Prospects for the Evolution of the Russian Segment of the Virtual Atomic and Molecular Data Center

Alexey Akhlyostin, Nikolay Lavrentiev, Alexey Privezetsev,

Alexander Z. Fazliev [0000-0003-2625-3156]

Institute of Atmospheric Optics SB RAS, V.E.Zuev sq.1, Tomsk 634055, Russia
lexa@iao.ru, remake@iao.ru, lnick@iao.ru, faz@iao.ru

Abstract. The paper discusses the prospects for the development of information systems associated with the Russian nodes that are part of the Virtual Atomic and Molecular Data Center. Atomic, ionic, and molecular spectral data are accumulated at these nodes. The key issues of the segment development are the semantization of tabular and graphical information resources and the personalization of data and its properties. This work lists the main information problems related to the three stages of the knowledge life cycle, reuse of accumulated data, and semantic search of spectral resources. It is shown how the issues of building researchers' own expert arrays based on the knowledge base of spectral information resources and physical quantities will be solved.

Keywords: Spectroscopy, Spectral Resources, Quality Control of Spectral Data, Russian Segment of the VAMDC

1 Introduction

Computer support of scientific research is focused on solving the complex tasks of integrating the efforts of researchers from different domains. It would seem that the emergence of the Internet information space has solved the access problem, however, it turned out that scientific activity requires systematized resources that must go through all the stages of the knowledge cycle [1] in order to turn from data into information and knowledge. A similar problem arose two centuries ago, when the needs of the industrial revolution led to technical progress and the development of science. The role of libraries as accumulators of fundamental information and knowledge was paramount at the time. At the end of the 20th century, the presentation of data, information and knowledge in digital form made access to them easier, but the main difference was the appearance of a new type of user - computers.

At the beginning of the 21st century the ideas of Semantic Web (SW) were introduced [2]. The SW approach was detailed in a digital science (e-Science) and cyberinfrastructure and finally in the open science data approach defined by FAIR principles. The main goal of the SW approach declared by its creators - the automatic accentuation of reliable and trustworthy information resources in the Internet infor-

mation space has not yet been achieved. The difficulty of achieving the SW goal consists, on the one hand, in the variety of domains of science and the variety of criteria determining the reliability of information resources in these areas of science, and, on the other hand, in the abundance of informal criteria for assessing the trustworthiness of information resources.

Evidently, the inaccessibility or low quality of necessary fundamental data is a limiting factor in the development of applied sciences. The removal of limitations, providing greater access to different types of information resources, containing fundamental data, and explicit evaluation of the quality of fundamental data are especially important for applied domains with intensive use of data. A good example of such domain is spectroscopy, in which constants, used in astronomy, climatology, atmospheric optics, ecology, etc. are measured and calculated.

Fundamental data play an essential role in scientific research, because they are used as input data for applied tasks, and their inaccuracy can distort research results at an early stage. Thousands of researchers form their own databases of fundamental data, combining them into dozens of virtual data centers. Such centers contain different types of data, among which the most used in practice are expert data. There are three main disadvantages of such databases. First, the update period of expert databases takes three to five years. Second, the requirements to the accuracy of the fundamental data used differ in different domains. Third, practical data optimization for solving some problems is unacceptable for solving other problems. It is possible to avoid these shortcomings by changing the established traditions and providing researchers with the means to independently compose personal fundamental data sets.

The prerequisites for creating such tools are: complete collections of published fundamental data; detailed description of the quality of information resources from which fundamental data are extracted, methods for analyzing the quality of different types of data, methods for assessing the confidence in expert resources, ontological knowledge bases that contain the results of quality research and assessment of confidence in expert fundamental data.

There is no doubt that spectroscopy is one of the data-intensive domains, and fundamental spectral data are relevant in dozens of domains, primarily in those in which data are studied remotely. For this reason, the practical application of the results of the solution of the indicated problem is relevant in the following domains: in astronomy when studying atmospheres of exoplanets and brown dwarfs; in climatology in the problems of calculations of radiation transport; in spectroscopy, for example, in the task of building sets of empirical energy level values close to the "measured" energy level values for molecules, etc.

From a practical point of view, the completeness of published data can be achieved in virtual fundamental data centers of different domains, e.g., VAMDC (Virtual Atomic and Molecular Data Center), in which the data integrity is checked and the confidence of information resources is evaluated.

The quality of primary data is determined by checking the constraints that are imposed by mathematical models of the processes in the domains under study. The data sets created by the experts are to some extent subjective, primarily in the choice of informal criteria for selection. Some criteria allow for variation in fundamental data

values in order to match experimental data in the applied subject domain. Fundamental data of a domain, modified as a result of their optimization to match the results of measurements in applied areas, cease to be fundamental data. This is exactly the situation in spectroscopy where the best known databases of spectral line parameters Hitran [3] and Geisa [4] contain a huge number of such optimized “fundamental” data. To find out the existence of such "optimization" of fundamental data, it is necessary to use quantitative estimates of confidence in expert data.

