

WWW-MINCRYST: ИНТЕРНЕТ-ОРИЕНТИРОВАННАЯ ИНФОРМАЦИОННО-ВЫЧИСЛИТЕЛЬНАЯ СИСТЕМА ПО КРИСТАЛЛОГРАФИИ И КРИСТАЛЛОХИМИИ МИНЕРАЛОВ¹

Д.А. Варламов, Т.Н. Докина, Н.А. Дрожжина, О.Л. Самохвалова

Описывается интернет-ориентированная информационно-вычислительная система (ИВС) WWW-MINCRYST, предназначенная для работы с кристаллическими структурами минералов, их синтетических аналогов и элементов. Основными компонентами ИВС являются собственно база данных (более 8000 записей для 3000 уникальных фаз), снабженная комплексом средств поиска и выбора информации, средствами мультимедийного представления информации (полиэдрические и шаровые интерактивные структуры, спектры и т.п.), возможностями обработки спектральной информации, в том числе с использованием пользовательских данных. ИВС размещена по адресу <http://mincryst.iem.ac.ru> и доступна пользователям без ограничений.

Ключевые слова: интернет-ориентированные базы данных, кристаллохимия, кристаллические структуры, минералы, PHP, MySQL.

Введение

WWW-MINCRYST (<http://mincryst.iem.ac.ru>) в качестве Интернет-ориентированной базы данных возник достаточно давно и стал одним из пионерских для РФ научных интерактивных Интернет-ресурсов (первый полностью работоспособный вариант WWW интерфейса был представлен пользователям уже в декабре 1997 года [1]). Ресурс призван обеспечить интерактивный доступ Интернет-пользователям к накапливавшимся с 1985 года данным по кристаллическим структурам (прежде всего литературным, а также авторским аналитическим) и многочисленным разработкам авторов проекта по обработке этих данных и анализу. Инициатором и основным идеологом работ стал заведующий группой рентгеноспектрального анализа выдающийся кристаллограф А.В. Чичагов (увы, ушедший от нас в 2010 году). Методология создания и технические детали текущей реализации информационно-вычислительной системы (ИВС) уже освещались ранее [2, 3] и здесь описана кратко, также освещены текущее состояние ИВС и перспективы ее развития.

В настоящее время ИВС WWW-MINCRYST (по оценкам пользователей и составителей каталогов информационных ссылок) входит в первые ряды рентгеноструктурных баз данных, связанных с изучением минерального вещества.

Созданная ранее и развиваемая ИВС с ее инструментарием во многом решает важную задачу информационного обеспечения исследований во всех областях науки, оперирующих с кристаллическим веществом (минералогия, кристаллография, физика твердого тела и т.п.), обеспечивая как поддержку многочисленных исследователей минерального

¹ Статья рекомендована к публикации программным комитетом Международной суперкомпьютерной конференции «Научный сервис в сети Интернет: поиск новых решений – 2012».

вещества и синтетических аналогов (включая студентов и аспирантов), так и давая дополнительный толчок в изучении кристаллографических данных.

WWW-MINCRYST зарегистрирован в Государственном Регистре Баз Данных (НТИЦ «Информрегистр») под номером 0229805169 (Регистрационное Свидетельство № 4873 от 11.02.99).

1. Стрoение и идеология ИВС WWW-MINCRYST

В основе WWW-MINCRYST лежит ряд принципиальных «идеологических» моментов, реализация которых уже обеспечила ему достаточное признание мировой кристаллографической и минералогической общественности.

Синтез информации о кристаллической фазе, рассматриваемой как «монокристалл» и/или как «поликристалл» с заменой экспериментальных поликристалл-стандартов расчетными – стержневая идея WWW-MINCRYST, которая и делает его универсальной, комплексной и оригинальной разработкой.

В настоящее время информационный фонд WWW-MINCRYST насчитывает чуть более 8000 записей для более чем 3000 уникальных фаз (минералов, их аналогов, элементов), т.е. большинства минералов (из 4800), 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 т.п.), близкие по свойствам к природным веществам. Информационный фонд содержит данные структурных работ из более 120 иностранных и отечественных журналов за период от 1930-х до 2012 года. Ежегодный прирост новых или заново переопределенных кристаллических структур минералов и их аналогов достаточно значителен, чтобы требовалась постоянная актуализация информационного фонда (в среднем, до 300-400 структур в год). В последние годы основной акцент сделан на новые минералы, а также на кристаллические структуры, описанные в советских и российских журналах (которые в последние десятилетия были менее доступны web-пользователю, нежели зарубежные). Суммарный объем базы составляет около 500 Мб.

