A Decade with VAMDC: Results and Ambitions

Damien Albert, Bobby K Antony, Yaye Awa Ba, Yuri L Babikov, Philippe Bollard, Vincent Boudon, Franck Delahaye, Giulio Del Zanna, Milan S Dimitrijević, Brian J Drouin, et al.

MDPI

Article

A Decade with VAMDC: Results and Ambitions

Damien Albert[8], Bobby K. Antony[2], Yaye Awa Ba[1], Yuri L. Babikov[3,4], Philippe Bollard[10], Vincent Boudon[6], Franck Delahaye[1], Giulio Del Zanna[9], Milan S. Dimitrijević[1,11], Brian J. Drouin[13], Marie-Lise Dubernet[1], Felix Duensing[34], Masahiko Emoto[12], Christian P. Endres[14], Alexandr Z. Fazliev[3], Jean-Michel Glorian[19], Iouli E. Gordon[16], Pierre Gratier[35], Christian Hill[7], Darko Jevremović[11], Christine Joblin[19], Duck-Hee Kwon[40], Roman V. Kochanov[3,4], Erumathadathil Krishnakumar[21], Giuseppe Leto[22], Petr A. Loboda[15,23], Anastasiya A. Lukashevskaya[3], Oleg M. Lyulin[3], Bratislav P. Marinković[25], Andrew Markwick[26], Thomas Marquart[17], Nigel J. Mason[27], Claudio Mendoza[28], Tom J. Millar[29], Nicolas Moreau[1], Serguei V. Morozov[15], Thomas Möller[41], Holger S. P. Müller[41], Giacomo Mulas[19,30], Izumi Murakami[12,20], Yury Pakhomov[31], Patrick Palmeri[32], Julien Penguen[36], Valery I. Perevalov[3], Nikolai Piskunov[17], Johannes Postler[34], Alexei I. Privezentsev[3], Pascal Quinet[32,37], Yuri Ralchenko[39], Yong-Joo Rhee[44], Cyril Richard[6], Guy Rixon[38], Laurence S. Rothman[16], Evelyne Roueff[1] , Tatiana Ryabchikova[31], Sylvie Sahal-Bréchot[1], Paul Scheier[34], Peter Schilke[41], Stephan Schlemmer[41], Ken W. Smith[29], Bernard Schmitt[10], Igor Yu. Skobelev[23,24], Vladimir A. Srećković[25], Eric Stempels[17], Serguey A. Tashkun[3], Jonathan Tennyson[18], Vladimir G. Tyuterev[4,42], Charlotte Vastel[19], Veljko Vujčić[11,43], Valentine Wakelam[35], Nicholas A. Walton[38], Claude Zeippen[1] and Carlo Maria Zwölf[1]

1 LERMA, Observatoire de Paris, PSL Research University, CNRS, Sorbonne University, UPMC Univ Paris 06, 5 Place Janssen, 92190 Meudon, France
2 Indian Institute of Technology (Indian School of Mines) Dhanbad, India
3 V.E. Zuev Institute of Atmospheric Optics, Siberian Branch, Russian Academy of Sciences, Zuev Sq. 1, Tomsk 634055, Russia
4 Tomsk State University, Tomsk 634050, Russia
5 Department of Physics and Astronomy, Drake University, Des Moines, Iowa 50311, USA
6 Laboratoire Interdisciplinaire Carnot de Bourgogne, UMR 6303 CNRS-Univ. Bourgogne Franche-Comté, 9 Avenue Alain Savary, BP 47 870, F-21078 DIJON Cedex, France
7 Nuclear Data Section, Division of Physical and Chemical Sciences, International Atomic Energy Agency (IAEA), Vienna International Centre, A-1400 Vienna, Austria
8 Université Grenoble Alpes, CNRS, OSUG, F-38000 Grenoble, France
9 DAMTP, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, UK
10 Université Grenoble Alpes, CNRS, IPAG, F-38000 Grenoble, France
11 Astronomical Observatory, Volgina 7, 11060 Belgrade, Serbia
12 National Institute for Fusion Science, National Institutes of Natural Sciences, Toki, Gifu 509-5292, Japan
13 California Institute of Technology, Jet Propulsion Laboratory, 4800 Oak Grove Dr. Pasadena, CA 91109, USA
14 Max-Planck-Institut für extraterrestrische Physik, Giessenbachstrasse, 85748 Garching, Germany
15 Russian Federal Nuclear Centre All-Russian Institute of Technical Physics (RFNC VNIITF), Snezhinsk, 456770 Russia
16 Center for Astrophysics | Harvard & Smithsonian, Atomic and Molecular Physics Division, MS50, 60 Garden St, Cambridge MA 02138-1516 USA
17 Department of Physics and Astronomy, Uppsala University, Box 516, 751 20 Uppsala, Sweden
18 Department of Physics and Astronomy, University College London, London WC1E 6BT, UK
19 Institut de Recherche en Astrophysique et Planétologie, Université de Toulouse (UPS), CNRS, CNES, 9 Av. du Colonel Roche, 31028 Toulouse Cedex 4, France
20 Department of Fusion Science, The Graduate University for Advances Studies, SOKENDAI, Toki, Gifu 509-5292, Japan
21 Raman Research Institute, C V Raman Avenue, Bangalore 560080, India
22 INAF - Osservatorio Astrofisico di Catania, Via S. Sofia 78, I-95123 Catania, Italy
23 National Research Nuclear University — Moscow Engineering Physics Institute (MEPhI), Moscow, 115409 Russia

[24] Joint Institute for High Temperatures, Russian Academy of Sciences, Moscow, 141570 Russia
[25] Institute of Physics Belgrade, University of Belgrade, P.O. Box 57, 11001 Belgrade, Serbia
[26] Jodrell Bank Centre for Astrophysics, School of Physics and Astronomy, University of Manchester, Oxford Road, Manchester M13 9PL, UK
[27] School of Physical Sciences, Ingram Building, University of Kent, Canterbury CT2 7NH, United Kingdom
[28] Department of Physics, Western Michigan University, Kalamazoo, MI 49008, USA
[29] School of Mathematics and Physics, Queen's University Belfast, University Road, Belfast BT7 1NN, Northern Ireland
[30] Istituto Nazionale di AstroFisica - Osservatorio Astronomico di Cagliari, via della Scienza 5, 09047 Selargius (CA), Italy
[31] Institute of Astronomy Russian Academy of Sciences, Pyatnitskaya 48, 119017 Moscow, Russia
[32] Physique Atomique et Astrophysique, Université de Mons, B-7000 Mons, Belgium
[33] ABB Corporate Research, Segelhofstrasse 1K, 5405 Baden-Dättwil, Switzerland
[34] Institute for Ion Physics and Applied Physics, University of Innsbruck, Technikerstr. 25/3, A-6020 Innsbruck, Austria
[35] Laboratoire d'astrophysique de Bordeaux, Univ. Bordeaux, CNRS, B18N, allée Geoffroy Saint-Hilaire, 33615 Pessac, France
[36] POREA, Observatoire Aquitain des Sciences de l'Univers, Univ. Bordeaux, CNRS, IRSTEA, Univ. La Rochelle, B18N, allée Geoffroy Saint-Hilaire, 33615 Pessac, France
[37] IPNAS, Université de Liège, B-4000 Liège, Belgium
[38] Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge, CB3 0HA, UK
[39] Atomic Spectroscopy Group, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA
[40] Nuclear Data Center, Korea Atomic Energy Research Institute, Daejeon, 34057, Republic of Korea
[41] I. Physikalisches Institut, Universität zu Köln, Zülpicher Str. 77, 50937 Köln, Germany
[42] GSMA, UMR CNRS 7331, Université de Reims, France
[43] Faculty of Organizational Sciences, University of Belgrade, Jove Ilića 33, 11000 Belgrade, Serbia
[44] Center for Relativistic Laser Science, Institute for Basic Science, Gwang-Ju 61005 Republic of Korea

Version October 2, 2020 submitted to Atoms

Abstract: This paper presents an overview of the current status of the VAMDC e-infrastructure including the current status of the VAMDC connected (or to be connected) databases, updates on the latest technological development within the infrastructure and a presentation of some application tools that make use of the VAMDC e-infrastructure. We analyse the past 10 years of VAMDC development and operation, and assess their impact both on the field of atomic and molecular (A&M) physics itself and on heterogeneous data management in international cooperation. The highly sophisticated VAMDC infrastructure and the related databases developed over this long term make them a perfect resource of sustainable data for future applications in many fields of research. However, we also discuss the current limitations that prevent VAMDC from becoming the main publishing platform and the main source of A&M data for user communities, and present possible solutions under investigation by the consortium. Several user applications examples are presented, illustrating the benefits of VAMDC in current research applications, which often need the A&M data from more than one database. Finally, we present our vision for the future of VAMDC.

Keywords: scientific databases; atomic and molecular data; interoperability; FAIR principles; open access

1. Introduction

The Virtual Atomic and Molecular Data Centre (VAMDC) has been developed to interconnect atomic and molecular databases, thus providing a single location where users can access atomic and molecular (A&M) data. The VAMDC portal currently provides access to 38 databases containing a wide range of data from atomic spectroscopy (VALD, section 2.1.2) to PAH theoretical data (PAH,

section 2.2.2). The paper presents the current status of the VAMDC project and its underlying databases, and gives some examples of the exploitation of VAMDC by some exemplar user communities; we describe open issues in the project and our vision for future developments.

