

# ИНФОРМАЦИОННО-ВЫЧИСЛИТЕЛЬНАЯ СИСТЕМА ПО КРИСТАЛЛОГРАФИИ И МИНЕРАЛОГИИ WWW-MINCRYST: 15 ЛЕТ ЭКСПЛУАТАЦИИ, ВОЗМОЖНОСТИ, ПЕРСПЕКТИВЫ

Д.А. Варламов

Ниже 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 (около 8500 записей для более чем 3000 уникальных фаз), снабженная комплексом средств поиска и выбора информации, средствами мультимедийного представления информации (полиэдрические и шаровые интерактивные структуры, спектры и др.), возможностями обработки спектральной и структурной информации, в том числе с использованием пользовательских данных.

ИВС WWW-MINCRYST (активные ссылки: <http://mincryst.iem.ac.ru>, <http://database.iem.ac.ru/mincryst>) в качестве Интернет-ориентированной базы данных возник достаточно давно и стал одним из пионерских научных интерактивных Интернет-ресурсов как для России, так и в области наук о Земле в мире. Первый полностью работоспособный вариант WWW интерфейса был представлен пользователям уже в декабре 1997 года [1], таким образом система успешно функционирует и востребована пользователями уже более 15 лет. На старте проекта ресурс был призван обеспечить интерактивный доступ Интернет-пользователям к накапливавшимся до этого с 1985 года данным по кристаллическим структурам (прежде всего литературным, а также авторским аналитическим), затем в состав ИВС были введены программные разработки авторов проекта по обработке этих данных и их анализу. Главным инициатором и основным идеологом работ стал заведующий группой рентгеноспектрального анализа Института экспериментальной минералогии РАН выдающийся кристаллограф А.В. Чичагов (увы, безвременно ушедший от нас в 2010 году).

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

Созданная 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 др.), обеспечивая как поддержку многочисленных исследователей (включая студентов и аспирантов) минерального вещества и синтетических аналогов, так и давая дополнительный толчок в изучении кристаллографических данных.

WWW-MINCRYST зарегистрирован в государственном регистре баз данных (НТЦ "Информрегистр") под номером 0229805169 (регистрационное свидетельство № 4873 от 11.02.99).

Сейчас WWW-МИНКРИСТ представляет собой многоуровневую информационную двуязычную (русский и английский, включая содержание и языки интерфейса) систему. В систему входят следующие компоненты: (а) комплексные поисковые интерфейсы, использующие как критерии поиска (в различных комбинациях) названия минералов, химический состав (в различных комбинациях), кристаллографические параметры, литературные ссылки и вспомогательную информацию, причем возможны комбинации поисковых параметров; (б) мультимедийные интерактивные формы представления структурной и спектральной информации (через Java-апплеты); (в) классификационные схемы; (г) системы динамически формируемых ссылок на внешние информационные ресурсы; (д) WWW-ориентированный инструментарий разработчика (включая систему импорта и проверки данных, а также их архивации); (е) прикладные программы по обработке кристаллографической информации (WWW-CrystPic и WWW-MixiPol).

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 стал в 1997 году (на базе Digital Unix) одной из первых интернет-ориентированных баз данных в России, использовавшим данную технологию, причем за 15 лет она несколько не потеряла свою технологичность на стороне сервера. Суммарный объем базы составляет около 500 Мб.

В теоретической основе WWW-MINCRYST лежит ряд принципиальных "идеологических" моментов, реализация которых и 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, поскольку позволяет существенно расширить информационную базу сравнительно с чисто аналитической.

Как уже сказано, информационный фонд WWW-MINCRYST на июнь 2013 насчитывает около 8500 записей для более чем 3000 уникальных фаз (минералов, их аналогов, элементов), т.е. большинства минералов

(из 4800 официально признанных - <http://www.ima-mineralogy.org/Minlist.htm>), 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 120 иностранных и отечественных 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 300-400 структур в год). В последние годы основной акцент был 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).

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 программной (на ПК) экспертизой по результатам расчета межатомных расстояний и других кристаллоструктурных характеристик и, в случае положительного решения, импортируется в ИВС специальными средствами 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, 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, 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 (Г.Б. Бокий). 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 данных, что позволяет проводить постоянную актуализацию информационного фонда. Также реализована возможность online редакции записей, их удаления, замены служебных файлов записей, архивации и восстановления базы данных через web-интерфейс.

#### 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, Чириотти, Бокий). Система поиска по этим пара-метрам 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-Xraypol 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 с использованием библиотек OpenGL) 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 "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. Более детальное описание всех возможностей апплета приведено здесь: <http://mincryst.iem.ac.ru/gus/crystpic.php>. В настоящее время из-за проблем с работой с библиотеками OpenGL (на которых реализован апплет) в составе новых версий Java проводится перенос 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 но-вые 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 данных по 8500 кристаллическим структурам 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 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 (4 млн. успешных единичных запросов за 2012 год, более 52 Гб скачанной информации, около 35000 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 не полностью совместимого 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 JSON, DHTML, HTML5, 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 HTML5 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, рисунки кристаллографических форм, дополнительные минералогические данные 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).

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 1997-2013 годов было поддержано несколькими грантами РФФИ (в том числе в настоящее время - грантом РФФИ 12-07-00742-а, рук. Варламов Д.А.).

#### ЛИТЕРАТУРА:

1. А.В. Чичагов, Д.А. Варламов, Е.В. Ершов, Т.Н. Докина, Н.А. Дрожжина, О.Л. Самохвалова «Кристаллографическая и кристаллохимическая база данных для минералов и их структурных аналогов (WWW-MINCRYST)» // Записки РМО, 2007, т.136, № 3, с.135-141
2. Д.А. Варламов, Т.Н. Докина, Н.А. Дрожжина, О.Л. Самохвалова "WWW-MINCRYST: Интернет-ориентированная информационно-вычислительная система по кристаллографии и кристаллохимии минералов" // Вестник ЮУрГУ, Серия «Вычислительная математика и информатика», 2013, т.2, вып.1, с.26-32
3. А.В. Чичагов, Д.А. Варламов WWW-MINCRYST-2007 - Интернет-ориентированная база данных по кристаллографии/кристаллохимии минералов и их аналогов // «Научный сервис в сети Интернет: технологии параллельного программирования. 15 лет РФФИ», Труды Всероссийской научной конференции, М., изд-во МГУ, 2007, с.390-392