

«Электронная структура атомов» и другие информационные системы по атомной спектроскопии: возможности специализации и кооперации в среде Интернет

А.С. Яценко^{1,2}, В.Г. Казаков^{1,3}, В.В. Казаков^{1,3}, В.С. Ковалев¹

¹Новосибирский государственный университет

²Институт автоматизации и электротехники СО РАН

³Новосибирский государственный университет экономики и управления
grotrian@nsu.ru, kazakov@phys.nsu.ru, vkazakov@phys.nsu.ru,
kovalev@mmedia.nsu.ru

Аннотация

Обсуждаются тенденции развития современных ресурсов по атомной спектроскопии, формулируются задачи развития информационной системы «Электронная структура атомов», связанные с основными трендами в эволюции аналогичных систем. Описаны работы по переводу системы на открытые форматы данных и включению в международную систему научной кооперации. Представлены новые возможности обработки и анализа данных, реализованные в системе. Рассмотрена реализация программных средств автоматизированной актуализации базы данных системы, основанная на инероперабельности современных информационных ресурсов по спектроскопии. Представлены некоторые планы дальнейшего развития системы.

Ключевые слова: атомные спектры; базы данных; инероперабельность; публикация в Интернет

1. Введение

К середине 2000-х годов был завершен процесс публикации информационных ресурсов в области спектроскопии высокого разрешения в Интернет в виде информационно-поисковых систем с Веб интерфейсом. Сегодня в области развития таких информационных ресурсов, в частности информационных систем по атомной спектроскопии, действуют две основные тенденции. Во-первых, информационные системы эволюционируют от информационно-поисковых систем к системам поддержки принятия решений: развиваются сложные алгоритмы обработки данных, средства научной графики и другие компоненты, характерные для таких систем [1,2]. Во-вторых, появляются общие стандарты представления данных, такие как XSAMS [3, 4],

на основе которых развивается интероперабельность ряда ведущих систем, что служит основой для международной научной кооперации и специализации, например, в проектах VAMDC [5, 6].

На базе института Института автоматизации и электротехники сибирского отделения РАН, Национального исследовательского новосибирского государственного университета и Новосибирского государственного университета экономики и управления в течение ряда лет развивается информационная система по атомной спектроскопии «Электронная структура атомов» (ИС ЭСА), не уступающая, а по ряду параметров превосходящая ведущие мировые аналоги [7, 8, 9]. В связи с указанными общими тенденциями развития подобных систем, перед коллективом, разрабатывающим и поддерживающим ИС ЭСА, стоит задача построения планов эффективного развития ресурса. При этом представляется, что для того, чтобы осуществлять на современном уровне поддержку научных исследований, совершенно необходимо не только уделять внимание базам спектральных данных, но и включать в систему развитые средства научной визуализации и анализа данных. С другой стороны, такое развитие системы требует значительных трудозатрат и может быть в современных условиях эффективно осуществлено в разумные сроки только при определенном разделении труда и использования возможностей кооперации между научными группами, осуществляющими поддержку таких систем.

Для того чтобы ИС ЭСА соответствовала современным требованиям к ведущим ресурсам по атомной спектроскопии, были сформулированы три задачи развития системы, учитывающие названные выше общие тенденции. В течение 2013–2014 гг. по каждой из этих задач были выполнены первый этап работ и получены результаты.

2. Создание основ интероперабельности

Первой задачей стало обеспечение возможности экспорта данных в формате XSAMS. Первоначально экспорт реализован в наиболее простом варианте. Модель данных системы и внутренние алгоритмы работы не претерпели изменений, а функции экспорта было решено осуществлять посредством внешнего слоя сопряжения, позволяющего стороннему пользователю получать данные в формате XSAMS. В текущем состоянии запрос может формироваться только средствами интерфейса системы, а результат представляется в виде файла, доступного для скачивания пользователем.

Следующим шагом в реализации функций экспорта, будет являться выдача данных по внешнему запросу. В этой связи встает вопрос о языке (языках) для запроса данных. Основным кандидатом на текущий момент является язык запросов проекта VAMDC – это SQL-подобный язык запросов, применяемый на портале VAMDC для формирования запросов спектроскопических данных. Поддержка внешнего языка запросов в сочетании со стандартизованным представлением результатов позволит внешним системам (в частности portalу VAMDC) использовать ИС ЭСА в качестве сервера спектральных данных для использования последних в собственных обработчиках и анализаторах.

