

Информационная система «Электронная структура атомов»: современное состояние и обеспечение интероперабельности и интеграции *

А.С. Яценко, В.В. Казаков, В.Г. Казаков, В.С. Ковалев, С. В. Ретунский, Е.В. Шевцов

Институт автоматизации и электротехники СО РАН, Новосибирский государственный университет, Новосибирский государственный университет экономики и управления
grotrian@nsu.ru, kazakov@phys.nsu.ru, kazakov@phys.nsu.ru, kovalev@mmedia.nsu.ru, kirshe13@yandex.ru, admin@evgeniy.ws

Аннотация

Описывается опубликованная в Веб-информационная система по атомной спектроскопии «Электронная структура атомов». Рассматриваются некоторые тенденции развития подобных ресурсов и возникающие в связи с этим проблемы. Обсуждаются возможности повышения эффективности систем по спектроскопии высокого разрешения на базе развития интероперабельности и интеграции. Приводятся сведения по планам и состоянию работ по переводу системы «Электронная структура атомов» на международный формат данных XSAMS и организации взаимодействия с другими ресурсами.

1. Современные информационные системы по атомной спектроскопии

Одним из примеров Интернет-ресурсов, критически важных для поддержки научных исследований, являются информационные системы по атомной спектроскопии. Анализ информации об оптических спектрах атомных и газовых систем занимает ключевое место в исследованиях в целом ряде научных направлений фундаментальной физики и прикладных областях, от астрофизики и физики газовых лазеров до элементного анализа и археологии. При этом, в природе существуют сотни атомов и ионов с десятками и сотнями уровней и разрешенных переходов, что составляет огромный, постоянно пополняемый, объем информации, организация эффективной работы исследователя с которым является сложной задачей. Наиболее перспективным механизмом предоставления спектральных данных сегодня являются информационные системы, основанные на базах данных и опубликованные в Интернет, способные

обеспечить специалиста актуальной информацией последних исследований и предоставить широкий спектр механизмов выборки и сортировки.

На сегодняшний день в мире существует достаточно много информационных ресурсов по спектральным данным атомных систем. В связи с большей оперативностью и доступностью большую популярность в научном сообществе имеют информационные системы на основе баз данных, обеспечивающие взаимодействие с пользователем по протоколу HTTP. Так в США национальным институтом стандартов и технологий (NIST) поддерживается система ASD (Atomic Spectra Database) [1]. Аналогичны ей системы NIFS (Япония), AMIDS (Австрия), OPD (Франция), COREX (Швеция). В России имеется несколько автоматизированных систем позволяющих оперировать со спектроскопическими данными атомов. Например, в ФИАН (г. Москва) используется программа «АТОМ» для расчетов различных характеристик атомов. База данных по ридберговским и автоионизационным состояниям используется в ИСАН (г. Троицк). В Снежинске имеется база данных SPECTR-W3 по оптическим спектрам атомов и ионов изоэлектронных рядов, опубликованная в 2006 г. в интернет и доступна для исследователей.

С 2003 года на базе Новосибирского государственного университета с участием авторов настоящей публикации ведется проект «Гротриан», направленный на создание информационной системы по атомным спектрам, сочетающей обширную базу данных с визуализацией спектральной информации [2]. В 2005 году созданная в рамках проекта информационная система «Электронная структура атомов» (ИС ЭСА) была опубликована в Интернет [3]. К настоящему времени данная система является наиболее полным отечественным ресурсом по информации о спектрах атомных систем и не уступает по основным параметрам ведущим мировым системам по атомной спектроскопии.

Технологии информационного общества в науке, образовании и культуре: сборник научных статей. Материалы XVII Всероссийской объединенной конференции «Интернет и современное общество» IMS-2014, Санкт-Петербург, 19 - 20 ноября 2014 г.