Технологически WWW-MINCRYST как сервис реализован на «классической» связке Linux-Apache-MySQL-PHP (LAMP) с использованием JavaScript и (для интерактивных апплетов) Java (на базе Sun/Oracle JDK), реализованной на сервере баз данных ИЭМ РАН (<http://database.iem.ac.ru>). WWW-MINCRYST был (в 1997 году) одной из первых интернет-ориентированных баз данных в России, использовавшим данную технологию, причем за 14 лет она нисколько не потеряла свою технологичность на стороне сервера. ИВС является двуязычной, включая содержание и языки интерфейса: русский и английский. В систему входят следующие компоненты: (а) собственно база данных; (б) комплексные поисковые интерфейсы, (в) мультимедийные интерактивные формы представления структурной и спектральной информации (через Java-апплеты); (г) классификационные схемы; (д) системы динамически формируемых ссылок на внешние информационные ресурсы; (е) WWW-ориентированный инструментарий разработчика; (ж) прикладные программы по обработке кристаллографической информации (WWW-GrayPol и WWW-MixiPol).

Основной источник информации для Базы данных по кристаллическим структурам – оригинальные журнальные статьи, опубликованные в открытой печати. Извлеченная кристаллоструктурная информация помещается (по специальному формату) в ASCII-файл с

последующей программой (на ПК) экспертизой по результатам расчета межатомных расстояний и других кристаллоструктурных характеристик и, в случае положительного решения, импортируется в ИВС специальными программами группового или одиночного импорта. Возможен импорт кристаллоструктурной информации через общепринятый в кристаллографии специализированный файл (Crystallographic Information File, CIF), содержащий кристаллоструктурные характеристики расшифрованной кристаллической фазы (формат – <http://www.sdsc.edu/pb/cif/cif.html>).

Базовая 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, НКЛ-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, а также по РТ-условиям синтеза. 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, в настоящее время вводится классификация по структурным типам минералов (Бокий Г.Б.). Для 2000 фаз сделаны экспресс-оценки потенциальной энергии кристаллической решетки.

На базе вводимой информации локальный программный пакет эксперта позволяет автоматически сформировать вторую, производную от первой, базу расчетных поликристалл-стандартов, проводя синтез двух типов информации о кристаллической фазе. Связка «Кристаллическая структура фазы и ее расчетная поликристалл-рентгенограмма» является не только информационной основой WWW-MINCRYST, но и служит важнейшим инструментарием в руках пользователя (в особенности рентгеновского кристаллографа). 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.

Для 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, что позволяет проводить постоянную актуализацию информационного фонда. Также реализована возможность on-line редакции записей, их удаления, замены служебных файлов записей, архивации и восстановления базы данных через web-интерфейс.

2. Основные возможности ИВС WWW-MINCRYST

В ИВС реализована эффективная система поиска фаз по комплексу параметров, включая название минерала (полное или частичное), химический, или элементный, состав в различных комбинациях (присутствие или отсутствие элементов и их комбинации), симметрию, кристаллоструктурные характеристики, а также 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, Чирiotти, Бокий). Система поиска по этим параметрам обеспечивает поиск (и группировку) минералов по указанным параметрам – низшим таксонам. По ряду параметров поиска WWW-MINCRYST до сих пор не имеет аналогов среди минералого-кристаллографических баз данных.

С помощью интегрированных в WWW-MINCRYST расчетных и презентационных модулей (в виде Java апплетов) последний превратился из первичной информационно-справочной в информационно-вычислительную систему, существенно расширив возможности пользователя по получению и представлению кристаллоструктурной и кристаллохимической информации о минеральном веществе. Разработанный и интегрированный (частично на конец 2012 года) в WWW-MINCRYST расчетный модуль WWW-Хгаурол в связках с модулями WWW-Crystpic и WWW-Mixipol позволяет квалифицированному пользователю модифицировать структуры из информационного фонда (не затрагивая его), а затем рассчитывать и получать графические изображения моделей кристаллических структур и квазиреальных полных профилей расчетных поликристалл-рентгенограмм на основе временно извлекаемых и модифицируемых базовых BDM-файлов и тем самым решая задачи по модификации реальных структур.

Наличие большого количества структур и хороших средств визуализации позволяет с помощью WWW-MINCRYST развивать принципиально новые подходы к представлению кристаллических структур – благодаря гибкому использованию полиэдров, позволяющему формировать различные варианты структурных моделей минералов. Используется принцип: в ряде случаев кристаллическое пространство можно организовать в смешанном шаровом и полиэдрическом изображении на основе любых атомов в структуре, не строго привязываясь к традиционному катионно-анионному изображению.