3. Создание средств анализа данных и выделения их в виде автономных сервисов

Второй задачей развития ИС ЭСА стало совершенствование встроенных в систему средств обработки и анализа информации. В соответствии с общей тенденцией модифицируемые средства предполагают значительно более развитые возможности анализа информации. При этом ставилась задача выделения таких средств в отдельные модули и сервисы, которые могли бы быть предложены для использования в других информационных ресурсах по спектроскопии в рамках международной научной кооперации. Первым средством обработки данных, который подвергся таким изменениям, стал виртуальный спектроскоп – средство обеспечивающее представление спектра нейтрального атома или атомного иона на основе базы данных ИС ЭСА [10, 11]. В то время, когда первая версия виртуального спектроскопа работала с внутренними форматами данных ИС ЭСА, новая версия виртуального спектроскопа рассчитана на работу со спектральными данными в формате XSAMS. Это обеспечивает возможность использования модуля вне ИС ЭСА: пользователь может подгружать для визуализации собственные XSAMS спектры, полученные, например, в других системах. Кроме того, программная компонента может быть развернута как часть любой системы, способной представлять данные в указанном формате.

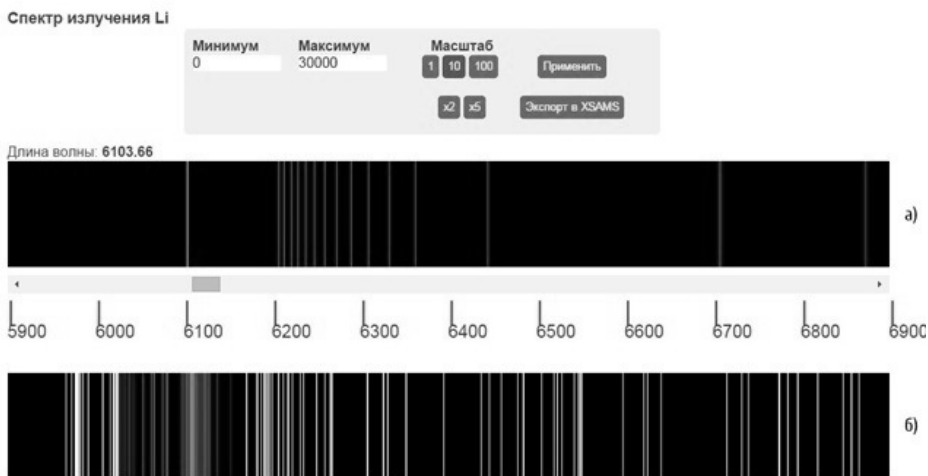


Рис. 1. Интерфейс виртуального спектроскопа на примере спектрограммы Li. Вверху рисунка расположены инструменты выбора участка спектрограммы и установка масштаба отображения:

- а) Спектрограмма участка диапазона рассматриваемого спектра. Вверху –
- б) Спектрограмма всего диапазона отображаемого спектра.

В новой реализации спектроскопа также существенно усложнены алгоритмы обработки и анализа данных. Он получил существенно обновленный и более

удобный интерфейс, включающий новые средства представления спектров. Так, например, «карта» спектральных линий (рис. 1.б) позволяет видеть все линии переходов, присутствующих в файле XSAMS, а также позицию участка, находящегося в области детального просмотра. В области детального просмотра (рис. 1.а) пользователь может увеличивать масштаб детализации, пользуясь стандартными множителями масштаба. В нижней части области детального просмотра находится шкала, по которой пользователь может определить диапазон длин волн. По каждой линии можно получить её XSAMS-описание, при нажатии курсором мыши. Между областью детального просмотра линий и «картой» находится элемент управления, позволяющий перемещать область детального просмотра по всему, существующему в XSAMS диапазону спектральных линий.

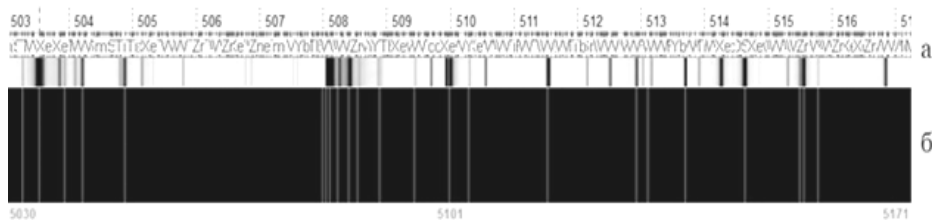


Рис. 2. Сравнение спектрограмм атома никеля, экспериментально полученной (а) и эмулированной из ЭСА (б) в видимой области (5030-5171) Å.