The VAMDC project[1] originated as a European Union framework 7 (FP7) research infrastructure project [1] that was funded between 2009 and 2012, and subsequently extended through the SUP@VAMDC project[2] [2] between 2012 and 2014. Since 2014 the VAMDC Consortium has operated as an independent body comprising some 35 research groups. This consortium has continued to operate and develop the data centre, as well as pursuing other related objectives in data science. The VAMDC consortium published its original aims in 2010 [1], which were updated in 2016 following the formal launch of the independent consortium [3]. The public entry of the VAMDC consortium is its public website [3] that provides an access to data and documentation, as well as consortium-related information such as current membership and how to join us [4].

2. Current Status of VAMDC Connected Databases

The VAMDC nodes are distributed such that they are located at the members' and partners' sites. At present most of the databases included in the VAMDC e-infrastructure are databases used primarily in astrophysics. The common output format is the XML Schema for Atoms, Molecules and Solids (XSAMS) (see section 3.1.1) that uses the tree-structured form of the data model. All of the current VAMDC databases are listed on the VAMDC portal[5] which also provides a short introduction to the database, a link to its public graphical interface and contact details of a member of that database team to whom inquiries may be directed. In addition, it is possible to list the atoms and the molecules that can be queried from those databases.

Table 1 provides a summary of the 38 databases currently accessible via VAMDC. It includes eight databases that are in the process of joining the consortium. It should be noted that a database might be off-line for a period of time due to maintenance issues.

Among the databases already present in the VAMDC e-infrastructure in 2016 and cited in our last publication [3], some have evolved with respect to their data content, functionalities and internal structure, and improved their interoperability with the VAMDC ecosystem. Some of them have not evolved, and two have been disconnected from the VAMDC ecosystem. Ten databases have been added. Among these new databases is a new node at NIFS in Japan.

2.1. Evolution of VAMDC Nodes/Connected Databases since 2016

Each individual database incorporated in VAMDC will now be discussed highlighting recent developments and improvements.

2.1.1. NIFS Databases

The National Institute for Fusion Science (NIFS) has compiled and developed an extensive atomic and molecular numerical database on collision processes which has been open for public access[6] since 1997. A data compilation on collisional cross sections for hydrogen, its isotopes and helium [41,42] was initiated by a collaborative working group of atomic and plasma researchers from Japanese universities, organised at the Institute of Plasma Physics, Nagoya University, in 1973. Subsequent compilations were published in IPPJ-AM reports in 1977–1989 and as NIFS-DATA reports since 1989

1 http://www.vamdc-project.vamdc.eu/
2 http://www.sup-vamdc.vamdc.org/
3 http://www.vamdc.org
4 http://www.vamdc.org/structure/how-to-join-us/
5 https://portal.vamdc.eu/vamdc_portal/nodes.seam
6 http://dbshino.nifs.ac.jp

Table 1. Table of databases connected to VAMDC (C) and to be connected (TC). User fields are labelled as follows for applications in astrophysics and planetary science: stellar physics (STEL), solar physics (SOL), interstellar medium (ISM), earth (E), planets (PL), exoplanets (EXO), brown dwarfs (BDW), comets (C). Partnerships are indicated as Full Members (FM) according to the VAMDC Consortium MoU and partners following the VAMDC technical quality chart only (P).

Database	Data Classification	Applications
NIFS Database[c]	atomic, molecular processes [4,5]	STEL, SOL, plasma, fusion
VALD[b]	atomic, molecular linelists [6]	STEL, SOL
VALD-Moscow[b]	atomic, molecular linelists [6]	STEL, SOL (subset) 2.1.2
NIST ASD[c]	atomic lineslists [7]	STEL, ISM
Spectr-W^3 [c]	atomic lineslists and collisions [8]	STEL, SOL, plasma, fusion
CHIANTI[b]	atomic lineslists and collisions [9]	SOL
TIPbase[b]	atomic linelists and collisions [10]	STEL, SOL, plasma
TOPbase[b]	atomic linelists and collisions [11]	STEL, SOL, plasma
Stark-B[b]	atomic line shifts, broadening [12,13]	STEL, plasma
CDMS[b]	molecular linelists [14]	ISM, E, C
JPL[c]	molecular linelists [15]	ISM, E, C
HITRAN[c]	molecular linelists, broadening coefficients [16]	E, PL, EXO
S&MPO[b]	O_3 linelists [17]	E, EXO
MeCaSDa[b]	CH_4 linelists [18]	E, EXO, PL, DBW
ECaSDa[b]	Ethene calculated linelists [18]	E, PL
TFMeCaSDa[b]	Tetrafluoro-Methane calculated linelists [18]	E
SHeCaSDa[b]	Sulfur Hexafluoride calculated linelists [18]	E
GeCaSDa[b]	GeH_4 linelists [18]	PL
RuCaSDa[b]	RuO_4 linelists [18]	Nuclear industry
TFSiCaSDa[b]	SiF_4 linelists [18,19]	E
UHeCaSDa[b]	UF_6 line lists ([a])	Nuclear industry
CDSD-296[b]	CO_2 linelists [3,20]	E, PL, EXO, BDW
CDSD-1000[b]	CO_2 linelists [3]	E, PL, EXO, BDW
CDSD-4000[b]	CO_2 linelists [21]	E, PL, EXO, BDW
NOSD-1000[b]	N_2O linelists [22]	E, PL, EXO
NDSD-1000[b]	NO_2 linelists [23]	E, PL, EXO
ASD-1000[b]	C_2H_2 linelists [24]	E, PL, EXO
SESAM[b]	VUV small molecules linelists [13]	ISM, STELL
W@DIS[b]	atmospheric molecule data sources [25]	E, PL
KIDA[b]	chemical kinetics [26,27]	ISM, PL
UDfA[b]	chemical kinetics [28]	ISM, PL
BASECOL[b]	molecular collisions [29,30]	ISM, C
MOLD[b]	photo-dissociation cross sections [31,32]	STEL
BeamDB[c]	molecule/atom-electron cross-sections [33]	Plasma, radiation damage
IDEADB[c]	dissociative electron collisions [34]	PL, EXO, ISM, radiation damage
GhoSSt[b]	solid spectroscopy data [35]	ISM, PL
LASp[b]	solid spectroscopy data [3]	ISM, PL
PAH[b]	PAH theoretical Data [3,36]	ISM, PL, E
ExoMolOP[e]	molecular opacities [37]	EXO, DBW, STEL, E
SSHADE[e]	solid spectroscopy data [35]	E, C, EXO, ISM, PL
AMBDAS[d]	collisions in plasmas (bibliography) ([a])	Nuclear Fusion
DESIRE[d]	radiative data for sixth row elements [3,38]	STEL, SOL, plasmas
DREAM[d]	radiative data for rare earths [39]	STEL, SOL, plasmas, lighting industry
IAMDB[d]	A+M spectroscopy, atomic collision ([a])	Astrophysics, other
PEARL[d]	atomic processes [40]	STEL, SOL, plasma, fusion
Clusters[d]	cluster size distributions, condensation ([a])	ISM, P, biology

[a] Paper in preparation [b] (FM, C) [c] (P, C) [d] (P, TC) [e] (FM, TC)

covering electron-impact ionisation and excitation cross sections, heavy-particle collision cross sections and plasma-wall interaction properties.

The first retrieval and display database system, the Atomic and Molecular Data Interactive System (AMDIS), was constructed in a mainframe computer for electron-impact ionisation and excitation cross sections of atoms and atomic ions in 1981 [43], with the database system being subsequently extended to include other collision processes. Currently the database system consists of eight sub-databases accessible via internet, full details are described by Murakami et al. [44].

Numerical data of collision cross sections and rate coefficients of various collision processes are compiled mainly from the published literature as well as bibliographic data and additional information such as experimental or theoretical methods used to derive the data are attached. Initially, data were collected for elements that were important mainly for fusion science; later the database was extended to include many more atoms and molecules applicable to various plasma applications.

Usually, several data sets obtained from different publications are stored for the same collision processes. Users can compare those data sets to check their reliability. Data can be queried by elements, charge states or various other query fields such as initial and final states, author, published year, and retrieved data are displayed as a numerical table or as a graph.

One of the sub-databases, AMDIS-IONIZATION has been available from the VAMDC portal with search functionality by element since 2017 [4]. AMDIS-IONIZATION has electron-impact ionisation cross sections and rate coefficients for atoms and atomic ions. Data sets of cross sections with collision energy or rate coefficients with electron temperature for single or multiple ionisation processes are available. Sometimes initial and final electronic states are available according to the original data source. In the future it is planned to add AMDIS-EXC (electron-impact excitation cross sections and rate coefficients for atoms and atomic ions) to the VAMDC portal.

2.1.2. VALD

The Vienna Atomic Line Database (VALD) compilation is a critically assessed collection of radiative transition data aimed primarily at the stellar astrophysics community. VALD contains data on energy levels, wavelengths, oscillator strengths and line broadening parameters for nearly all stable elements in the periodic table and for a few simple molecules. For atoms, VALD includes 6 ionisation stages. The VALD interface allows a search of the whole data collection or on a restricted subset that is based on high-precision laboratory measurements. The Moscow VAMDC node serves high-precision data only, while the Uppsala VAMDC node provides the most complete VALD data set.

VALD also includes atomic data for individual rare-earth elements/ions, also provided by the DREAM database[7] [45] (see Section 2.3.4). These atomic data are extracted via VALD extraction tools following an adopted quality ranking. Therefore, only a part of DREAM data collection is accessible through the VALD-VAMDC nodes.