2. Современные тенденции в развитии информационных систем по атомной спектроскопии

Информационные системы по атомной спектроскопии появились как публикация в Web баз данных спектров, которые в свою очередь были переводом в компьютерную форму печатных справочных изданий, существовавших еще с начала 20 века. Новые технологии существенно определяют современные тенденции развития таких систем. В их числе повышенные требования к полноте и актуальности баз данных, визуализация спектральных данных, средства обеспечения целостности и достоверности данных и другие. Можно сказать, что основной тенденцией развития информационных систем по атомной (и молекулярной) спектроскопии является эволюция от информационно-поисковых систем к системам поддержки принятия решений.

2.1. Базы спектральных данных

Для современных информационных систем по атомной спектроскопии одними из важнейших показателей систем является объем, полнота и актуальность их баз данных. Число записей в наиболее авторитетных системах составляют десятки и сотни тысяч. Примером и своеобразным эталоном может служить информационная система NIST ASD, база данных которой является одной из наиболее полных. В этой базе содержатся следующие данные о спектральных линиях: длина волны в воздухе, относительная интенсивность, коэффициент Эйнштейна, энергии верхний и нижних уровней, их конфигурация и термы. Имеется также оценка точности приведенных величин. В базе хранятся данные о переходах 99 элементов и уровнях энергии 52 элементов. Всего описано 70000 уровней энергии и 91000 линий, для 40000 линий имеются вероятности перехода. База данных системы постоянно пополняется данными новейших исследований.

База данных ИС «Электронная структура атомов» является одной из самых больших и включает около 200000 записей по радиационным уровням и переходам нейтральных атомов, содержащих широкий перечень фиксируемых параметров. Для уровней это: электронная конфигурация, энергия, время жизни. Для переходов: энергия перехода, длина волны, сила осциллятора, интенсивность, вероятность перехода, сечения возбуждения, а также ссылки на образующие переход уровни. В базе данных представлены все элементы периодической таблицы. Основными особенностями базы данных являются большая, по сравнению с аналогами, полнота представления трансурановых и редкоземельных элементов с большим числом Z , а также то, что большинство введенных в базу данных переходов классифицированы. База данных системы также постоянно расширяется как за счет

включения данных последних исследований, так и спектральной информацией по многократным ионам.

2.2. Визуализация спектров в виде спектрограмм

Одним из наиболее эффективных видов визуализации спектральных данных являются спектрограммы. Такая возможность реализована в ряде информационных систем по атомной спектроскопии. Так, например, в ASD NIST имеются возможности построения спектрограмм - изображений атомных спектров. Спектрограмма в ASD реализована как черно-белая растровая графика, которая на лету генерируется системой по запросу пользователя, включающему атомную систему и спектральный диапазон.

Имеется система представления спектров в виде спектрограмм и в ИС «Электронная структура атомов». Однако здесь они реализованы по-другому, оригинальным способом [4]. Спектрограмма в ИС ЭСА является «насыщенным» Интернет приложением (RIA, Rich Internet Application), имитирующем просмотр исследователем спектра с помощью спектроскопа [5]. Пользователь может интерактивно выбирать масштаб спектра и просматриваемый диапазон, а линии в оптической его части соответствующим образом окрашены. Сущность метода состоит в том, чтобы на основе базы данных спектральных линий численно моделировать изображения, получаемые спектральным прибором.

При этом в распоряжении исследователя имеется целый набор инструментов: лупа, всплывающие подсказки с информацией о параметрах выбранной линии и опирающихся на нее переходах и т.д.

Общий вид интерфейса для отображения спектрограмм приведен на Рис.1.

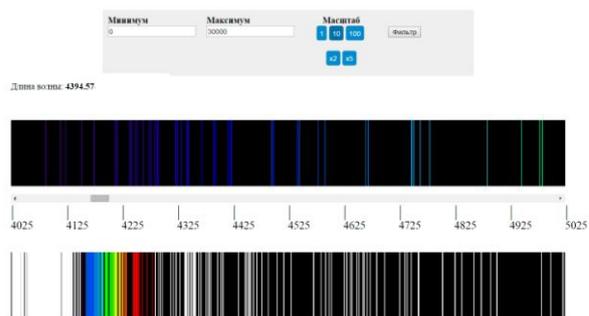


Рис.1 Общий вид интерфейса спектрограммы Na I.