Для всех записей через встроенный апплет WWW-Crystpic (Java-3D) доступны динамически создаваемые интерактивные изображения моделей кристаллических структур в шарах-сферах и в полиэдрических проекциях (до 138 позиций и до 1500(!) атомов на структуру) в соответствии с основными канонами минералогии и кристаллографии. Программа позволяет делать всевозможные манипуляции с моделью структуры, включая масштабирование, непрерывное и/или автоматическое дискретное вращение вокруг «экраных» 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 («плотными») межатомными расстояниями. Более детальное описание всех возможностей апплета приведено здесь: <http://mincryst.iem.ac.ru/rus/crystpic.php>). В настоящее время из-за проблем с работой с библиотеками OpenGL (на которых реализован апплет) в составе новых версий Java рассматривается возможность переноса апплета на новые программные платформы типа встроенных средств HTML5

Также в WWW-MINCRYST интегрирован апплет WWW-Mixipol, предназначенный 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 одновременно) при возможности варьирования относительными содержаниями компонентов смеси.

Как для структур, так и для спектров минералов предусмотрены упрощенные варианты представления (для старых браузеров и маломощных персоналок) в виде традиционных шаровых структур и линейчатых спектров.

В состав WWW-MINCRYST встроен модуль Xгаурol, который в связках с программами WWW-Crystpic и WWW-Mixipol, позволяет квалифицированному пользователю менять «на лету» ряд параметров структуры и рассчитывать новые варианты изображений моделей кристаллических структур и квазиреальных полных профилей расчетных поликристалл-рентгенограмм на основе временно извлекаемых и модифицируемых базовых BDM-файлов, тем самым решая конкретные задачи по модификации реальных структур, но не затрагивая основной фонд WWW-MINCRYST.

Одними из первых среди научных интернет-ориентированных баз данных была разработана система динамически формируемых перекрестных веб-ссылок для связи записей с записями для конкретных минералов в ведущих минералогических базах данных, размещенных в Интернете. Система генерации динамических гиперссылок на внешние информационные ресурсы (в основном, на минералогические базы данных и поисковые системы) позволяет «прозрачно» для пользователя подключать большие внешние массивы данных, используя метод «генеральных» запросов. 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 подобным же образом ссылаться этим базам уже на наши информационные объекты, что резко повышает востребованность WWW-MINCRYST внешними пользователями.

Наличие в WWW-MINCRYST более 8 тысяч кристаллических структур и встроенный универсальный расчетный комплекс позволили помимо ориентированных на поиск и предоставление информации возможностей использовать ИВС в разработке нетрадиционных научных подходов к интерпретации и представлению некоторых кристаллических структур. WWW Xгаурol позволил, например, выявить в традиционных структурах возможность гибкого использования полиэдров и почти автоматически формировать различные варианты структурных моделей минералов. Как выяснилось, для части минералов кристаллическое пространство можно не строго привязываться к традиционному катионно-анионному изображению, а формировать структуры на основе любых атомов, входящих в ее состав. Метод особенно эффективен для сложных «неправильных» бескислородных структур (например, фосфиды, сложные сульфиды и др.).

Востребованность WWW-MINCRYST хорошо подтверждается статистикой обращений (7.5 млн. успешных единичных запросов за 2012 год, более 60 Гб скачанной информации, около 33000 уникальных сайтов), 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, возникла необходимость кардинальной переработки нынешнего клиентского интерфейса, уже не отвечающего как новым технологиям (и зачастую несовместимого с новым ПО), так и требованиям пользователей. Она будет включать изменение способа представления записи путем сведения разобщенных сейчас полей (основные данные, CPDS карта, атомные позиции, структура, спектры, ссылки и др.) в единое информационное пространство, формируемое на базе динамического HTML (с использованием технологий Jason, DHTML, HTML-5, оверлейных структур). Предусматривается как переделка и рекомпиляция Java

апплетов (в связи с устареванием и неполной совместимостью кода с последними реализациями Java VM), так и их замена на модули, использующие мультимедийные технологии стандартов HTML-5 или Adobe Flash (с использованием оригинального авторского расчетного кода). Также будут изменены в сторону повышения дружелюбности и простоты) дизайн и элементы управления предоставляемой пользователю информации. В ходе проекта предусмотрено пополнение базы данных дополнительной сугубо минералогической информацией (фото минералов и рисунки кристаллографических форм и др.). Несмотря на то, что по ряду параметров поиска WWW-MINCRYST до сих пор не имеет аналогов, будет расширен круг потенциальных поисковых запросов (прежде всего в области поисков по составу и кристаллографическим данным).