VALD is under continuous development both in terms of functionality and data content. New data sets are systematically compared with existing ones and with experimental data in order to establish data quality ranking. For any given transition, all data elements are merged according to the ranking offering the best quality result for the end user while preserving all relevant bibliographic references. Recent developments include the gradual introduction of energy level and spectral line data for individual isotopes and isotopologues. Currently the work is complete for Li, Ca, Ti, Cu, Ga, Ba, Eu atoms and for isotopologues of CN, TiO, C_2, CH, CO, OH, MgH and SiH. VALD now also provides data for the hyperfine-structure splitting of ^{6}Li I, ^{7}Li I, ^{23}Na I, ^{27}Al I, ^{27}Al II, ^{39}K I, ^{40}K I, ^{41}K I, ^{45}Sc I, ^{45}Sc II, ^{47}Ti I, ^{49}Ti I, ^{47}Ti II, ^{49}Ti II, ^{50}V I, ^{51}V I, ^{51}V II, ^{55}Mn I, ^{55}Mn II, ^{57}Fe I, ^{59}Co I, ^{59}Co II, ^{61}Ni I, ^{63}Cu I, ^{65}Cu I, ^{67}Zn I, ^{67}Zn I, ^{69}Ga I, ^{69}Ga I, ^{71}Ga I, ^{71}Ga II, ^{85}Rb I, ^{87}Rb I, ^{87}Sr II, ^{89}Y II, ^{93}Nb II, ^{95}Mo II, ^{97}Mo II, ^{127}I I, ^{127}I I, ^{135}Ba II, ^{137}Ba II, ^{139}La II, ^{141}Pr I, ^{141}Pr II, ^{151}Eu II, ^{153}Eu II, ^{155}Gd II, ^{157}Gd II, ^{159}Tb II,

[7] https://hosting.umons.ac.be/html/agif/databases/dream.html

^{165}Ho I, ^{165}Ho II, ^{171}Yb II, ^{173}Yb II, ^{175}Lu I, ^{175}Lu II, ^{176}Lu I, ^{181}Ta I, ^{181}Ta II, ^{191}Ir I, ^{193}Ir I, ^{203}Tl I, ^{205}Tl I and ^{209}Bi III. The latest developments of VALD and its interface to VAMDC have been presented in several refereed publications [46–49] and conference proceedings [50–52]. It should be noted that the VALD database is linked to the VAMDC Query Store (see Section 3.2.2).

2.1.3. NIST ASD

The Atomic Spectra Database (ASD) [7] at the National Institute of Standards and Technology (USA) contains critically evaluated atomic data including energy levels, radiative transition probabilities and oscillator strengths, ionisation potentials, and observed and accurately calculated wavelengths of spectral lines between the hard X-ray and infrared regions of spectra. As of July 2020, ASD provides data on more than 112,000 energy levels, 280,000 spectral lines and 118,000 radiative transition probabilities for elements from H ($Z = 1$) to Ds ($Z = 110$). The majority of the data on energies and spectral lines were collected and evaluated from experimental papers, and in most cases realistic uncertainties are provided as well. ASD is directly linked to the set of the NIST Atomic Spectroscopy Bibliographic Databases [53], which allows an immediate access to the original sources of data. In addition to a tabular output of data in various formats, ASD also offers rich graphical services, e.g. online-generated Grotrian diagrams of levels and transitions. ASD serves as the basis for the NIST LIBS (Laser-Induced Breakdown Spectroscopy) Database [54], which allows on-the-fly generation of realistic spectra for Saha–Boltzmann plasmas that are used for analysis of elemental abundances and diagnostics of terrestrial and astrophysical plasmas.

2.1.4. Spectr-W^3

The Spectr-W^3 project[8] is a long-term collaboration between the Russian Federal Nuclear Centre – All-Russian Institute of Technical Physics (RFNC-VNIITF) and the Joint Institute for High Temperatures of the Russian Academy of Sciences (JIHT RAS) [8]. At present, Spectr-W^3 is the largest available database providing information on spectral properties of multicharged ions. The database contains over 450,000 records of experimental, theoretical and compiled numerical data on ionisation potentials, energy levels, wavelengths, radiative and autoionisation widths, satellite-line intensity factors for free atoms and ions and also fitting parameters, and analytic formulae to represent electron-collisional cross sections and rates (optional). References to the original sources and comments on the methods of data acquisition, etc. are also provided. Since 2016, a new section of Spectr-W^3 providing graphical data on X-ray emission spectra (densitograms) recorded from various plasma sources has been made available to users. Densitogram graphic images are characterised with a set of fields enabling one to perform queries specifying the shortest and longest wavelengths of interest, the identification of up to five chemical elements and ionic isosequences, emission from which contributes to the recorded spectra, the upper and lower principal quantum numbers of optical single-electron transitions corresponding to the spectral lines pictured and the reference to the relevant publication. Densitogram database records are also supplied with comments elucidating details of the experimental measurements.

2.1.5. CHIANTI

CHIANTI[9], first released in 1996 [55], is a well-established atomic database and modelling code for optically thin plasmas. For a recent overview of the status of the database and future plans, see [9]. The database consists of a series of ASCII files with all the relevant atomic collisional and radiative rates (collisional excitation rates by electron and proton impact, transition probabilities, as well as theoretical/experimental wavelengths) required to calculate spectral line emissivities. The database

[8] http://spectr-w3.snz.ru
[9] www.chiantidatabase.org

originally only included ions of astrophysical importance but in recent versions data for minor ions and some neutrals have also been added.

Radiative transition rates and wavelengths from CHIANTI versions 6 and 7.1 were included in VAMDC. The CHIANTI database also has ionisation/recombination rates and various other types of data that are specific for the modelling codes but these were not included in VAMDC. CHIANTI version 8 [56] included new atomic rates for ions in the Li, B and Ne isoelectronic sequences, plus atomic data for several iron ions participating in emission of light from the solar corona. Excitation rates among all states have been included (for modeling of high-density plasma), which has increased the size of the database significantly. Version 8 also changed the format of some of the main ASCII files. CHIANTI version 9 [57] then introduced a significant change in the structure of the data for several ions to allow the calculation of the emissivities of satellite lines, with autoionising states and rates having been added to those of the bound levels.

These changes mean that a straightforward update to include all the changes of CHIANTI version 8 onwards within VAMDC could not be carried out, so at present the VAMDC portal only accesses CHIANTI version 7.1. CHIANTI version 10 is now in development and it is planned to integrate this version with the latest CHIANTI data into VAMDC.

2.1.6. TIPbase-TOPbase

The Iron Project (IP) and The Opacity Project (OP) databases provide energy levels/terms, radiative transitions probabilities, photoionisation cross sections and collision strengths for a large selection of ions in the range H to Ni. The TOPbase contains the OP data for radiative processes [11,58,59], and the TIPbase the IP data for collisional and radiative processes [10]. The data have been calculated using state-of-the-art computer programs developed and maintained by the IP–OP members including the *R*-matrix suite of codes [60,61] and the AUTOSTRUCTURE code [62]. These data are relevant for experiment analysis, theoretical comparisons and for various astrophysical or laboratory experimental applications. The OP radiative data are used to calculate monochromatic and mean opacities [63] required in stellar codes and for the analysis of experiments. The TOPbase and TIPbase are hosted at the University of Strasbourg in France [10]. Sets of new (there were no data of that type in the database before) and recalculated data (with a better methodology) for existing data are being implemented in TOPbase for Fe I (new), Fe II (new) and Fe XVII (recalculated) as well as all the Ni ions (new); other recalculated ions will follow. A new service for opacity table is available on the TIPTOP webserver [64] (but not yet accessible through VAMDC). It should be noted that the TIPbase-TOPbase databases are linked to the VAMDC Query Store (see Section 3.2.2).

2.1.7. Stark-B

Stark-B [12,65–67] is devoted to the modelling and spectroscopic diagnostics of various plasmas in astrophysics, laboratory experiments, laser equipment design, laser produced plasma analysis and inertial fusion research. It contains Stark widths and shifts of isolated lines of atoms and ions due to collisions with electrons, protons, ionised helium (the most important colliders for stellar atmospheres) and other ions. The data are calculated using the semi-classical perturbation [68] and modified semi-empirical methods [69]. Stark-B is continuously updated by adding new data and by introducing new facilities for their use. Under the website option ”Data history” there are “New datasets” and “Updated datasets“, which make available a description of newly added data including the date of import and details of the modifications for revised data. In order to enable inclusion of data from Stark-B in the computer codes used for stellar atmospheres modelling and other numerical calculations, it is possible to fit tabulated data with temperature. In particular, a fitting formula for interpolation of the displayed data for different temperatures has been derived [70]. Thus, to each

[10] http://cdsweb.u-strasbg.fr/OP.htx

table of Stark widths and shifts, an additional table has been added with the coefficients needed for the fitting formula. The latest updates are further described in the current issue [13]. It should be noted that the Stark-B database is linked to the VAMDC Query Store (see Section 3.2.2).

2.1.8. CDMS and JPL Spectral Line Catalog

Both these databases[11] provide spectral data for molecular species, which are or may be observed in various astronomical sources (usually) by radio astronomical means or for use in remote sensing. The JPL catalog [15,71,72] and the CDMS [14,72–74] have been previously described to some extent in the literature. Briefly, the database content is generally restricted to effective Hamiltonian predictions and associated assigned experimental data for quantum transitions with entry fields including line position with accuracy, intensity, lower state energy and quantum numbers. These restrictions facilitate transfer of high-precision laboratory data into comprehensive predictions within the associated range of quantum numbers, and reconciliation and/or extension of models from multiple laboratory studies. Separate entries exist for different isotopic species and usually also for different vibrational states. Updates and new entries to the spectral data are performed aperiodically, with focus on detection in the interstellar medium as well as candidates for terrestrial remote sensing target molecules.