На настоящий момент средства визуализации атомных спектров в ИС ЭСА, по-видимому, являются наиболее развитыми среди всех ИС по атомной спектроскопии, опубликованных в Web.

2.3. Графические представления электронной структуры атомов

Еще одним элементом сложной обработки и визуализации спектральной информации, развиваемой в современных информационных системах является визуализация электронной структуры атомов, традиционно реализуемая в виде диаграмм Гротриана [6]. Дело в том, что исследователю, наряду с точными числовыми данными об уровнях и переходах, требуется графическое представление таких данных, незаменимое для целей общего анализа электронной структуры атомов. Построение таких диаграмм является сложной интеллектуальной задачей, решение которой до последнего времени было возможно только для квалифицированных специалистов и осуществлялось в «вручную». Ситуация усугубляется и тем, что такой способ построения диаграмм весьма трудоемок и поэтому ощущается их постоянный дефицит.

В последние годы различными научными коллективами был сделан ряд попыток компьютеризировать процесс построения диаграмм Гротриана, что могло бы сократить его трудоемкость. В настоящее время возможность построения диаграмм Гротриана имеется в уже упоминавшейся ASD NIST. В ней присутствует возможность полностью автоматического построения диаграмм Гротриана, осуществляемая динамически по запросу пользователя. Сами же полученные диаграммы, также как и спектрограммы статичны и представлены растровыми графическими изображениями.

ИС ЭСА является на сегодня одной из двух информационных системой по атомной спектроскопии, в которой развито автоматическое построение диаграмм Гротриана, которое осуществляется динамически по запросу пользователя и по спектральной информации, хранящейся в базе данных атомных уровней и переходов системы на момент построения. Ряд решений в ИС ЭСА обеспечивает лучшее по сравнению с ASD качество получаемых диаграмм [7].

Главным и наиболее принципиальным отличием является наличие специального алгоритма, обеспечивающего отбор информации для размещения на диаграмме. Данный алгоритм содержит основы интеллектуального анализа данных и отбирает линии, имитируя интеллектуальную деятельность специалиста по построению диаграмм. Алгоритм классифицирует уровни и переходу по ряду признаков и, затем, отбирает из каждой группы для размещения на диаграмме такое количество, чтобы при соблюдении общей читаемости представить все основные особенности отображаемого спектра.

При этом диаграммы Гротриана, сгенерированная ИС ЭСА, так же как и спектрограммы, являются RIA приложениями,

выполнены в векторной графике, интерактивны и предоставляют пользователю ряд инструментов для дополнительной настройки вида диаграммы, например, выбор энергетического диапазона для отображения.

Пример диаграммы сгенерированной ИС ЭСА приведен на Рис.2.

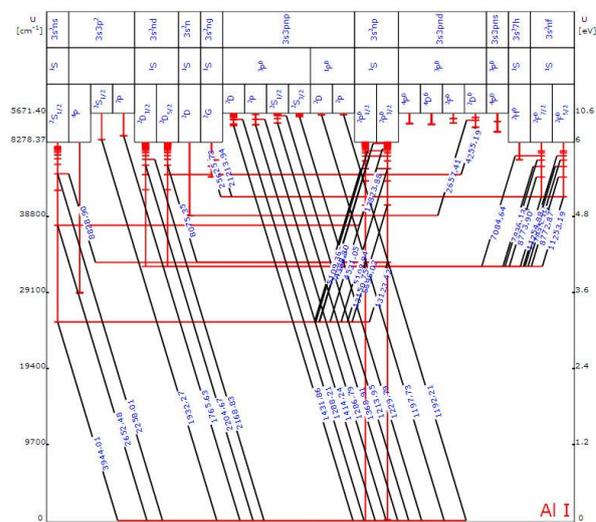


Рис. 2. Диаграмма Гротриана для атома алюминия Al I.

Можно констатировать, что на сегодняшний день ИС ЭСА обладает наиболее сложными и эффективными средствами визуализации электронной структуры атомных систем в виде диаграмм Гротриана среди всех подобных спектроскопических Интернет ресурсов.