2 Information Resources and Related Information Tasks in Spectroscopy

Most of the published spectral resources in quantitative spectroscopy are associated with solutions to seven spectroscopic problems [5] - calculation of energy levels, vacuum transitions and spectral line parameters of atoms, ions, molecules or measurements of spectral functions. Such resources are presented in publications in tabular, graphical, and analytical form. Transformation of these resources into a digital representation to be used by software has not yet been completed. Currently, most of the Web data is associated with information extracted from published tables. The extraction of graphical information resources data is in its early stages. Despite the widespread use of graphical resources in domains, the research of inter-machine processing of such resources is at the initial stage. We have found only two publications related to graphical resources in nuclear physics [6] and atmospheric chemistry [7]. Practically, the work on accumulation of analytical information in the form suitable for software use (for example, selection rules, classification of quantum numbers, etc.) has not begun, and it is worth mentioning the creation of XSAMS - scheme [8] and applied ontologies of information resources on spectroscopy and spectroscopic physical quantities [9, 10].

Information tasks that need to be solved when creating and maintaining spectroscopic information resources can be divided into six groups, closely related to the stages of the knowledge life cycle [1]. In this paper, we note only the main relevant tasks related to the three stages of the knowledge life cycle: modeling, reusability, and knowledge retrieval.

Knowledge modeling. Let's note the necessity to create a model of spectroscopy (the part of it in which data and information are produced) and information subjects related to it. The task of systematization of quantum numbers and selection rules, as well as translation table for different units of physical quantities, are of great importance. Within the spectroscopy model it is important to create software for automatic generation of applied ontologies.

Knowledge reuse. It is necessary to create data and metadata formats for graphical and analytical resources, to develop a typical ontological description for reusable data, to formalize the construction of applied ontologies (information resources, physical values and restrictions on the values used in applied domains), and to develop methods for analyzing the quality of spectral data in graphical representation. The key objectives are to create an intelligent decision support system for the formation of

spectral data arrays for individual information needs and an application for data and information quality analysis.

Knowledge retrieval. Our aim is to create software for semantic search of data and information in the ontology base.

Solving these problems will provide an explicit statement of the problem of consistent use of software agents and web-services in spectroinformatics.

3 W@DIS Information System

IOA SB RAS hosts two VAMDC nodes supported by Theoretical Spectroscopy Laboratory (original computational data) and Integrated Information Systems Group (complete sets of published spectral data on atmospheric molecules). The W@DIS is an information system for developing and testing new information and knowledge handling functionalities in molecular spectroscopy.

The basic approach to the formation of the W@DIS information system was established in the early 2000s. It was based on the model of spectroscopy as a domain with intensive production of data and information, obtained by solving 7 problems defined in [5] and presented in three forms - tabular, graphical and analytical. These seven problems include direct and inverse problems of finding energy levels, vacuum wavenumbers, spectral line parameters and measuring spectral functions.

A key requirement for collections of published spectral data is their completeness. We managed to achieve this for tabular resources. Following information tasks were solved: generation of data representation of atoms, isotopes, ions, molecules and their mixtures; data import and export (with XML schema validation); generation of data properties; analysis of data sources reliability as well as presentation of data and their properties. This phase was completed in the late 2000s. At the same time an attempt was made to unite the resources of different institutions involved in molecular spectroscopy in Russia [11]. It turned out to be unsuccessful in organizational terms (there was no domestic Russian consumer, but it was found abroad).

The next set of problems arose while analyzing the quality of published primary measurement data, prepared to produce reference energy levels [12]. After achieving an acceptable quality of the primary data, they were used to analyze the confidence assessment of the wave number values in the expert data, the quality of which from the formal point of view was and still is mediocre.

Formally, to create expert data, it is necessary to know a significant number of properties of spectral data. To accumulate properties of tabular data we began to use ontologies of information resources and spectral values [13-17] and later properties of data extracted from graphical information [18, 19]. The analysis of the quality of primary data turned out to be significantly simpler than the analysis of expert data sources. This is due to the fact that statistical processing of experimental and theoretical data does not depend on the researcher, while subjective opinions of experts compiling expert data and independently determining the selection criteria require an assessment of confidence in expert data. It was necessary to use the full set of published data to find out which values used in the expert data were not published [20, 21].

Decomposition of the expert data revealed values of physical quantities in the expert data that were questionable.

The most difficult task was the systematization of the parameters of spectral lines. The complexity of this task is due to the use of more than two dozen models of spectral line contours. The first satisfactory results in constructing database structures for inverse problem solutions were obtained for the carbon dioxide molecule [22]. A series of works on finding the reference energy levels of molecules, which began with water isotopologues [12], allowed us to significantly reduce the number of doubtful transitions in the complete data collection in the W@DIS information system [23].

