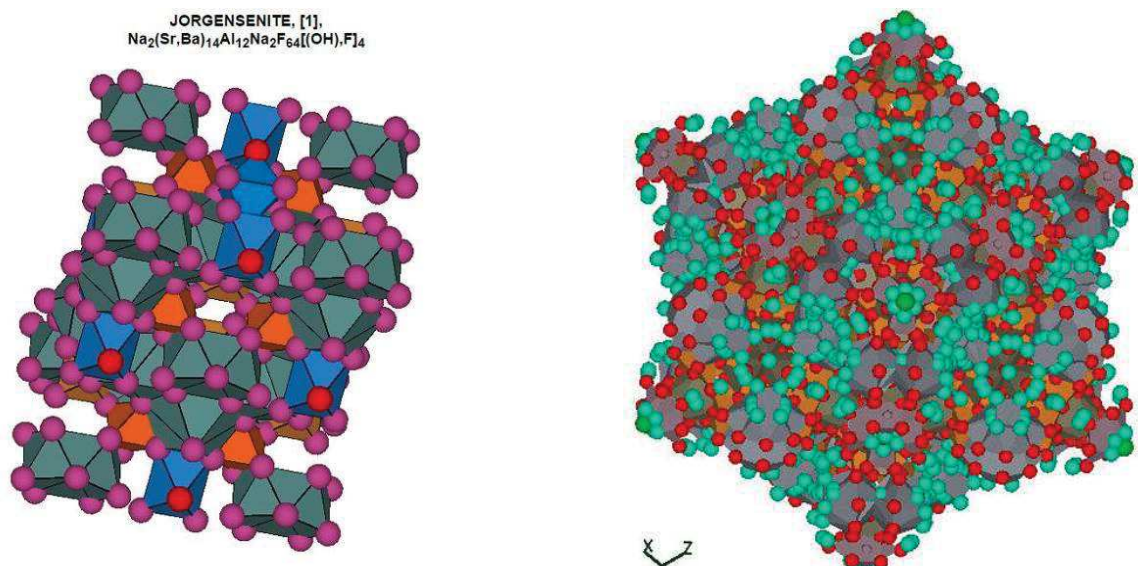


Рис. 1. Примеры интерактивных моделей кристаллической структуры - смешанное представление в шарах и полиэдрах (справа цеолит, размеры атомов уменьшены вдвое)

Программа изображает любые полиэдры, включая "дефектные" с необычно малыми ("плохими") межатомными расстояниями. Примеры моделей кристаллических структур (до 600 атомов) приведены на рис.1.

В настоящее время существуют проблемы с работой с новыми версиями WWW-браузеров из-за их параноидальных настроек безопасности для языка Java, поэтому начат перенос модулей на новые программные платформы типа встроенных средств HTML5 или WebGL. Модуль в виде апплета WWW-Mixipol предназначен для графического представления полных расчетных спектральных профилей поликристалл-рентгенограмм с возможностями манипулирования спектрами для разных источников излучения и разных типов спектральных шкал.

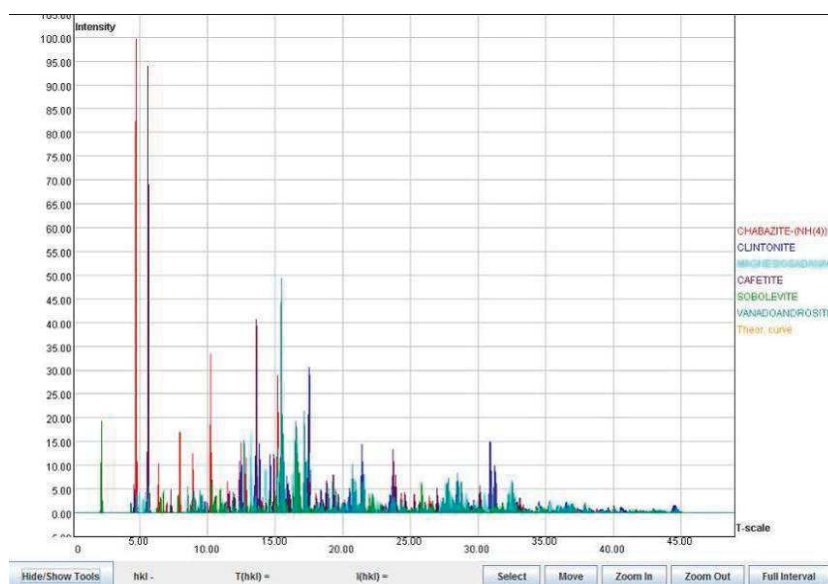


Рис. 2. Пример комбинации расчетных спектральных профилей поликристалл-рентгенограмм (смесь 6 фаз)

Также модуль способен формировать рентгенограммы смесей фаз (до 6 фаз одновременно) при возможности варьирования относительными содержаниями компонентов смеси. Пример расчетных спектров смеси шести фаз приведен на рис.2.

Как для структур, так и для спектров минералов предусмотрены «упрощенные» варианты представления в виде интерактивных традиционных шаровых структур и линейчатых спектров.

Одной из первых среди научных интернет-ориентированных баз данных для ИВС WWW-MINCRYST была разработана система динамически формируемых перекрестных веб-ссылок для связи объектов с записями для конкретных минералов в ведущих минералогических базах данных в Интернете. Система генерации динамических гиперссылок на внешние информационные ресурсы (на минералогические базы данных и поисковые системы) позволяет "прозрачно" для пользователя подключать большие внешние массивы данных, используя метод "генеральных" запросов (Плечов, 2002; Варламов 2016). При этом пользователь сразу получает доступ к информации по интересующему его объекту, минуя стадии поиска или просмотра всей внешней базы. Кроме того, данный механизм реализует обратную связь, позволяя таким же образом ссылаться этим базам уже на наши информационные объекты, что резко повышает востребованность ИВС внешними пользователями.