The incorporation of these two databases in VAMDC, and thus the application of its standards (see Section 3.1.1), not only greatly simplified the interoperability between the databases and improved the readability of the content (e.g. quantum number format), but also provided additional features and information. A listing of energy levels as well as all files that have been used to generate the data are commonly provided. Partition functions are now tabulated for 110 temperatures up to 1000 K. In the context of VAMDC, a Python library (`vamdclib`) has been developed that allows queries of data via VAMDC protocols with the information being stored in a local SQLITE3 database for use in third-party applications (e.g. XCLASS, see below in the section of use cases). Both databases are linked to the VAMDC Query Store (Section 3.2.2), and thus support VAMDC's generation of digital object identifiers (DOIs) for individual queries that can be used to cite this data in publications.

Creating new entries or updating existing ones is an important part of the work for the CDMS. As of July 2020, there are 1020 entries in the CDMS, up from 808 four years ago. Recent activities have emphasised species that are or may be observable with the Atacama Large Millimeter/submillimeter Array (ALMA), the Northern Extended Millimeter Array (NOEMA) and similar facilities. Examples include metal-containing di- and triatomics, which are relevant for the circumstellar envelopes of late-type stars, small to moderately sized organics, including cyclic molecules, which were or may be detected in star-forming regions, and several radicals or cations. Some effort has been made to make vibrationally excited state data available since transitions pertaining to low-lying excited vibrational states have been observed for several small to mid-sized organics (< 15 atoms) in the dense and warm parts of star-forming regions. Very highly excited states of diatomics and some larger molecules, such as HCN, have been detected in particular in the envelopes of C- and O-rich late-type stars. A detailed account of these activities is planned in the near future.

2.1.9. HITRAN

The HITRAN molecular spectroscopic database [16] is a canonical compilation of molecular spectroscopic parameters that are required for the input into the radiative-transfer codes. HITRAN provides the best unique parameter values assimilated from both experimental and theoretical studies. Before adapting data to the database, the HITRAN group performs independent evaluations (controlled laboratory, atmospheric retrievals and theoretical analyses) wherever possible. While the target audience of HITRAN are scientists that study terrestrial and planetary atmospheres, the applications of HITRAN span a great many fields of science, engineering and medicine. The database

[11] https://cdms.astro.uni-koeln.de,https://spec.jpl.nasa.gov

was first established several decades ago [75] and has been under continuous development since then [16,76–83]. The latest edition of the database is HITRAN2016 [16] and is distributed through HITRAN*online* [84] that is accessible through the VAMDC portal. The VAMDC portal provides access to a traditional line-by-line high-resolution section of HITRAN, which contains spectral parameters for 49 molecules along with their significant isotopologues appropriate for terrestrial and planetary atmospheric applications. HITRAN provides references to the original sources for the majority of the parameters for every transition. The details of how such referencing was incorporated is described by Skinner et al. [85] in this special issue. It is worth pointing out that with an exception of some diatomic molecules (including hydrogen halides) for which the line lists can be used at temperatures up to 5000 K or higher, the majority of the HITRAN line lists are intended to be used at the lower temperatures encountered in the terrestrial atmosphere.

For higher temperature applications, it is recommended to use the HITEMP database [86], which mimics the format of HITRAN but contains a substantially larger amount of lines. However, at the present time, the HITEMP database provides data for only eight molecules [86–89], and is not currently accessible through VAMDC due to the large number of transitions it contains.

HITRAN also provides a section that contains high-resolution experimental cross sections for molecules with very dense spectra that are not amenable to a full quantum-mechanical description. A recent update [90] provided spectra of almost 300 molecules at various pressures and temperatures. There is also a section of the compilation that provides collision-induced absorption cross sections, which has been updated recently [91]. These data are not accessible through VAMDC yet but can be accessed through HITRAN*online*[12] or through the HITRAN Application Programming Interface (HAPI) [92]. A very extensive effort is currently underway to release the new HITRAN2020 edition of the database.

2.1.10. S&MPO

S&MPO (Spectroscopy and Molecular Properties of Ozone) is an information system [17] jointly developed by Reims University (France), Institute of Atmospheric Optics (Russia) and Tomsk State University (Russia). The line-by-line list of vibration-rotation transitions, energy levels, transition moments and other related information can be requested from S&MPO via the VAMDC portal. Since the previous edition of VAMDC [3], considerable new spectroscopic information has been included using experimental data and extended analysis of ozone bands in the infrared [93] and isotopic spectra enriched by ^{17}O and ^{18}O oxygen [94,95]. Supplementary information concerning theoretical spectra simulation, comparisons with experimental records, dipole moment functions [96] and *ab initio* intensities for strong lines [97] can be accessed via the Reims[13] and Tomsk[14] sites. The S&MPO interactive software was completely rewritten to make it relevant to current trends in internet application development. A new version 3.0 of the S&MPO relational database and information system is now operational and provides supplementary functionalities for the fully updated graphical interface. More information about the version 3.0 S&MPO information system can be found on its public website[15]. Applications of the S&MPO may include education/training in molecular and atmospheric physics, studies of radiative processes and spectroscopic analysis.

2.1.11. MeCaSDa, ECaSDa, TFMeCaSDa, SHeCaSDa, GeCaSDa, RuCaSDa, TFSiCaSDa, UHeCaSDa

These databases contain calculated rovibrational transitions (mostly infrared absorption and also some Raman scattering lines) for highly symmetrical molecules. They result from the analysis

[12] https://hitran.org
[13] http://smpo.univ-reims.fr
[14] http://smpo.iao.ru
[15] https://smpo.univ-reims.fr/news/en_2019-09-25_02-22-19

and fit of effective Hamiltonian and transition moments using experimental spectra recorded at high resolution. The calculation uses a model [98] and programs [99] developed by the group at the Laboratoire Interdisciplinaire Carnot de Bourgogne, Dijon, France, and based on group theory and tensorial formalism [100].

The MeCaSDa database [101] contains methane (CH_4) lines, which constitutes the main research subject of the Dijon group. It is Dijon's biggest database and is of importance for atmospheric and planetary applications. Its contents have recently been updated, and it now has more than 16 million lines [18].

Other databases for several molecules of either atmospheric, planetological or industrial applications have been significantly updated recently [18]: ECaSDa (C_2H_4 found in Earth's and planetary atmospheres [101]), TFMeCaSDa (CF_4 found in Earth's atmosphere), SHeCaSDa (SF_6 found in Earth's atmosphere), GeCaSDa (GeD_4 found in giant planets' atmospheres) and RuCaSDa (RuO_4 important for the nuclear industry).

Two new databases have been added recently: TFSiCaSDa [19] concerning the SiF_4 molecule with applications to volcanic gases and UHeCaSDa (UF_6) for applications for the nuclear industry. This latter database is an exception: no experimental data are publicly available for this peculiar radioactive molecule and calculations are only based on literature data [102,103]. More databases of such highly symmetrical molecules maybe developed in the near future. The databases SHeCaSDa, MeCaSDa, GeCaSDa, TFMeCaSDa and RuCaSDa are linked with the Query Store service (Section 3.2.2), while the databases ECaSDa, TFSiCaSDa and UHeCaSDa are currently being connected to the Query Store.

2.1.12. CDSD-296, CDSD-1000, CDSD-4000, ASD-1000, NOSD-1000 and NDSD-1000

Six molecular databanks in VAMDC of atmospheric and astrophysical interest have been provided by the Laboratory of Theoretical Spectroscopy, V.E. Zuev Institute of Atmospheric Optics, Siberian Branch, Russian Academy of Sciences, Russia. These include three versions of the Carbon Dioxide Spectroscopic Databank (CDSD-296, CDSD-1000, CDSD-4000), the Acetylene Spectroscopic Databank (ASD-1000), the Nitrous Oxide Spectroscopic Databank (NOSD-1000) and the Nitrogen Dioxide Spectroscopic Databank (NDSD-1000). These databanks provide positions, intensities, air- and self-broadened half-widths, coefficients of temperature dependence of the half-widths and quantum numbers associated with the transitions. The line positions and intensities are calculated within the framework of the method of effective operators. The line shape parameters are calculated using different theoretical approaches or empirical equations in terms of the rotational quantum numbers.

CDSD-296 and CDSD-1000 have been described in our previous paper [3]. CDSD-296 has been updated recently [20] to improve the line parameters accuracy. CDSD-4000 contains more than 628 million lines of the four most abundant isotopologues of carbon dioxide and covers the 226–8310 cm^{-1} wavenumber range [21]. The reference temperature is $T_{ref} = 296$ K and the intensity cutoff is 10^{-27} $cm^{-1}/(molecule\,cm^{-2})$ at 4000 K. ASD-1000 has more than 30 million lines of the principal isotopologue of acetylene in the 3–10000 cm^{-1} spectral range [24]. The database is adopted for the temperatures up to 1000 K with an intensity cutoff 10^{-27} $cm^{-1}/(molecule\,cm^{-2})$. The line intensities and pressure broadening coefficients are given for $T_{ref} = 296$ K. NOSD-1000 contains more than 1.4 million lines of the principal isotopologue of nitrous oxide and covers the 260–8310 cm^{-1} region [22]. The intensity cutoff is at 10^{-25} $cm^{-1}/(molecule\,cm^{-2})$ at $T_{ref} = 1000$ K. Finally, the NDSD-1000 has more than 1 million lines in the 466–4776 cm^{-1} wavenumber range [23]. The intensity cutoff is at 10^{-25} $cm^{-1}/(molecule\,cm^{-2})$ at $T_{ref} = 1000$ K. The broadening parameters are given for two reference temperatures: $T_{ref} = 296$ K and $T_{ref} = 1000$ K.