The creation of ontologies of information resources on spectroscopy made it possible to organize the semantic search of such resources taking into account their quality. The ontology of physical quantities enabled semantic search for specific physical quantities (describing states and transitions) in W@DIS collections [17]. The ontology of graphical spectroscopic resources provided an opportunity for semantic search of the most cited resources in the spectroscopic problems of continuum absorption and description of properties of weakly coupled molecular complexes [24].

The information problem of creating a toolkit for constructing an expert array of spectral data with the maximum computer assistance was formulated in [25]. Its partial solution is given in [26, 27]. Thus, conditions have been created to solve the problem of personalization of expert spectral data within the framework of the created data structures and their properties. The solution of this problem allows a qualified researcher to create his own expert data or, on the basis of existing expert data, to modify them with the understanding of their limitations.

The fundamental elements of the W@DIS IS and the associated software are described in our paper [28]. There are databases of substances, bibliography, data sources as well as properties of all data sources. Much of the data is housed in ontological knowledge base, containing spectral information resources and physical quantities. The ontology knowledge bases are detailed catalogs of atmospheric spectroscopy information resources and all published solutions of spectroscopy problems in the W@DIS IS.

In the near future W@DIS development will be focused on an intelligent decision-making support system for the formation of spectral data arrays for individual information needs. A further prospect is the use of agent-based and web-service technologies to solve the same problem.

4 Prospects for the Development of the Russian Segment of the VAMDC

The Russian segment of the VAMDC emerged in early 2012, when Russian participants of the VAMDC project added their information resources to those of the European Virtual Atomic and Molecular Data Center. The segment consists of 4 nodes: one at the Institute of Astronomy (atomic spectroscopy), one at the Institute of Technical Physics (ion spectroscopy), and two at the Institute of Atmospheric Optics (molecular spectroscopy). Currently, all 4 nodes are independent of each other. Three

nodes have their own sites (VALD [29], Spectr-W^3 [30], and W@DIS [31]), which provide extensive information about their resources and services that handle these resources.

Over the past 9 years since the end of the project, the functionality of VAMDC and the Russian segment information systems have developed in different directions. The consortium maintained the infrastructure of the virtual center and created services to work with integrated data [34], while data producers provided new spectral data for their nodes and developed their own services to work with them.

In the Russian segment, the information systems for atomic and ion spectroscopy emerged before the system for molecular spectroscopy. For this reason they differ from each other in the information problems they solve.

In this section, we consider the differences that exist for all major components of the spectroscopy knowledge layer. The closest to the W@DIS molecular spectroscopy model is the Spectr-W^3 ion spectroscopy model. Information content of the Spectr-W^3 database (created in 2001-2013) includes spectroscopic constants of atoms and ions, such as wavelengths and probabilities of radiative transitions, energies of atomic and ion levels, ionization potentials, autoionization rates, and parameters of approximations of cross sections and rates of collision transitions in atoms and ions by analytical expressions. The database contains about 450,000 records and is the largest factual database on the properties of the spectra of multivalent ions in the world [30].

The VALD atomic spectroscopy database [29] was created in 1992 and contains data on the spectral lines of atoms used in the analysis of stars. The evolution of this system from VALD to VALD3 is presented in [29, 32, 33].

At present, the W@DIS IS has proven to be the most technologically advanced information system for spectral resources. It contains three layers of information resources: data and applications layer, information layer and knowledge layer. The information layer contains data properties, and the knowledge layer contains ontologies of spectral information resources and physical quantities that characterize spectra. Atomic and ion spectroscopy ISs contain only the data and applications layer.

The VAMDC has only data layer information resources, and supplementing them with an information layer and a knowledge layer is rather labor-intensive. For this reason, it was decided to create an information layer and a knowledge layer on the atomic and ion spectroscopy resources of the Russian segment of VAMDC.

The prospects for the development of the Russian segment are determined by the development of information systems of Russian organizations in the consortium. The first stage of development has been completed and is described below. It includes the formation of the data layer of the information system on ion spectroscopy and, partially, the information layer. The databases of atoms and ions, the data sources database, the database of properties of ion spectroscopy solutions and the bibliographic database were created. The first phase finishes with the construction of a prototype data layer for the existing VALD and Spectr-W^3 atomic and ionic spectroscopy ISs, using the iron atom and multi-charged iron ions as examples.

Table 1.

Database	VALD	Spectr-W^3	W@DIS
Substances	-	- (+)	+
Bibliography	-	+ (+)	+
Data sources	-	- (+)	+
Properties of solutions to spectroscopic problems	-	-	+
Ontologies	-	-	+

Table 1 lists the main databases of the W@DIS IS and shows the presence or absence of similar databases in the VALD and Spectr-W^3 IS. The round brackets with a plus sign indicate the data-bases created in the prototype version of new information system on ion spectroscopy.