Переход на открытые форматы позволил построить на базе виртуального спектроскопа средство анализа спектров. Идея состоит в том, что в виртуальный спектроскоп загружается два спектра: анализируемый спектр, который представлен в одном из форматов, используемых для хранения экспериментальных спектрограмм, и эталонный спектр в формате XSAMS. Анализатор спектров позволяет синхронно позиционировать и масштабировать спектры, давая специалисту возможность их сравнивать и делать заключение об их схожести. Наиболее простым вариантом использования анализатора спектров является определение наличия в экспериментальном спектре, характерных линий какой-либо атомной системы. В качестве примера на рис. 2.а, 2.б приведены две спектрограммы – экспериментальная и эмулированная. На рис. 2.а – экспериментальная спектрограмма атома Ni, содержащая кроме основного элемента 23 вида других примесей. На рис. N.b – эмулированная спектрограмма того же элемента. При сопоставлении двух диаграмм и выявляются примесные элементы.

4. Актуализация базы данных

Третьей задачей развития ИС ЭСА, является актуализация базы данных. До сих пор основным способом пополнения баз данных, как, впрочем, и в других аналогичных системах, являлся ручной ввод данных специалистом, основанный на анализе первоисточников и отборе необходимой информации. Таким

образом была создана уникальная база данных ИС ЭСА содержащая на настоящий момент около 200 тысяч записей по параметрам уровней и переходов нейтральных атомов, однократных и многократных атомным ионам.

В то же время, была сделана попытка использовать для актуализации баз данных возможности интероперабельности ведущих систем по атомной спектроскопии. Поскольку специалисты, поддерживающие каждую такую систему, осуществляют анализ первоисточников для пополнения баз данных, представляется разумным организовывать в этой работе специализацию научных групп по областям анализа и последующую кооперацию, выражающуюся в использовании полученной информации для актуализации всех ресурсов, для которых эти данные представляют интерес. При этом за счет интероперабельности систем, подобный информационный обмен может осуществляться в автоматическом или полуавтоматическом режимах.

В ИС ЭСА построен модуль, который в полуавтоматическом режиме позволяет импортировать данные из других систем, поддерживающих формат XSAMS. При этом проводится автоматический анализ поступившей спектральной информации, и для каждого уровня или перехода определяются его статус. Например, данный уровень (переход) присутствует в базе данных, уровень (переход) присутствует в ИС ЭСА, однако в импортируемых данных представлены отсутствующие параметры, данная запись в ИС ЭСА отсутствуют и т.д. В соответствии с определяемым статусом выбирается алгоритм импорта записи: она отклоняется, информация используется для уточнения уже существующей в БД ИС ЭСА записи, генерируется новая запись и др.

При анализе импортируемых данных могут возникать весьма сложные и запутанные ситуации, поэтому на данном этапе окончательная проверка и принятие решения в процессе импортирования остается за экспертом. Однако даже в таком случае эффективность работ по актуализации базы данных системы возрастает многократно.

4. Заключение

Таким образом, в ИС ЭСА проводятся работы по развитию интероперабельности и созданию средств поддержки принятия решений, что соответствует общим тенденциям, наблюдаемым в развитии информационных ресурсов по атомной спектроскопии. Работы ведутся по трем основным направлениям: экспорт данных в открытых форматах, построение сложных средств обработки и анализа спектральных данных с оформлением в виде отдельных веб-приложений, автоматизированная актуализация баз данных на основе обмена данными с авторитетными ресурсами по атомной спектроскопии. На каждом из этих направлений получены конкретные результаты, позволяющие говорить о перспективности выбранного направления для общего прогресса в развитии ИС ЭСА.

Тенденции развития информационных ресурсов по спектроскопии позволяют говорить о том, что в ближайшее время можно ожидать новое поколение информационных ресурсов, ориентированных на поддержку научных исследований.

С одной стороны, такие ресурсы будут эволюционировать от информационно-поисковых систем к системам поддержки принятия решений. Будут активно развиваться средства визуализации и анализа данных, построенные на основе сложных алгоритмов. Будут повышаться требования к полноте и достоверности данных, лежащих в основе ресурсов. Все это потребует существенного увеличения трудозатрат для поддержания информационного ресурса в актуальном состоянии.