Заключение

Разработанная и развиваемая на протяжении 15 лет ИВС WWW-MINCRYST является общедоступным, дружелюбным пользователю интерфейсом к большому объему 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 1997–2012 годов было поддержано несколькими грантами РФФИ (в том числе в настоящее время – грантом РФФИ 12-07-00742-а, рук. Варламов Д.А.).

Литература

1. МИНКРИСТ – кристаллографическая база данных для минералов: локальный и сетевой (WWW) варианты / А.В. Чичагов, Д.А. Варламов, Р.А. Диланян и др. // Кристаллография. – 2001. – Т. 46. – № 5. – С.950–954.
2. Кристаллографическая и кристаллохимическая база данных для минералов и их структурных аналогов (WWW-MINCRYST) / А.В. Чичагов, Д.А. Варламов, Е.В. Ершов и др. // Записки Российского минералогического общества. – 2007. – Т. 136. – № 3. – С.135–141.
3. Чичагов, А.В. WWW-MINCRYST-2007 – Интернет-ориентированная база данных по кристаллографии/кристаллохимии минералов и их аналогов / А.В. Чичагов, Д.А. Варламов // «Научный сервис в сети Интернет: технологии параллельного программирования. 15 лет РФФИ», Труды Всероссийской научной конференции. – М.: изд-во МГУ, 2007. – С.390–392.

Варламов Дмитрий Анатольевич, с.н.с., Институт экспериментальной минералогии РАН, dima@iem.ac.ru

Докина Татьяна Николаевна, инженер-исследователь, Институт экспериментальной минералогии РАН, tdokina@mail.ru

Дрожжина Наталья Алексеевна, инженер-исследователь, Институт экспериментальной минералогии РАН, nadron@mail.ru

Самохвалова Ольга Леонидовна, инженер-исследователь, Институт экспериментальной минералогии РАН, olsamoh@mail.ru

WWW-MINCRYST: THE INTERNET-ORIENTED INFORMATION AND CALCULATION SYSTEM ON THE CRYSTALLOGRAPHY AND THE CRYSTAL CHEMISTRY OF MINERALS

D.A. Varlamov, Institute of Experimental Mineralogy RAS (Chernogolovka, Russian Federation),

T.N. Dokina, Institute of Experimental Mineralogy RAS (Chernogolovka, Russian Federation),

N.A. Drozhzhina, Institute of Experimental Mineralogy RAS (Chernogolovka, Russian Federation),

O.L. Samokhvalova, Institute of Experimental Mineralogy RAS (Chernogolovka, Russian Federation)

The Internet-oriented WWW-MINCRYST information-calculation system (ICS) intended for work with crystal structures of minerals, their synthetic analogs and elements is described. The ICS main components are actually a database (more than 8000 records for 3000 unique phases), supplied with a complex of means of search and information choice, means of multimedia submission of information (polyhedral and spherical interactive crystal structures, spectrum, etc.), opportunities of processing of spectral information, including using of the loaded user data. ICS is placed to the address <http://mincryst.iem.ac.ru> and is available to users without restrictions.

Keywords: Internet-oriented databases, crystal chemistry, crystal structures, minerals, PHP, MySQL

References

1. 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*, Vol. 46, Issue 5, P. 876–879.
2. Chichagov A.V., Varlamov D.A., Ershov Ye.V., Dokina T.N., Drozhzhina N.A., Samokhvalova O.L. Kristallograficheskaja i kristallohimicheskaja baza dannyh dlja mineralov i ih strukturnyh analogov (WWW-MINCRYST) [Crystallographic and crystal-chemical database for minerals and their structural analogues (WWW-MINCRYST)]. *Zapiski Rossijskogo mineralogicheskogo obshhestva* [Proceedings of the Russian Mineralogical Society]. 2007. Vol. 136. № 3. P. 135–141.
3. Chichagov A.V., Varlamov D.A. WWW-MINCRYST-2007 – Internet-orientirovannaya baza dannyh po kristallografii i kristallokhimii mineralov i ih analogov [WWW-MINCRYST-2007 – The Internet-oriented database on a crystallography and crystal chemistry of minerals and their analogs]. *Trudy vserossiiskoi konferentsii «Nauchnyi servis v seti Internet: tekhnologii paralelnogo programmirovaniya, 15 let RFFI» (Novorossiisk, sentyabr 2007)* [«Scientific service in Internet: technologies of parallel programming»: Proceedings of the Russian Scientific Conference (Novorossiysk, September 2007)]. Moscow. Moscow University Press, 2007. P. 390–392.

Поступила в редакцию 15 января 2013 г.