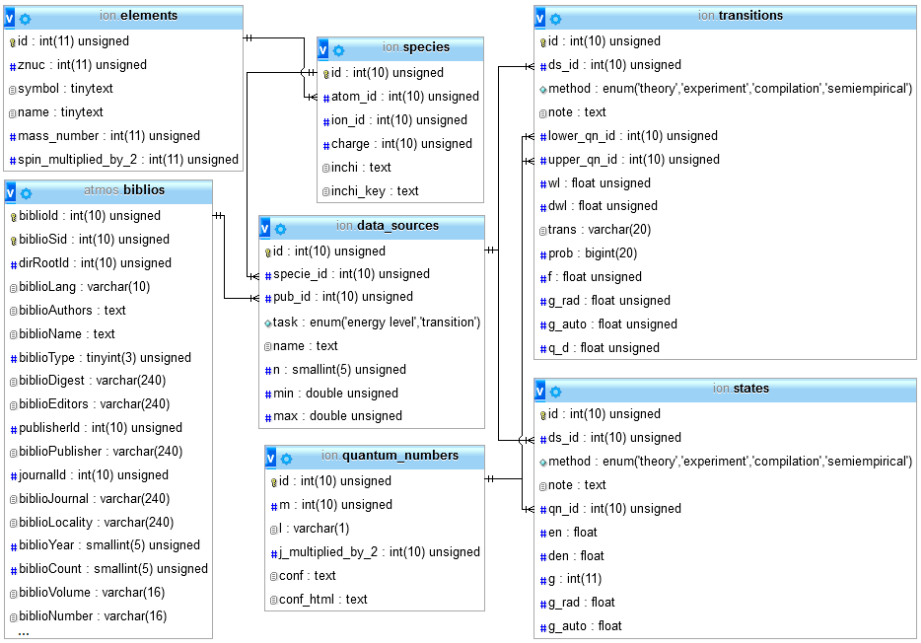


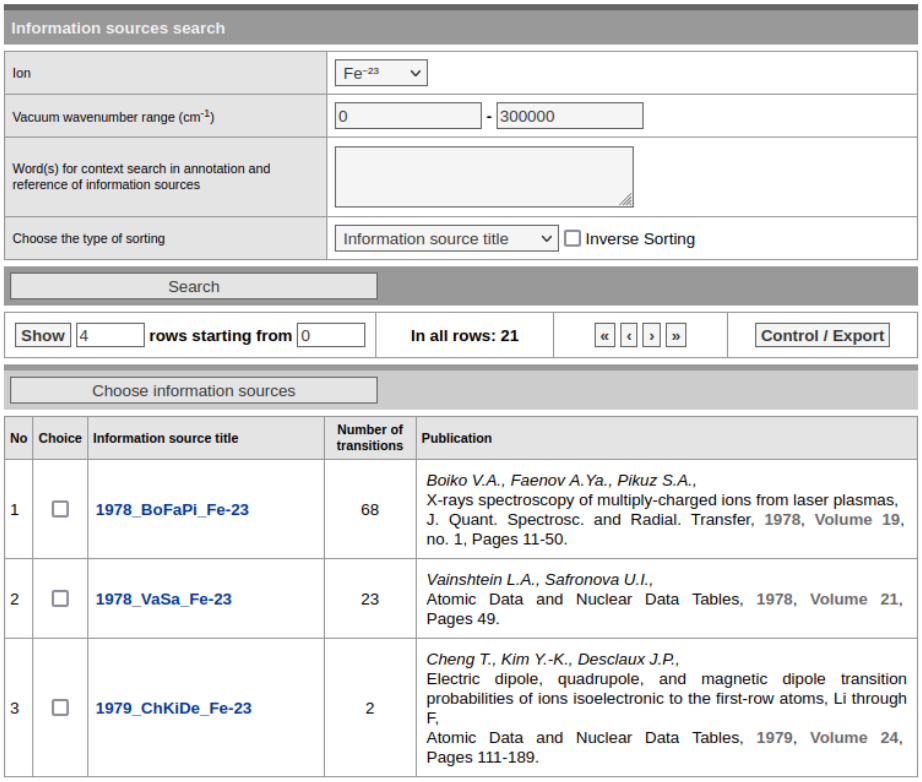
Fig. 1. Database structure for the new version of the ion spectroscopy IS.

Fig. 1 demonstrates the database structure for the new version of the information system on ion spectroscopy. It includes seven tables: *elements*, *species*, *data sources*, *biblios*, *states*, *quantum numbers* and *transitions*. Tables *states* and *transitions* contain the results of solving direct and inverse problems of finding energy levels and parameters characterizing transitions. The *data sources* table contain data on the substance for which the source contains data, the bibliographic reference to the publication from

which the data were extracted, the identifier of the spectral problem, the data source name, and the maximum and minimum values of the physical quantities in this data source.