Currently, the above mentioned databanks are presented in VAMDC as separate nodes with a similar basic set of "restrictables" [16] parameters (wavenumber, wavelength, line strength, InChIKey

[16] Restrictables refer to the type of data queried by VAMDC (https://standards.vamdc.eu/dictionary/restrictables.html).

and state energies). The common set of returnable entities includes radiative transitions, sources, isotopic species, environment signatures and function descriptions for the temperature and pressure dependencies of line shape parameters. The parametric content of these returnables however, depends on the original structure of the particular databank. Unification of these databanks, which is a step necessary for their merging into a single VAMDC node, is subject of future work.

Besides being available through VAMDC, the databanks can be downloaded from the IAO ftp server[17] where they are presented in the original tabular format.

2.1.13. SESAM

SESAM[18] is restricted to the electronic spectra of molecular hydrogen and its deuterated substitutes, as well as CO. The hydrogen spectra includes the Lyman and Werner band systems as well as B'-X, D-X electronic bands. The SESAM database allows queries, within a specific wavelength range, of the properties of available transitions of a selected molecule. The transitions are in the Vacuum Ultra Violet (VUV) range. It is also possible to download the full range of data for a particular goal. Different additions are considered, e.g. the query of these molecular transitions at any redshift, which can be interesting for extragalactic observations where the spectrum is shifted in the visible. Since 2016 the CO molecule has been added. Further information and the latest updates are described in this Special Issue [13]. It should be noted that the SESAM database is linked to the VAMDC Query Store (Section 3.2.2).

2.1.14. W@DIS

The W@DIS databases are part of the information system [25] designed to provide access to both tabular and graphical data, as well as information [104][19] and ontologies [105] for quantitative molecular spectroscopy necessary for solving fundamental and applied problems in a number of subject areas: atmospheric spectroscopy of planets and exoplanets, astronomy, etc. The semantic information system W@DIS is the next-generation molecular spectroscopy information system, based on application of Semantic Web technologies to its tabular and graphical information resources [106,107]. The W@DIS information system is available and hosted at the V.E. Zuev Institute of Atmospheric Optics in Tomsk, Russia[20]. This system can be a prototype for semantic information systems in the atomic, ionic and solid-state spectroscopy to be used by the VAMDC consortium. Meanwhile some data such as transitions are accessible from the VAMDC portal; they can be displayed via the "Molecular XSAMS to HTML" visualisation tool.

2.1.15. KIDA

The KInetic Database for Astrochemistry (KIDA[21]) [26,27] is a compilation of kinetic data (chemical reactions and associated rate coefficients) used to model chemistry in astrophysical environments (interstellar medium, protoplanetary disks, planetary atmospheres, etc.). In addition to detailed information on each reaction (e.g. temperature range of validity of the rate coefficients, reference and uncertainty), particular attention is given to the quality of the data, which is evaluated by a group of experts in the field. Since 2016, KIDA also compiles data used to compute chemical reactions occurring on the surface of interstellar dust grains (branching ratios, activation energies and barrier width) and desorption/diffusion of species on these surfaces (desorption and diffusion energies).

---

17 ftp://ftp.iao.ru/pub/
18 http://sesam.obspm.fr/
19 "information" is interpreted in accordance with the same term defined in the given reference
20 http://wadis.saga.iao.ru
21 http://kida.astrophy.u-bordeaux.fr

366 2.1.16. UDfA

367 The UMIST Database for Astrochemistry (UDfA), first released to the public in 1991 [108], contains
368 basic chemical kinetic data and associated software codes and documentation for modelling the
369 chemical evolution of interstellar clouds and the circumstellar envelopes of evolved AGB stars. The
370 core of the data, which can be accessed via its website[22] in addition to VAMDC, consists of reaction
371 rate coefficients of several thousand gas-phase reactions and is supplemented by more restricted
372 data sets concerning the chemistry of deuterium fractionation. The database does not contain any
373 surface chemistry but a file of surface binding energies, which allows processes such as reaction and
374 desorption to be taken into account, is supplied. Where possible, and in line with VAMDC policy,
375 a great deal of effort has been made to identify the precise source of each datum entry through the
376 application of its DOI, in particular for those rate coefficients measured experimentally. Software codes
377 for calculating the chemical evolution of interstellar and circumstellar regions are also provided as
378 are codes that generate UDfA output files in the form needed for radiative transfer codes such as
379 RADEX [23], RATRAN [24] and RADMC-3D [25] used to calculate emergent molecular line profiles from
380 these regions. A major revision of the 2013 release [28], including the review of current data and the
381 identification of new reactions, particularly those associated with chlorine chemistry and with the
382 formation of metal oxides, hydroxides and chlorides in oxygen-rich AGB stars, is underway with a
383 public release due by the end of 2020.

384 2.1.17. BASECOL

385 The BASECOL database collects, from the refereed literature, the rate coefficients for the excitation
386 of rotational, vibrational and rovibrational levels of molecules by atoms, molecules and electrons. The
387 processes are described in the temperature ranges relevant to the interstellar medium, to circumstellar
388 atmospheres and to cometary atmospheres. The BASECOL database is currently the sole VAMDC
389 connected database that implements the Java version of the node software[26]. It can be displayed
390 from the VAMDC portal with the "Collisional data XSAMS to HTML" processor and is accessible
391 from the SPECTCOL tool [109,110]. Since its last review paper in 2013 [29] the scientific content of
392 the database is updated with published data. Since the last VAMDC review paper in 2016 [3], the
393 technical components of the database have been entirely replaced [30] with the internal structure of
394 the database, and the data ingestion files have been made compliant with the metadata necessary for
395 VAMDC interoperability. The public graphical interface has been changed to a simpler system. The
396 connection of BASECOL with the Query Store (Section 3.2.2) is in a final testing phase. A complete
397 description of the new BASECOL technical design and updates can be found in this Special Issue [30].

398 2.1.18. MOlD

399 The MolD VAMDC node [31,32] provides data for plasma modelling, e.g. for modelling different
400 stellar atmospheres, early Universe chemistry and analysis of the kinetics of laboratory plasma. MolD
401 contains photodissociation cross sections for individual rovibrational states of diatomic molecular ions
402 as well as corresponding data on molecular species and molecular state characterisations calculated
403 using a quantum mechanical method described in [111]. Since the previous VAMDC review [3] large
404 amounts of new data concerning alkali molecular ions have been included from reference [112].
405 The node is hosted at the Belgrade Astronomical Observatory[27]. It has enabled fairly easy access
406 to data (in tabulated and graphical form) of thermally averaged photodissociation cross sections

22 http://www.udfa.net
23 https://personal.sron.nl/~vdtak/radex/index.shtml
24 https://personal.sron.nl/~vdtak/ratran/frames.html
25 www.ita.uni-heidelberg.de/~dullemond/software/radmc-3d/
26 www.vamdc.org/activities/research/software/java-nodesoftware/
27 http://servo.aob.rs/mold

across the available spectrum at a requested temperature in order to facilitate atmospheric modelling and other numerical calculations [113]. Future plans are to include new (i.e. complex) molecules of astrophysical importance. The MolD database is linked to the Query Store service (Section 3.2.2).

2.1.19. BeamDB

BeamDB (the Belgrade electron-atom/molecule DataBase) is a collisional database, in which electrons are projectiles while targets are considered to be atoms and molecules [33]. Interactions of electrons with atoms and molecules are presented as both differential and integral cross sections for processes such as elastic scattering, excitation and ionisation [114]. Since the previous review of VAMDC [3] a considerable volume of new collisional data on metal atom targets have been included from the published sources (e.g. for bismuth [115] or zinc [116]) and selected molecules (e.g. for methane [113] or nitrous oxide [32]). Curation and maintenance of electron collisional data is relevant in many research areas such as astrophysics [117], plasma [118], radiation damage [119] or in lighting applications [120]. The plan for the future expansion of the database is to include ions as new targets [121]. BeamDB[28] is hosted at the Belgrade Astronomical Observatory and is linked to the Query Store service (Section 3.2.2).

2.1.20. IDEADB

The Innsbruck Dissociative Electron Attachment (DEA) database node collects relative partial cross sections for dissociative electron attachment processes of the form: $AB + e^- \rightarrow A^- + B$, where AB is a molecule. Queried identifiers are searched in both products and reactants of the processes. XSAMS files (see Section 3.1.1) are then returned which describe the processes found including numeric values for the relative partial cross sections of the dissociative electron attachment reactions. Additionally, a visual representation of the cross sections can be viewed on the web site. Since 2016 the possibility to add cross sections for cationic products and resolution of several minor issues have been addressed. It is planned to modify the database structure, so that measurements in matrices such as water or helium nanodroplets can be added. The node[29] is hosted and maintained by the group of Paul Scheier at the University of Innsbruck, Austria.

2.2. VAMDC Data Nodes that Have not Evolved since 2016

There have been no changes to some databases and nodes since 2016. These will be described in this section, and more information can be found in our previous publication [3]. It should be noted that these databases contain quite unusual species and processes for the VAMDC e-infrastructure, for which the current VAMDC visualisation tools and even the current VAMDC standards lack features that would allow a full display of the databases contents. Therefore at present VAMDC mainly provides only a simple view of those databases.

2.2.1. LASp

The LASp (Laboratorio di Astrofisica Sperimentale) database is hosted at the Catania Astrophysical Observatory in Italy[30] and additional information can be found at its URL[31]. LASp spectra are taken by using *in situ* techniques and equipment especially developed to analyse the effects of irradiation (ion and/or UV photons) and thermal cycling (down to 10 K) by infrared, Raman and UV-VIS-NIR spectroscopy. Analysed materials include frozen gases, solids samples and meteorites.