3. Интеграция и интероперабельность

Следование данным тенденциям требуют от научно-технических коллективов, поддерживающих и развивающих современные информационные системы по атомной спектроскопии все больших усилий. Анализ первоисточников предполагает просмотр все возрастающего количества оригинальных печатных работ по спектроскопии с целью как поиска информации о вновь выявленных или рассчитанных линиях и переходах, так и уточнения параметров уже известных.

Средства визуализации спектральных данных, включая построение спектрограмм и графических представлений электронной структуры атомов, требуют составления все более сложных алгоритмов и все более развитых графических интерфейсов, построенных как RIA приложения.

Все это требует для создания и развития современных информационных ресурсов в области спектроскопии создания команд из высококвалифицированных специалистов не только в области физики, но и различных областей информационных технологий: базы данных, веб интерфейсы, научная растровая и векторная

графика и т.д. С учетом того, что сегодня каждая научная держава поддерживает по крайней мере одну информационную систему по атомной спектроскопии, общие затраты на их развитие весьма значительны. В связи с этим весьма актуальным является развитие различных методов сокращения издержек на поддержку и развитие таких ресурсов: специализация отдельных систем, обмен данными между системами, переиспользование алгоритмов и программных модулей для обработки данных.

До последнего времени в области информационных ресурсов по спектроскопии высокого разрешения было весьма затруднительным, поскольку каждая из систем создавалась отдельно и автономно. Конечно, существовала практика обмена данными или их получения путем синтаксического разбора генерируемых системой Web страниц, однако из-за разных схем данных и отсутствия семантической разметки это каждый раз осуществлялось специалистом вручную.

В последнее время ситуация изменилась в связи с появлением общепризнанных форматов представления данных, служащих основой для обмена информации о спектрах между различными системами. Под эгидой МАГАТЭ была разработана спецификация языка XSAMS (XML Schema for Atoms, Molecules and Solids), предназначенного для унификации представлений спектральных данных [8]. Первая версия спецификации XSAMS v.1.0 опубликована в 2009 году и представлялась на различных научных мероприятиях [9].

В связи с появлением XSAMS в информационной системе «Электронная структура атомов» была развита новая функциональность, направленная на развитие международной научной специализации на основе интероперабельности информационных систем по атомной спектроскопии. Мы уже отмечали, что ИС ЭСА обладает развитыми средствами формирования интерактивных спектрограмм. Теперь эти средства могут применяться не только по отношению к спектрам из базы данных системы, но и строиться для любых спектров, сформированных в формате XSAMS. Таким образом, утилита становится доступной для визуализации спектральных результатов, более двух десятков информационных систем по спектроскопии, представляющих данные на запросы пользователей в формате XSAMS. В дальнейшем планируется перевод на XSAMS других средств обработки спектральных данных, которые могут быть интересны международному сообществу. В первую очередь это средства построения графических представлений электронной структуры атомных систем в виде диаграмм Гротриана, а также пакет программ, реализующих алгоритмы обеспечения достоверности информации об атомных спектрах.

С другой стороны появление формата XSAMS позволяет актуализировать базу данных ИС ЭСА в

автоматизированным режиме через анализ интересующих спектральных данных из других систем по атомной спектроскопии, представленных в том же формате. Такие средства в настоящее время в ИС ЭСА активно разрабатываются.

Очень интересен опыт международного проекта Virtual Atomic and Molecular Data Center (VAMDC) [11]. В рамках этого проекта, по существу, построена площадка, портал, для интеграции информационных ресурсов по атомной и молекулярной спектроскопии, для которых реализована интероперабельность на основе формата XSAMS [10]. Центральные службы портала, во-первых, позволяют пользователю отправить запрос на получение спектральной информации, который будет оттранслирован всем информационным системам, имеющим в портал шлюз, и получить от тех систем, которые обладают необходимой информацией, наборы данных в формате XSAMS. Во-вторых, пользователь портала может провести обработку полученных от какой-либо системы данных с помощью одной из целого набора приложений, предоставляемых порталом. Роль таких порталов неочевидна: они позволяют собрать усилия многих коллективов в одном месте и дают возможность исследователю работать с различными наборами данных и средствами их обработки, по сути, объединяя несколько систем в одну.