Atomic spectroscopy data are a set of ranked alternative records used to synthesize unique expert data. Each record may contain data from different publications. The relationship between the concepts of "record" and "data source" is obvious: A “record” is a “composite data source”, i.e., a data source containing data extracted from more than one publication.

The transition from records (VALD IS) to data sources (W@DIS IS) requires the creation of applications that W@DIS does not currently have. Analogous to the notion of “record” in W@DIS is the notion of “composite data source”, which is used to describe expert data sources.



Information sources search

Ion	Fe^{-23}
Vacuum wavenumber range (cm^{-1})	0 - 300000
Word(s) for context search in annotation and reference of information sources	
Choose the type of sorting	Information source title ☐ Inverse Sorting

Search

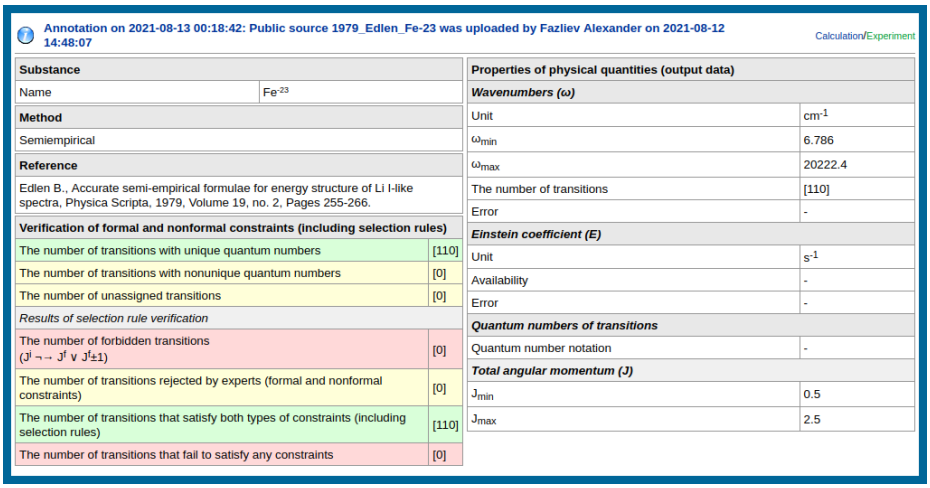
Show 4 rows starting from 0 | In all rows: 21 | « ‹ › » | Control / Export

Choose information sources

No	Choice	Information source title	Number of transitions	Publication
1	☐	1978_BoFaPi_Fe-23	68	*Boiko V.A., Faenov A.Ya., Pikuz S.A.,* X-rays spectroscopy of multiply-charged ions from laser plasmas, J. Quant. Spectrosc. and Radial. Transfer, 1978, Volume 19, no. 1, Pages 11-50.
2	☐	1978_VaSa_Fe-23	23	*Vainshtein L.A., Safronova U.I.,* Atomic Data and Nuclear Data Tables, 1978, Volume 21, Pages 49.
3	☐	1979_ChKiDe_Fe-23	2	*Cheng T., Kim Y.-K., Desclaux J.P.,* Electric dipole, quadrupole, and magnetic dipole transition probabilities of ions isoelectronic to the first-row atoms, Li through F, Atomic Data and Nuclear Data Tables, 1979, Volume 24, Pages 111-189.

Fig. 2. Fragment of the results of the data source search for Fe^{-23} ion.

Fig. 2 demonstrates the interface for the program module implementing search for data sources containing ion transitions in the prototype ion spectroscopy information system.



Annotation on 2021-08-13 00:18:42: Public source 1979_Edlen_Fe-23 was uploaded by Fazliev Alexander on 2021-08-12 14:48:07

Calculation/Experiment

Substance	
Name	Fe^{23}

Method
Semiempirical

Reference
Edlen B., Accurate semi-empirical formulae for energy structure of Li I-like spectra, Physica Scripta, 1979, Volume 19, no. 2, Pages 255-266.

Verification of formal and nonformal constraints (including selection rules)	
The number of transitions with unique quantum numbers	[110]
The number of transitions with nonunique quantum numbers	[0]
The number of unassigned transitions	[0]
Results of selection rule verification	
The number of forbidden transitions ($J^i \neg\rightarrow J^f \vee J^i \pm 1$)	[0]
The number of transitions rejected by experts (formal and nonformal constraints)	[0]
The number of transitions that satisfy both types of constraints (including selection rules)	[110]
The number of transitions that fail to satisfy any constraints	[0]

Properties of physical quantities (output data)	
Wavenumbers (ω)	
Unit	cm^{-1}
ω_{min}	6.786
ω_{max}	20222.4
The number of transitions	[110]
Error	-
Einstein coefficient (E)	
Unit	s^{-1}
Availability	-
Error	-
Quantum numbers of transitions	
Quantum number notation	-
Total angular momentum (J)	
J_{min}	0.5
J_{max}	2.5

Fig. 3. Metadata representation for the data source.