---

[28] http://servo.aob.rs/emol
[29] https://ideadb.uibk.ac.at/
[30] http://vamdclasp.oact.inaf.it/GUI/index
[31] http://www.oact.inaf.it/weboac/labsp/index.html

The main application field up until now has been in astrophysics, and over the years many hundreds of ice mixtures of various compositions and of solids have been studied.

Through the VAMDC portal, transmission and optical depth data for water-ice experiments are available [122,123]. The public database also includes optical constants of solid-state CO, CO_2 and CH_4 deposited at various temperatures.

2.2.2. PAH

The Cagliari–Toulouse PAH theoretical spectral database [36] is a joint effort by the groups of G. Mulas (INAF-OAC) and C. Joblin (Univ. Toulouse III/CNRS-IRAP) aimed at providing all the "ingredients" needed for modelling the photophysics of individual polycyclic aromatic hydrocarbons (PAHs) in space, mainly in photon-dominated interstellar and circumstellar environments. It includes basic structural properties of PAHs in four charge states (-1, 0, +1 and +2), ionisation potentials and electron affinities, harmonic vibrational analyses and vertical photoabsorption electronic spectra. The link to the "old" database with flat files, which includes more data than available through VAMDC, is still online[32]. The web interface to the relational database, holding the data available through VAMDC, is hosted at the Cagliari Observatory in Italy[33]. An effort is underway to develop import tools that will feed the relational database, and we expect some substantial changes within a year's time. Currently, from the VAMDC portal it is possible to obtain molecular structures, corresponding energies and vibrational analyses for a number of PAHs in different ionisation states. The electronic photoabsorption spectra are not available yet through VAMDC but will be in the near future.

2.3. Databases in the Process of Being Connected to VAMDC

2.3.1. ExoMolOP

The ExoMol project provides molecular line lists for exoplanet and other atmospheres [124] with a particular emphasis on studies of hot atmospheres. Besides line lists, which are stored as states and transitions files [125], the ExoMol database[34] stores a variety of other associated data including partition functions, state lifetimes, cooling functions, Landé g-factors, temperature-dependent cross sections, opacities, pressure broadening parameters, k-coefficients and transition dipoles. These data and the associated data structure are described in the database release papers [126,127]. The line lists provided by ExoMol are huge and are too big to be handled by the VAMDC portal; this issue is described below.

Recently a new offshoot of the ExoMol project called ExoMolOP has been created, which contains opacity cross sections and k-tables [128] for molecules of astrophysical interest [37]. This database is built on ExoMol data but contains input from HITRAN [16], the empirical MolList database of Bernath [129] and NIST for selected atoms. ExoMolOP provides data on a grid of 22 pressures and 27 temperatures on a grid of wavelengths for each species. By comparison with the unprocessed ExoMol data, these provide a comparatively compact representation of the absorption property of each species. An implementation of the ExoMolOP data within the VAMDC portal is currently in progress.

2.3.2. SSHADE in VAMDC

Currently VAMDC allows access to the original GhoSST database[35](Grenoble astropHysics and planetOlogy Solid Spectroscopy and Thermodynamics) to search for a few pure and mixed molecular solids through their constituent species and to retrieve their infrared spectra, either as

[32] https://astrochemistry.oa-cagliari.inaf.it/database/
[33] https://qchitool-pah-dev.oa-cagliari.inaf.it/
[34] http://www.exomol.com
[35] https://ghosst.osug.fr/

absorption coefficients or as optical constants. SSHADE[36] [35] is a database infrastructure of solid-state spectroscopy that hosts spectral data of many different types of solids including ices, snows, minerals, carbonaceous matters, meteorites, interplanetary dust particles and other cosmo-materials covering a wide range of wavelengths: from X-rays to millimeter wavelengths. The data are collected from a consortium of partners[37], which provide their data in their own database of the SSHADE infrastructure. Currently a "band list" database of molecular solids is under development in the frame of the Europlanet-2024 research infrastructure program. It will host critical compilations of the position, intensity, width and vibration modes of absorption bands (visible-infrared or Raman active) of pure and mixed molecular solids as well as for several types of molecular compounds such as hydrates and clathrate hydrates. Both the SSHADE spectral databases and the band list database will be linked to VAMDC, but this will require an upgrade of the XSAMS data model (see Section 3.1.1) in order to describe the fundamental solid constituents better. With such databases, VAMDC will allow the user to retrieve and compare the band parameters of molecular species in both gaseous and solid states, and therefore, will allow them to determine which one does actually contribute to observed absorption features. This capability is particularly useful in environments where both phases coexist, such as planetary atmospheres with aerosols.

2.3.3. AMBDAS

AMBDAS (Atomic and Molecular Bibliographic Data System)[38] is a library of around 50,000 references to publications in the scientific literature concerning collisional and plasma-material interactions of relevance to nuclear fusion energy research. An online, browser-based searchable interface allows the database to be queried by reactant species, charge state and process type as well as by author, journal and title keyword.

Integration of AMBDAS within the VAMDC infrastructure is planned for release in 2020 as part of an upgrade to the database software, which includes a recently developed DOI-centred reference management library [85] and an updated classification of plasma processes [130]. In addition to a new search interface, VSS2 queries and XSAMS output (see Section 3.1.1) are supported. The AMBDAS system will be queried by species and processes, and so will be accessible from the species database. In addition, since the AMBDAS system is a bibliographic database, it will be queried through the new bibliographic service described in Section 6.5.

2.3.4. DREAM-DESIRE

The Database of Rare Earths At Mons University (DREAM) contains information concerning the radiative parameters (wavelengths, transition probabilities and oscillator strengths) for more than 72,000 spectral lines belonging to the lower ionisation stages of lanthanide elements ($Z = 57$ to 71), from neutral to triply ionised species. This database, originally created by Biémont et al. [131] and recently updated by Quinet and Palmeri [39], is hosted by Mons University[39]. All the data tabulated in DREAM have been determined from detailed pseudo-relativistic Hartree–Fock calculations including core-polarisation effects [132,133] carried out by the Atomic Physics and Astrophysics group at Mons University, Belgium. The accuracy of the theoretical results have been assessed through comparisons with experimentally measured radiative lifetimes using laser-induced fluorescence spectroscopy. The Database on Sixth Row Elements (DESIRE) contains the same type of information as DREAM, but is dedicated to the elements of the sixth row elements of the periodic table ($Z = 72$ to 86). This database[40] is described in [38].

---

[36] http://www.sshade.eu
[37] https://wiki.sshade.eu/sshade/databases
[38] https://amdis.iaea.org/databases/
[39] http://hosting.umons.ac.be/html/agif/databases/dream.html
[40] http://hosting.umons.ac.be/html/agif/databases/desire.html

2.3.5. IAMDB

The volume of high-precision data generated by the Indian atomic and molecular community is quite substantial. This is evident from the number of articles published by various groups. However, such data are not well organised and, hence, very difficult to retrieve for further use. This is the reason for the requirement of an atomic and molecular data repository in India, which has been envisaged for many years now. The Indian Atomic and Molecular DataBase (IAMDB[41]) is the outcome of such an effort. With the help of the VAMDC interface, the data generated and gathered in IAMDB can be easily retrieved by any user.

In the last few years a great deal of electron scattering data has been produced in India. In particular, partial and total electron ionisation cross sections have been measured over an extended energy regime for about two dozen organic molecules important to radiation biology [134,135]. Since many of these systems are in solid form at room temperature, a new experimental technique was employed for the measurements. The result of some important ones like DNA and RNA bases are reported in [134,135]. The same group has also measured dissociative electron attachment cross sections for several organic molecules [136]. On the other hand, Antony and coworkers have produced a large quantity of electron scattering data, in particular for those molecules and radicals which are difficult to measure [137,138]. Recent calculations from this group include positron scattering from a variety of atoms and molecules [139,140] and photoionisation cross sections for polyatomic molecules [141,142]. Once these electron/positron/photon collision data are incorporated into IAMDB, it will be integrated with VAMDC.

2.3.6. PEARL

The PEARL (Photonic Electronic Atomic Reaction Laboratory) database in the Korea Atomic Energy Research Institute (KAERI) includes electron impact ionisation (EII), recombination and photoionisation for atoms and ions. The EII [143] and the dielectronic recombination (DR) [144] have been calculated using a relativistic distorted wave approximation, and the photoionisation [145] has been calculated using a non-iterative eigenchannel R-matrix method for the ground and lower excited levels of atomic ions below $Z = 30$, which are of interest in astrophysics. The EII [146] and the DR [147] calculations have also been performed for tungsten (W, $Z = 74$) ions, which are essential in fusion tokamak research. The calculated cross section data can be graphically displayed on the PEARL web site[42], together with other available experimental and theoretical data for comparison, the numerical data can be also downloaded. Recently a collisional–radiative model (CRM) for low-temperature plasma has been developed and the electron impact excitation (EIE) data for the levels considered in the model have been calculated and compiled. The calculated line ratios of He I can be displayed as a function of the electron temperature and density on the website. CRM results for Ar I [148] will be uploaded in the near future.

2.3.7. Clusters

Clusters are complexes of two up to several million atoms and/or molecules which bridge the gap between molecular physics and solid-state physics. The addition or removal of a single atom or molecule from a cluster may dramatically change its properties. Interesting attributes in cluster physics are e.g. cluster structures, bond lengths and bond dissociation energies. These properties can be extracted by combining experimental techniques such as mass spectrometry and/or spectroscopy in combination with quantum chemical simulations.