С другой стороны, будет развиваться интероперабельность ресурсов на основе открытых форматов и протоколов. Это позволит развивать научную кооперацию ресурсов на базе специализации отдельных коллективов. Так функциональность по визуализации и анализу данных перестанет быть связана с конкретными информационными системами и может быть организована в виде самостоятельных сервисов, к которым могут обращаться все информационные ресурсы. Также может быть организован информационный обмен между базами данных различных систем с целью избежать дублирования работ по подготовке и вводу новых данных.

Работа выполнена при поддержке Российского фонда фундаментальных исследований, грант № 13-07-00973.

Литература

- [1] Казаков В.Г., Раутиан С.Г., Яценко А.С. Компьютерное представление характеристик электронных оболочек атомов. // Оптика и спектроскопия. 2008. Т. 105. № 1. С. 53 – 58.
- [2] Казаков В.Г. и др. Информационные системы по атомной спектроскопии: от информационно-поисковых к системам поддержки принятия решений. / Казаков В.Г., Казаков В.В., Ковалев В.С., Федотов А.М., Яценко А.С. // Вестник Новосибирского государственного университета экономики и управления. 2014. № 2. С. 268 – 279.
- [3] XSAMS: XML Schema for Atoms, Molecules and Solid. // International Atomic Energy Agency. Vienna, Austria. URL: <https://www-amdis.iaea.org/xsams/> (дата обращения: 20.09.2014).
- [4] Ralchenko Y., Clark R.E.H. et al. Development of New Standards for Exchange of Atomic and Molecular Data. // AIP Conference Proceedings 6th International Conference on Atomic and Molecular Data and Their Applications. Сер. "ICAMDATA-2008 - 6th International Conference on Atomic and Molecular Data and Their Applications". Beijing, 2009. С. 207 – 216.
- [5] VAMDC // Virtual Atom and Molecular Data Centre. VAMDC Consortium. URL: <http://www.vamdc.eu/index.php/> (дата обращения: 20.09.2014).
- [6] Dubernet M.L. Virtual Atomic And Molecular Data Centre / Dubernet M.L., Roueff E., Boudon V. et al. // Journal of Quantitative Spectroscopy and Radiative Transfer. 2010. Т. 111. № 15. С. 2151–2159.
- [7] Казаков В.Г., Яценко А.С., Казаков В.В. и группа разработчиков ИС ЭСА. Информационная система «Электронная структура атомов» (версия. 1.03) // Новосибирский государственный университет, Новосибирск. URL: <http://grotrian.nsu.ru> (дата обращения: 20.03.2015).

- [8] Казаков В. Г., Тюменцев А. С., Яценко А. С. Информационная система «Электронная структура атомов» с динамическим построением графического представления спектральных данных. // Автометрия. 2005. Т.41. №6. С.115 – 123.
- [9] Казаков В.Г., Яценко А.С., Казаков В.В. Информационная система «Электронная структура атомов»: основные возможности и особенности. // Вестник Новосибирского государственного университета. Серия: Физика. 2011. Т. 6. № 3. С. 64 – 70.
- [10] Казаков В.Г., Яценко А.С., Казаков В.В. Эмулированный анализатор оптического спектра. // Патент на полезную модель RUS 117178 23.11.2011
- [11] Казаков В.В. и др. Цифровая эмуляция спектрографа. / Казаков В.В., Яценко А.С., Ковалёв В.С., Казаков В.Г // Вестник Новосибирского государственного университета. Серия: Информационные технологии. 2011. Т. 9. № 3. С. 30 – 36.

System “Electronic Structure of Atoms” and Other Databases on Atomic Spectroscopy: Opportunities for Specialization and Cooperation in Internet

A.S. Yatsenko^{1,2}, V.G. Kazakov^{1,3}, V.V. Kazakov^{1,3}, V.S. Kovalev¹

¹Novosibirsk State University, ²Institute of Automation and Electrometry, Siberian Branch of RAS, ³Novosibirsk State University of Economy and Management

The problems of the development of an information system for atomic spectroscopy "Electronic structure of atoms", associated with the main trends in the evolution of similar systems. The transformation system to open formats of data, and inclusion in the international system of scientific cooperation are described. Presents new opportunities and data analysis processing implemented in the system. Realization of automated updating of the database system based on modern information interoperability resources spectroscopy. Presented some plans for further development of the system.

Keywords: atomic spectra; database; interoperability; publication on the Internet