Fig. 3 shows *the 1979_Elden_Fe-23* data source properties.

The second stage in the development of atomic and ionic information systems entails the creation of an information and knowledge layer in them. The third stage in the development of information systems of the Russian segment is devoted to the creation of decision support systems for the construction and modification of personalized expert spectral data.

5 VAMDC Consortium

The VAMDC digital infrastructure currently contains 38 heterogeneous databases, containing atomic and molecular data [34]. It emerged from a European infrastructure project in 2009-2012 with ambitions to integrate data provider resources for open access to spectroscopy resources [35]. Later, with the expansion of VAMDC information resources and the inclusion of additional data providers, the VAMDC consortium was created. Its resources and activities are described in [36]. The current state of the resources (data and applications) and the prospects for the development of VAMDC resources are described in paper [34], which was published last year.

The infrastructure components are nodes, a metadata registry, a portal and a substance database. A data node is a database, pre-existing or created for VAMDC purposes, encapsulated in specialized software that implements a web service. The VAMDC metadata registry is a list of properties of data nodes. Applications use the registry to decide which databases should be queried and then locate services for those databases on the Internet. The VAMDC portal manages the infrastructure elements to provide seamless access to the interconnected databases. Through this unique interface, the user can query any part of the VAMDC infrastructure database and retrieve data in the common format of shared VAMDC-XSAMS files. A central-

ized chemical repository was created to solve the problem of substance identification. When designing the infrastructure, VAMDC developers follow the FAIR Principles.

The uniform description of data in XSAMS has enabled the interoperability of databases, but the problems of data quality analysis have not yet been fully solved, as well as the problems of presenting the results of this analysis in an explicit form.

The tasks formulated in the previous section of the article are being recognized and the graphical resources of the W@DIS system are mentioned in [34] (note that the Spectr-W^3 IS contains a collection of emission spectrograms (or densitograms) of plasma objects glow in different X-ray spectrum ranges). The organization of the semantic search for spectral resources is also on the list of problems to be solved.

6 Conclusion

The paper highlights the main information tasks associated with the three stages of the knowledge life cycle for the part of spectroscopy with intensive data production and use. A brief overview of solved and planned VAMDC information tasks is given. Using the example of the W@DIS information system, it is shown which partial solutions to the main information problems in spectroscopy have been obtained. The information systems VALD and Spectr-W^3 are briefly characterized, the prospects of their development and the first results of the plans for the development of the Russian VAMDC segment are described. This work is largely declarative in nature, describing the development prospects for five years.

Acknowledgements. The authors are grateful to Pakhomov Yu., Ryabchikova T.A., and Loboda P.A. for discussing the prospects of the VAMDC segment and providing test data as well as documentation that allowed understanding the data structures of atomic and molecular data collections in short time. This work was supported by the Ministry of Science and Higher Education of the Russian Federation.

References

1. De Roure, D., Jennings, N.R., Shadbolt, N.R.: The Semantic Grid: A Future e- Science Infrastructure, chapter 17. In book: Grid Computing: Making the Global Infrastructure a Reality, pp.437 – 470.
2. Berners-Lee, T., Hendler, J., Lassila, O.: The Semantic Web. Scientific American, May 17, 2001.
3. Gordon, I.E., Rothman, L.S., Hill, C., Kochanov, R.V., Tan, Y., Bernath, P.F., Birk, M., Boudon, V., Campargue, A., Chance, K.V., Drouin, B.J., et al.: The HITRAN2016 molecular spectroscopic database. Journal of Quantitative Spectroscopy and Radiative Transfer, vol. 203, pp. 3-69.
4. Jacquinet-Husson, N., Armante, R., Scott, N.A., Chédin, A., Crépeau, L., Boutammine, C., Bouhdaoui, A., Crevoisier, C., Capelle, V., Boonne, C., et al.: The 2015 edition of the GEISA spectroscopic database. Journal of Molecular Spectroscopy, vol. 327, pp. 31-72.