---

41 www.iamdb.org.in
42 http://pearl.kaeri.re.kr

Clusters are not yet addressed by VAMDC. Inclusion of clusters in VAMDC requires the addition of a section to the schema of VAMDC to describe the cluster data. We propose to develop a two-layer implementation of the node. The first layer – which will be available in the portal as well – is for selecting the dominant species of interest (for example "$(CO_2)(H_2O)$"). This will return all datasets with clusters of "$(CO_2)_n(H_2O)_m$", $(m, n > 0)$. Often, the produced mass spectra display features of many different cluster ions, because impurities are attached to the clusters of interest. Fragmentation of larger molecules may also display features of further non-stoichiometric cluster progressions. Thus, different queries return the same dataset. On the website of the node itself we will offer more filter options:

- Method of cluster formation (supersonic expansion, seeded beam, gas aggregation, electrospray ionization, helium nanodroplets, ...)
- Method of ionization (electrospray ionization, matrix-assisted laser desorption/ionization, electron impact, photo ionization, ...)
- Steps in between (tandem mass spectrometry, collision-induced dissociation, ...)
- Analysis method (time-of-flight, quadrupole, ion cyclotron resonance, ...)
- Environment (temperature, pressure, ...)
- Others (evaluation of data, publication, magic numbers, solvation-effect, ...)

The most important returned data will be mass spectra (which can also be visualized in a browser). Published papers related to the species asked for as well as possible evaluation programs [149] and their output will be made available. Larger files will be available by a download link found in the XSAMS data file (see par. 3.1.1). Since 1986 a group in Innsbruck [150,151] has been producing mass spectra of clusters via different approaches. These results are planned to be made available in this database.

3. Current Status of VAMDC e-Infrastructure

3.1. Overview of the VAMDC e-Infrastructure Components

The e-infrastructure currently connects, in an interoperable way, 38 heterogeneous databases with atomic and molecular data. By providing data producers and compilers with a large dissemination platform for their work, VAMDC successfully removes the bottleneck between data producers and the wide community of atomic and molecular data users. The "V" of VAMDC stands for "virtual" in the sense that the e-infrastructure itself does not contain the data: it is a wrapper that exposes the heterogeneous databases in a unified way. The wrapping software, called the node-software [152], integrates a stand-alone database into a VAMDC federated database to become a data-node. Each data-node accepts queries submitted in a standard grammar (see Section 3.1.1) and provides output in a standard format. Each data-node is entered into the VAMDC registry (see Section 3.1.2) that enables a standardised application programming interface (API) to discover the available resources.

3.1.1. Data Nodes, Query Language and Data Formats

A data-node is a database, either pre-existing or created for the purpose of VAMDC, wrapped in the node-software that implements the web-service[43] protocol VAMDC-TAP, which is derived from the International Virtual Observatory Alliance (IVOA)[44] Table Access Protocol (TAP).[45]

All VAMDC-TAP services support a common data model, query language and output format. The data model, expressed in the VAMDC Dictionary, represents the data both in a tree structure and

[43] In this context, a *web service* is a data source on the WWW designed for access by application software. C.f. a *web page* designed to be interpreted by a human intellect.
[44] http://ivoa.net/
[45] http://www.ivoa.net/documents/TAP/20190927/ TAP enables an application to query a remote database.

as a standardised set of virtual tables. The query language VAMDC SQL Subset 2 (VSS2)[46] operates on these virtual tables. The common output format is the XML Schema for Atoms, Molecules and Solids (XSAMS) [47] [153] that uses the tree-structured form of the data model. All VAMDC-TAP services allow XSAMS output while the nodes may optionally support other formats. XSAMS is highly flexible and expressive, but other formats may be preferred for compactness.

Libraries in Python of the node software are provided by the VAMDC consortium. The node owner configures these libraries for the database of choice by adding small translation-functions from VSS2 to SQL operating on the actual database, and from the query results to XSAMS or other formats. Additional information and examples of best practices were described by Regandell et al. [152].

3.1.2. Registry

The VAMDC metadata registry[48] lists the details of the VAMDC data-nodes. Applications use the registry to decide which databases should be queried and to locate the services for those databases on the internet. The VAMDC registry is based on the work of Astrogrid, which was the UK's Virtual Observatory development project from 2001 to 2010 [154]. They developed a registry whose interface is based on the then-current IVOA standard. To simplify the access to this registry, the VAMDC consortium provides Java and Python libraries that are used by the VAMDC portal, among other applications.

3.1.3. The Portal

The VAMDC portal[49] [155] relies on the infrastructure elements previously described in this section to provide seamless access to the inter-connected VAMDC databases. Through this unique interface a user can query any database member of the VAMDC infrastructure and can retrieve data in the common shared file format VAMDC-XSAMS (see Section 3.1.1). The page displaying the resulting data recalls the exact query processed by the infrastructure to produce the data (for example, see Fig. 1).

---

[46] http://vamdc.eu/documents/standards/queryLanguage/vss2.html
[47] https://standards.vamdc.eu/#data-access-protocol-query-language-and-dictionaries
[48] http://registry.vamdc.eu/registry-12.07/main/index.jsp
[49] https://portal.vamdc.eu/vamdc_portal/home.seam

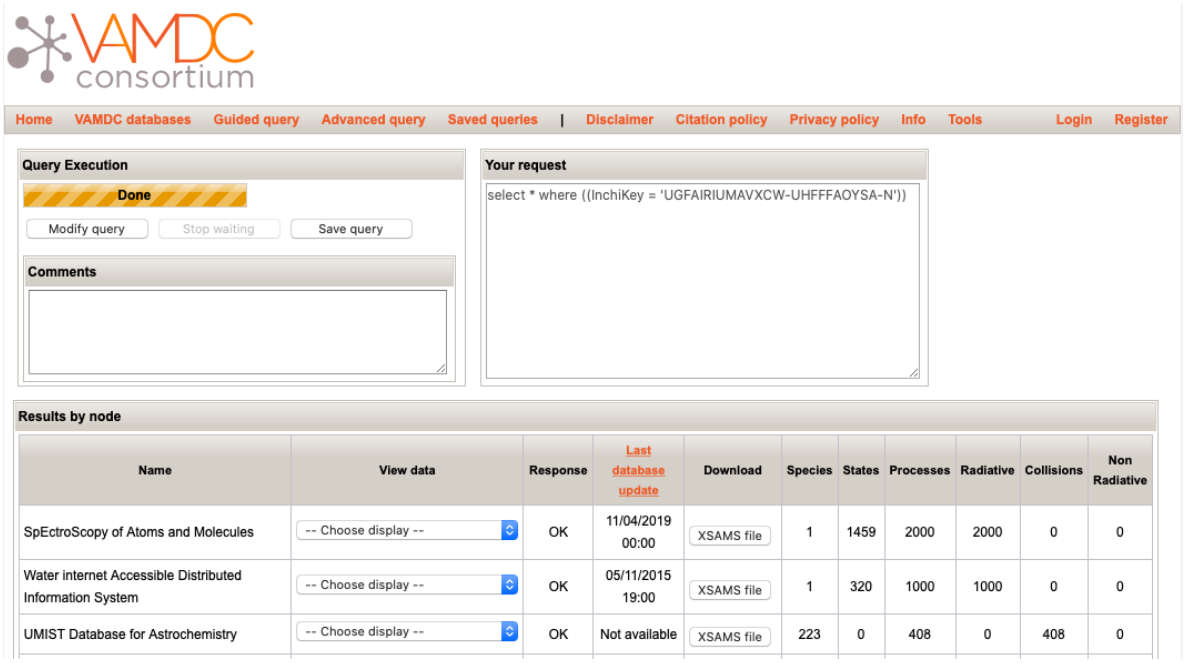

Figure 1. Example of the displayed results by the VAMDC portal after performing a given query. The processed query is highlighted in the *Your Request* box

The portal embeds processors to convert data from the XSAMS format into several formats (chosen by user) and has several graphical tools to visualise the extracted data. Moreover, the portal implements the IVOA-SAMP protocol [156] created to connect scientific tools when working with multiple data types. VAMDC data can be directly piped from the portal into any tool implementing the SAMP protocol, e.g. TOPCAT[50].

3.2. Services Built over Existing VAMDC Infrastructure since 2016

3.2.1. The Species Database

To overcome the problem of species identification, the VAMDC consortium created in 2016 a centralised chemical-species repository called the species database. Updated daily by automated collation of data from the VAMDC data-nodes, it contains the list of all the species in each VAMDC database. Every species is identified uniquely by an InChIKey [157], a hash code generated from an InChI description[51]. In the species database each InChiKey is associated with the different ways of identifying a species; e.g. their chemical names, formula, stoichiometic formula and CAS number. By adding a REST API and a web graphical interface to this species database[52], we provide a versatile tool to explore the species content of the atomic and molecular VAMDC connected databases. Using this REST API[53] of the species database, the VAMDC portal provides both an autocomplete suggestion for species names and identifiers and a feature to discover the isotopologues of a species. Thus, it is possible to specify, very precisely, which species is the most relevant to a user's search.

---

50 http://www.star.bris.ac.uk/~mbt/topcat/

51 https://iupac.org/who-we-are/divisions/division-details/inchi/, https://iupac.org/who-we-are/divisions/division-details/inchi/

52 https://species.vamdc.eu

53 A RESTful API is a interface between a web-based client and a server that exploits representational state transfer (REST) constraints.

3.2.2. The Query Store

By using the VAMDC facilities, scientists can easily discover the atomic and molecular resources and access their data in a unique and practical way. However, as VAMDC has been adopted by a larger range of communities it has revealed a new set of challenging issues linked with citation and data reuse:

- The VAMDC infrastructure data are dynamic. A database displayed through the VAMDC infrastructure may evolve over time: the most recent and precise version of given data may replace old ones. We therefore needed mechanisms to allow the citation of dynamic data.
- The data set provided by the VAMDC infrastructure always contains the references of the papers used for compiling the data sets. However, the citation process may become cumbersome when the extracted data sets come from many sources.