В настоящее время одной из приоритетных задач VAMDC является включение информационных и программных ресурсов в портал VAMDC. На данный момент коллективом разработчиков ИС «ЭСА» подготовлены для включения в портал средства построения спектрограмм, отлаживаются программный шлюз, осуществляются работы по подготовке к включению в портал средств генерации диаграмм Гротриана.

Заключение

Современные web ресурсы по атомной спектроскопии являются сложными информационными системами, в настоящее время эволюционирующими от информационно-поисковых к системам поддержки принятия решений и могут служить некоторым индикатором развития многих систем поддержки научных исследований, основанных на обширных базах данных.

Отечественная информационная система «Электронная структура атомов» соответствует основным требованиям и не уступает лучшим зарубежным аналогам.

Постоянное повышение сложности таких систем требует эффективной международной кооперации, основанной на интероперабельности и интеграции. Развитие интероперабельности систем по атомной и молекулярной спектроскопии возможно на базе формата XSAMS с интеграцией на портале VAMDC.

В рамках развития информационной системы «Электронная структура атомов» реализован переход на формат XSAMS, что позволяет предоставить в общее пользование ряд средств обработки спектральных данных, разработанных коллективом ИС «ЭСА».

Литература

- [1] Ralchenko Yu., Kramida A.E., Reader J., and NIST ASD Team (2010). NIST Atomic Spectra Database (ver. 4.0.1) [Электронный ресурс] / National Institute of Standards and Technology, Gaithersburg, USA. URL: <http://physics.nist.gov/asd>.
- [2] 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.
- [3] Казаков В.Г., Яценко А.С., Казаков В.В. и группа разработчиков ИС ЭСА. Информационная система «Электронная структура атомов» (версия. 1.03) [Электронный ресурс] / Новосибирский государственный университет, Новосибирск, Россия. URL: <http://grotrian.nsu.ru> (дата обращения: 9 февраля 2011).
- [4] Казаков В.Г., Яценко А.С., Казаков В.В. Эмулированный анализатор оптического спектра. Патент на полезную модель RUS 117178 23.11.2011.
- [5] Казаков В.В., Яценко А.С., Ковалёв В.С., Казаков В.Г. Цифровая эмуляция спектрографа. Вестник Новосибирского государственного университета. Серия: Информационные технологии. 2011. Т. 9. № 3. С. 30-36.
- [6] 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.
- [7] 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.
- [8] International Atomic Energy Agency (IAEA): Official Web Site of the IAEA [Электронный ресурс]. URL: <http://www.iaea.org> (дата обращения 10 сентября 2014 г.).
- [9] 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.
- [10] XSAMS: XML Schema for Atoms, Molecules and Solid [Электронный ресурс] / International Atomic Energy Agency. Vienna, Austria. URL: <https://www-amdis.iaea.org/xsams/> (дата обращения 10 сентября 2014 г.).
- [11] VAMDC – Virtual Atom and Molecular Data Centre [Электронный ресурс] / VAMDC Consortium. URL: <http://www.vamdc.eu/index.php> (дата обращения 10 сентября 2014 г.).
- [12] Dubernet M.L., Roueff E., Boudon V. et al. Virtual Atomic And Molecular Data Centre // Journal of Quantitative Spectroscopy and Radiative Transfer. 2010. Vol. 111 (15). P. 2151-2159.

Information System "Electronic Structure of Atoms": Current Status and Ensure Interoperability and Integration

A.S.Yatsenko, V.V. Kazakov, V.G. Kazakov, V.S. Kovalev, S.V. Retunskiy, E.V. Shevtsov

The article describes published in the web information system for atomic spectroscopy, "Electronic structure of atoms." The article described some trends and problems of the development such information resources and considered the possibility of increasing the efficiency of systems for high-resolution spectroscopy based on the development of interoperability and integration. Authors provide information on the plans and work of the shifting "Electronic structure of atoms" to the international format data XSAMS and interaction with other resources.

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