5. Bykov, A.D., Voronin, B.A., Kozodoev, A.V., Lavrent'ev, N.A., Rodimova, O.B., Fazliev, A.Z.: Information system for molecular spectroscopy. 1. Structure of information resources. Atmospheric and oceanic optics, 2004, vol. 17, no. 11, pp. 816-820.
6. Praczyk, P.A.: Management of scientific images: an approach to the extraction, annotation and retrieval of figures in the field of High Energy Physics. Thesis Doctoral, Universidad de Zaragoza (2013).
7. Keller-Rudek, H., Moortgat, G.K., Sander, R., and Sörensen, R.: The MPI-Mainz UV/VIS spectral atlas of gaseous molecules of atmospheric interest. Earth System Science Data, vol. 5, pp. 365–373.
8. XSAMS: XML Schema for Atoms, Molecules and Solids, http://www.vamdc.eu/documents/vamdc-xsams-doc-1.0.
9. Privezentsev, A.I., Tsarkov, D.V., Fazliev, A.Z.: Computed knowledge base for description of information resources of molecular spectroscopy. 3. Basic and applied ontologies. Digital Library Journal, 2012, vol. 15(2).
10. Voronina, S S., Privezentsev, A.I., Tsarkov, D.V., Fazliev, A.Z.: An Ontological Description of States and Transitions in Quantitative Spectroscopy. Proc. of SPIE XX-th International Symposium on Atmospheric and Ocean Optics: Atmospheric Physics, 2014, vol. 9292, pp. 92920C.
11. Bykov, A.D., Kozodoev, A.V., Privezentsev, A.I., Sinitsa, L.N., Tonkov, M.V., Filippov, N.N., Fazliev, A.Z., Tretyakov, M.Yu.: Distributed information system on molecular spectroscopy. Proc. of SPIE, International Symposium on High Resolution Molecular Spectroscopy, 2006, vol. 6580, pp. 65800W.
12. Tennyson, J., Bernath, P.F., Brown, L.R., Campargue, A., Carleer, M.R., Császár, A.G., Gamache, R.R., Hodges, J.T., Jenouvrier, A., Naumenko, O.V., Polyansky, O.L., Rothman, L.S., Toth, R.A., Vandaele, A.C., Zobov, N.F., Daumont, L., Furtenbacher, T. Fazliev, A.Z., Gordon, I.E., Mikhailenko, S.N., Shirin, S.V.: IUPAC Critical Evaluation of the Rotational-Vibrational Spectra of Water Vapor. Part I. Energy Levels and Transition Wavenumbers for $H_2{}^{17}O$, $H_2{}^{18}O$. Journal of Quantitative Spectroscopy and Radiative Transfer, 2009, vol. 110, no. 9-10, pp. 573-596.
13. Privezentsev, A., Tsarkov, D., Fazliev, A., Tennyson, J.:, Computed Knowledge Base for Description of Information Resources of Water Spectroscopy. In: Evren Sirin, Kendall Clark (eds.) Proc. of the 7th International Workshop on OWL: Experiences and Directions (OWLED 2010), vol. 614.
14. Lavrentiev, N.A. Privezentsev, A.I. Fazliev, A.Z.: Computed knowledge base for description of information resources of molecular spectroscopy. 2. Data model of quantitative spectroscopy. Digital Library Journal, 2012, vol. 14(2) (in Russian).
15. Privezentsev, A.I., Tsarkov, D.V., Fazliev, A.Z.: Computed knowledge base for description of information resources of molecular spectroscopy. 3. Basic and applied ontologies. Digital Library Journal, 2012, vol. 15(2) (in Russian).
16. Fazliev, A., Privezentsev, A., Tsarkov, D., Tennyson, J.: Ontology-Based Content Trust Support of Expert Information Resources in Quantitative Spectroscopy. In book: Knowledge Engineering and the Semantic Web, Communications in Computer and Information Science, vol. 394, Springer Berlin Heidelberg.
17. Voronina, S S., Privezentsev, A.I., Tsarkov, D.V., Fazliev, A.Z.: An Ontological Description of States and Transitions in Quantitative Spectroscopy. Proc. of SPIE XX-th International Symposium on Atmospheric and Ocean Optics: Atmospheric Physics, 2014, vol. 9292, pp. 92920C.
18. Lavrentiev, N.A., Privezentsev, A.I., Fazliev, A.Z., Tabular and Graphic Resources in Quantitative Spectroscopy. In: Manolopoulos Y., Stupnikov S. (eds.) Data Analytics and