Наличие в ИВС WWW-MINCRYST данных по 10200 кристаллическим структурам, встроенного расчетного комплекса и развитых средств визуализации позволили помимо возможностей, ориентированных на поиск и предоставление информации, использовать ИВС в разработке нетрадиционных научных подходов к интерпретации и представлению некоторых кристаллических структур. Обработка комплексными SQL запросами массива накопленных данных позволил получить весьма нетривиальные статистические выборки, например, по распределению в природе минералов среди групп симметрии кристаллов. WWW Xгаурol позволил выявить в традиционных структурах возможность гибкого использования полиэдров и почти автоматически формировать различные варианты структурных моделей минералов. Благодаря гибкому использованию полиэдров, позволяющему формировать различные варианты структурных моделей минералов, для части минералов можно нестрого привязываться к традиционному катионно-анионному изображению, а формировать структуры на основе любых атомов, входящих в ее состав. Используется принцип: в ряде случаев кристаллическое пространство можно организовать в смешанном шаровом и полиэдрическом изображении на основе любых (а не общепринятых) атомов в структуре. Метод особенно эффективен для сложных "неправильных" бескислородных структур (фосфиды, сложные сульфиды и др.).

Востребованность WWW-MINCRYST хорошо подтверждается статистикой обращений (за 2017 год – 9 млн. успешных единичных запросов, более 109 Гб скачанной информации, около 48000 уникальных сайтов-клиентов), а также большим количеством отзывов, описаний и внешних ссылок на WWW-MINCRYST (см. раздел "Ссылки" на сайте).

О перспективах развития WWW-MINCRYST следует сказать отдельно. В связи с очень быстрым развитием web-технологий и (соответственно) браузеров, не ослабевают потребности в кардинальной переработке нынешнего клиентского интерфейса, во многом уже не отвечающего как новым технологиям и зачастую неполностью совместимого с новым клиентским ПО (особенно в области web-security и мультимедиа-технологий), а также возросшим требованиям пользователей и возможностям конкурирующих баз данных. Также будут изменены в сторону повышения дружелюбности и простоты дизайн и элементы управления предоставляемой пользователю информации. Предусмотрено пополнение базы данных дополнительной сугубо геолого-минералогической информацией (реальные составы минералов, фото минералов, наиболее близких по форме идеальным, рисунки кристаллографических форм, геоинформационные данные – места находок и др.). Предполагается усовершенствование интерфейса разработчиков для ввода, проверки и редактирования исходных данных, что значительно облегчит пополнение и ревизию данных. Также конструируется единый поисковый интерфейс в отличие от ныне разобщенных поисковых страниц (химический состав, кристаллографические параметры, имена минералов и т.п.), с возможностью создания комплексных запросов с включением /исключением поисковых критериев, а также потенциально новых типов поисковых запросов.

ИВС WWW-MINCRYST является общедоступным, дружелюбным пользователю интерфейсом к большому массиву кристаллоструктурной и кристаллохимической информации с развитыми средствами поиска, представления и обработки и может служить мощным инструментарием для всех исследователей в минералогии, кристаллографии, физике твердого тела, материаловедении и прочих смежных областях науки.

Проведение работ по ИВС WWW-MINCRYST в течение 1997-2017 годов было поддержано 6-ю грантами РФФИ (в том числе в настоящее время – грантом РФФИ № 15-07-08399-а, рук. Варламов Д.А.).

ЛИТЕРАТУРА

- Варламов Д.А., Докина Т.Н., Дрожжина Н.А., Самохвалова О.Л.** WWW-MINCRYST: Интернет-ориентированная информационно-вычислительная система по кристаллографии и кристаллохимии минералов // Вестник ЮУрГУ, Сер. «Вычислительная математика и информатика», 2013, т.2, вып.1, с.26-32
- Варламов Д.А., Докина Т.Н., Дрожжина Н.А., Самохвалова О.Л.** Мультимедийные Web-ориентированные базы данных в науках о Земле на примере ИВС WWW-MINCRYST (кристаллохимия и минералогия) // Научный сервис в сети Интернет: труды XVIII Всероссийской научной конференции (19-24 сентября 2016 г., г.Новороссийск). М.: ИПМ им.М.В.Келдыша, 2016. С.104-110 <http://keldysh.ru/abrau/2016/29.pdf>
- Плечов П.Ю., Варламов Д.А., Трусов С.В.** Типы информационных потоков в области наук о Земле // Научный сервис в сети Интернет, изд-во МГУ, 2002, с.142-144.
- Чичагов А.В., Варламов Д.А., Ершов Е.В., Докина Т.Н., Дрожжина Н.А., Самохвалова О.Л.** Кристаллографическая и кристаллохимическая база данных для минералов и их структурных аналогов (WWW-MINCRYST) // Записки РМО, 2007, т.136, № 3, с.135-141
- Chichagov A.V., Varlamov D.A., Dilanyan R.A., Dokina T.N., Drozhzhina N.A., Samokhvalova O.L., and Ushakovskaya T.V.** MINCRYST: a Crystallographic Database for Minerals, Local and Network (WWW) Versions // Crystallography Reports, 2001, Volume 46, Issue 5, pp.876-879