Management in Data Intensive Domains. DAMDID/RCD. 2018. Communications in Computer and Information Science, vol 1003, p. 55-69.
19. Lavrentiev, N., Privezentsev, A., Fazliev, A.: Applied Ontologies for Managing Graphic Resources in Quantitative Spectroscopy. In: Elizarov A., Novikov B., Stupnikov S. (eds.) Data Analytics and Management in Data Intensive Domains, DAMDID/RCDL 2019. Communications in Computer and Information Science, vol. 1223, pp. 82-93.
20. Akhlyostin, A.Y., Lavrentiev, N.A., Privezentsev, A.I., Fazliev, A.Z.: Computed knowledge base for describing information resources in molecular spectroscopy. 5. Expert data quality. Digital Library Journal, 2013, vol. 16(4) (in Russian).
21. Lavrentyev, N.A., Makogon, M.M., Fazliev, A.Z.: Comparison of the HITRAN and GEISA Spectral Databases Taking into Account the Restriction on Publication of Spectral Data. Atmospheric and Oceanic Optics, 2011, vol. 24, no. 5, pp. 436–451.
22. Kozodoev, A.V., Privezentsev, A.I., Fazliev, A.Z., Filippov, N.N.: Systematization of Sources of Data on Spectral Line Parameters for the CO_2 Molecule and Its Isotopologues in the W@DIS Information System. Atmospheric and Oceanic Optics, 2018, vol. 31, no. 2, pp. 201–215.
23. Akhlestin, A., Lavrentiev, N., Kozodoev, A., Kozodoeva, E., Privezentsev, A., Fazliev, A.: Improvement of the Data Quality Assessment Procedure in Large Collections of Spectral Data. CEUR Supplementary Proceedings of the XXII International Conference on Data Analytics and Management in Data Intensive Domains (DAMDID/RCDL 2020), 2020. vol. 2790, pp. 263-279.
24. Akhlestin, A.Yu., Lavrentiev, N.A., Rodimova, O.B., Fazliev, A.Z.: The continuum absorption: trust assessment of published graphical information. Proc. SPIE 11208, 25-th International Symposium on Atmospheric and Ocean Optics: Atmospheric Physics, pp. 112080P.
25. Kozodoev, A.V., Fazliev, A.Z.: Information system for molecular spectroscopy. 2. Array operations for transformation of data on spectral line parameters. Atmospheric and oceanic optics, 2005, vol. 18, no. 09, pp. 680-684.
26. Kozodoev, A.V., Kozodoeva, E.M.: Universal module "unary operations" in the IS "Molecular Spectroscopy". Vestnik NSU Series: Information Technologies, 2015, vol. 13, no. 1, pp. 46-54 (in Russian).
27. Kozodoev, A.V., Kozodoeva, E.M: Binary operations in the information system "Molecular spectroscopy". Vestnik NSU Series: Information technologies., 2018., vol. 16, no. 2, pp. 70-77 (in Russian).
28. Akhlyostin, A.Yu. Kozodoev, A.V. Lavrentiev, N.A. Privezentsev, A.I. Fazliev, A.Z.: Computed knowledge base for description of information resources of molecular spectroscopy. 4. Software. Digital Library Journal, 2012, vol. 15(3) (in Russian).
29. Piskunov, N., Kupka, F., Ryabchikova, T.A., Weiss, W.W., Jeffery, S.: VALD: The Vienna Atomic Line Data Base, Astronomy and Astrophysics Supplement Series, vol. 112, p. 525.
30. Skobelev, Yu., Loboda, P.A., Gagarin, S.V., Ivliev, S.V., Kozlov, A.I., Morozov, S.V., Pikuz, S.A., Pikuz, T.A., Popova, V.V., Faenov, A.Ya.: The Spectr-W^3 database on the spectroscopic properties of atoms and ions, Optics and Spectroscopy, 2016, vol. 120(4), pp. 507-514.
31. Akhlyostin, A., Apanovich, Z., Fazliev, A., Kozodoev, A., Lavrentiev, N., Privezentsev, A., Rodimova, O., Voronina, S., Csaszar, A.G., Tennyson, J.: The current status of the W@DIS information system. In: G.G. Matvienko, O.A. Romanovskii (eds.) Proc. SPIE of 22-nd International Symposium Atmospheric and Ocean Optics: Atmospheric Physics, vol.10035, pp. 100350D.

32. Kupka, F., Piskunov, N., Ryabchikova, T.A., Stempels, H.C., Weiss, W.W.: VALD-2: Progress of the Vienna Atomic Line Data Base. Astron. Astrophys. Suppl., vol. 138, pp. 119-133.
33. Pakhomov, Yu., Piskunov, N., Ryabchikova, T.: VALD3: current developments, 30 Oct 2017, arXive-1710.10854v1 [astro-ph.IM].
34. Albert, D., Antony, B.K., Ba, Y.A., Babikov, Y.L., Bollard, P., Boudon, V., Delahaye, F., Del Zanna, G., Dimitrijevíc, M.S., Drouin, B.J., Dubernet, M.-L., et al.: A Decade with VAMDC: Results and Ambitions, Atoms 2020, vol. 8(4), p. 76.
35. Dubernet, M.L., Boudon, V., Culhane, J.L., Dimitrijevic, M.S., Fazliev, A.Z, Joblin, C., Kupka, F., Leto, G., Le Sidaner, P., Loboda, P.A., et al: Virtual atomic and molecular data centre. J. Quant. Spectrosc. Radiat. Transf., 2010, vol. 111, pp. 2151–2159.
36. Dubernet, M.L., Antony, B., Ba, Y.A., Babikov, Y., Bartschat, K., Boudon, V., Braams, B., Chung, H.K., Daniel, F., Delahaye, F., et al.: The Virtual Atomic and Molecular Data Centre (VAMDC) Consortium. J. Phys. B At. Mol. Opt. Phys., 2016, vol. 49, pp. 074003.