VAMDC is addressing these issues at the data-community level and in 2014 VAMDC joined the Research Data Alliance. The RDA, through its Data Citation Working Group and RDA/WDS Scholarly Link Exchange (Scholix) Working Group, has defined new citation models in the digital era. We succeeded in implementing the RDA recommendations to provide the VAMDC users with a Query Store [158], a tool to facilitate citation of the data extracted from VAMDC for scientific reproducibility and for giving due credits to data producers. Through the Query Store:

- each query served by the infrastructure is identified by a persistently unique resolvable identifier;
- the query-produced data may be assigned a DOI[54];
- data become directly citable by their DOI.

When registering a DOI, the authors of the papers used for compiling the data appear in the "references" part of the DOI-metadata schema. The data-set sources become the references of the DOI. The authors/papers referenced in the VAMDC extracted data set will automatically get credit when the data set is cited (using the DOI) into a paper.

Table 2 lists the VAMDC databases currently integrated with the Query Store service, whose queries may be discovered by accessing the URL[55]; this service is the most complete collection of VAMDC queries that may be cited as example of the queries performed through the infrastructure.

Table 2. List of databases connected with the Query Store service. The databases marked with a star ($\star$) are currently being connected to the Query Store and the test phase is in progress.

VALD	Topbase	Tipbase
SHeCaSDa	MeCaSDa	GeCaSDa
TFMeCaSDa	RuCaSDa	TFSiCaSDa*
UHeCaSDa*	ECaSDa*	Stark-B
CDMS	JPL	BeamDB
Mold	SeSAM	Basecol*

3.3. Pending Technical Issues

In addition to all of the advancements of the individual databases and VAMDC as a whole, there are still technical issues that need to be addressed in the future.

3.3.1. Treatment of Big Data

The growth in large data sets, which is common to all areas of science, is also true in A&M physics. VAMDC uses a tightly defined and rigorous data structure that is also, as a result, relatively

[54] DOI: Digital Object Identifier, a formal name for a document or data-set in a standard format intelligible by software.
[55] https://cite.vamdc.eu

verbose. This means that the data returned by queries to some databases such as VALD and KIDA are already limited by the software so as not to cause problems. Moreover, if many current VAMDC molecular spectroscopic databases contain data with a relatively limited amount of spectral lines for each molecule (HITRAN, JPL, CDMS, ...), some others (MeCaSDa,...) contain extensive calculated line lists, and in this case, their number for a given species can be very large (several millions of spectral lines at least) and could increase even more with subsequent updates. At present there is a limit within the VAMDC data retrieval volume. Queries producing large data extracts time out or are truncated by the data-nodes to avoid timeouts. This prevents downloads of line lists over a wide spectral range or for many molecules at once. The actual limits vary from node to node according to the computing resources invested and the details of the database.

For example, when asked for all data connected with lines between 150 nm and 800 nm, seven nodes produced a complete result, three produced a truncated result, two failed and one timed out; the others quickly reported that they had no relevant data. The largest response was for 67 million lines in 218MB of XSAMS, and this is close to the practical limit of the system. For this data-set, three of the recommended displays and format-converters were able eventually to process the data, one failed and one produced no results within 10 minutes.

This is a problem for applications in atmospheric planetary physics and exoplanet studies. A solution currently under investigation is the possibility of asynchronous queries that run slowly in the background and leave their results cached and accessible by an ephemeral URL.

There are also whole databases that are currently really too big to be usefully probed via the VAMDC portal. A particular area that produces very large data sets is the provision of molecular line lists for studies of hot atmospheres. The ExoMol [126,127], TheoReTS [159] and the NASA Ames group (e.g. [160]) all employ theoretical methods to compute very extensive lists of molecular transitions particularly for high-temperature applications in astrophysics. Many of these line lists are huge; for example, the ExoMol line list CH_3Cl contains 166 billion transitions for each of the two main isotopologues [161]. Even at room temperature, complete line lists for relatively heavy long-lived greenhouse molecules with low-frequency vibrations like CF_4 [162] or NF_3 [163] require billions of transitions to converge opacity calculations over infrared wavelengths. Data sets of this size are completely outside the current capabilities of the VAMDC project.

Some work has been performed on data compacting, in particular using the so-called super-lines approach [164,165]. Another method of big-data compression is to use "effective" lines as implemented in a recent addition of methane [89] to the HITEMP database [86]. Use of these protocols can lead to a very significant reduction in the size of the effective line list, but these compressed lists still contain many millions of lines making it difficult for the current VAMDC infrastructure to process or download the full line list even after it is compacted. We note that the ExoMol line lists have recently been used to generate a set of temperature and pressure dependent opacities [37] creating the so-called ExoMolOP database. These opacity functions provide a much more compact representation of the molecular data, and an implementation of the ExoMolOP database within VAMDC is currently being explored. A similar direction is currently being followed for the TheoReTS information system[56] containing a theoretical and experimental atlas of methane absorption/emission cross sections [166] up to 1000 K.

3.3.2. Selection and Comparison of Atomic States

It is a long-standing goal of VAMDC to be able to combine data from different databases that relate to a given state or energy level of an atom or molecule. This allows comparison of measurements or theoretical values from different sources, and also permits a fuller set of data to be built up: e.g. energies for radiative transitions from one database combined with broadening values from another. This comparison needs a common labelling-scheme for states and energy levels.

---

[56] http://theorets.univ-reims.fr, http://theorets.tsu.ru

This common labelling is already achieved for simple molecules in the "case by case" formalism of XSAMS. The tool SPECTCOL [109,110], which takes spectral data from CDMS and collisional excitation rates from BASECOL, exploits this. The labelling of states simplifies the current, theoretical understanding of the molecules sufficiently for the software to suggest states that *might* be treated as equivalent: i.e. it performs a "fuzzy" match from which an expert user can pick states to treat as true matches.

It is currently much harder to combine data on atomic states and levels. This part of the XSAMS data-model allows greater precision and flexibility in the description of a level at the cost of easy matching between databases. We hope to retain the current accuracy of the data model while augmenting it with extra notations that are less precise but more consistent between databases. Thus we may recover the ability to make "fuzzy" matches between databases. The restricted notations introduced by the PyValem software-package[57] may be relevant here.

We emphasize that we are not seeking *exact and complete*, quantum-mechanical descriptions of all states. We understand that such a description may not be possible with the current, theoretical models of atoms, which are themselves approximate and incomplete (except for the simplest atoms). Furthermore, state descriptions precise to the limits of current theory may not be unique e.g. due to underlying assumptions about coupling and state mixing. Thus, the VAMDC consortium is exploring hierarchical structures for describing atomic terms and electronic configurations. At the highest, most-precise level, these would describe the atoms to the full extent allowed by current theory. At the lower, more-approximate levels, they would allow automatic matching of data that probably describe the same states, leaving final interpretation to expert users. As a bonus, we hope to make it possible to select from the databases just the data relating to particular atomic levels and states.

3.3.3. Update of Node-Software

The common Python software on which most data-nodes are built relies on a small suite of third-party libraries. We wish to keep this infrastructure baseline constant as long as possible to enhance stability and reduce maintenance for the node operators. However, it is now necessary to upgrade to newer versions of the libraries. Notably, Django[58] v2, which is no longer supported, must be replaced with v3. This may force some changes in the customisation of each node since Django upgrades are rarely compatible in all APIs.

Provided that old and new versions of the node-software support the same version of VAMDC standards, we do not require all the data-nodes to be upgraded at the same time.

4. Applications and User Case Studies

Any database lives from the use of its data, the more the better. Many applications need concerted access to different databases, just what is provided through VAMDC. Here, we present a few of these applications that demonstrate the use and benefit of VAMDC.

4.1. CASSIS Software

CASSIS[59] (Centre d'Analyse Scientifique de Spectres Instrumentaux et Synthétiques) is a standalone software written in Java, freely delivered to the community for help in visualising, analysing and modelling observations from ground or space-based observatories. It has been developed at IRAP since 2005 and is part of the OVGSO data centre that aims at promoting VO technology. It displays any spectra (ASCII, FITS or GILDAS/CLASS format or the result from a query to any SSAP[60] VO or

[57] https://github.com/xnx/pyvalem
[58] https://www.djangoproject.com/
[59] http://cassis.irap.omp.eu
[60] http://www.ivoa.net/documents/SSA/20120210/

EPN-TAP[167] VO service from IVOA registries) and identifies atomic and molecular species through its link to databases such as JPL [15] and CDMS [74] via a SQLite database. Species identification may also proceed via a direct access to VAMDC for any available spectroscopic database using TAP (the Table Access Protocol), which is based on the IVOA format.

Modelling is performed in Kelvin units, and a conversion tool for Jansky units is available. LTE (Local Thermodynamic Equilibrium) and non-LTE modelling tools are provided. CASSIS gives $\sim$ 80 collision files constructed from the LAMDA [168] and BASECOL [29,30] databases, as well as from local files, such that their quantum numbers, Einstein coefficients, upper energy levels and rest frequencies match the request made within CASSIS. Indeed, some molecular species appear both in JPL and CDMS with sometimes different spectroscopic parameters, hindering the request for the modelling of line profiles. In the near future CASSIS will use the SPECTCOL tool [109,110], developed by the VAMDC consortium, to associate spectroscopic data provided by spectroscopic databases with collisional data provided by collisional databases to obtain the most up-to-date rate coefficients. Analysis modules for line and baseline fitting, resampling, rotational diagram analysis and χ^2 minimisation with Jython[61] scripts are available to find best-fit model parameters such as column density, T_k (or T_{ex}), $n(H_2)$ or source size.

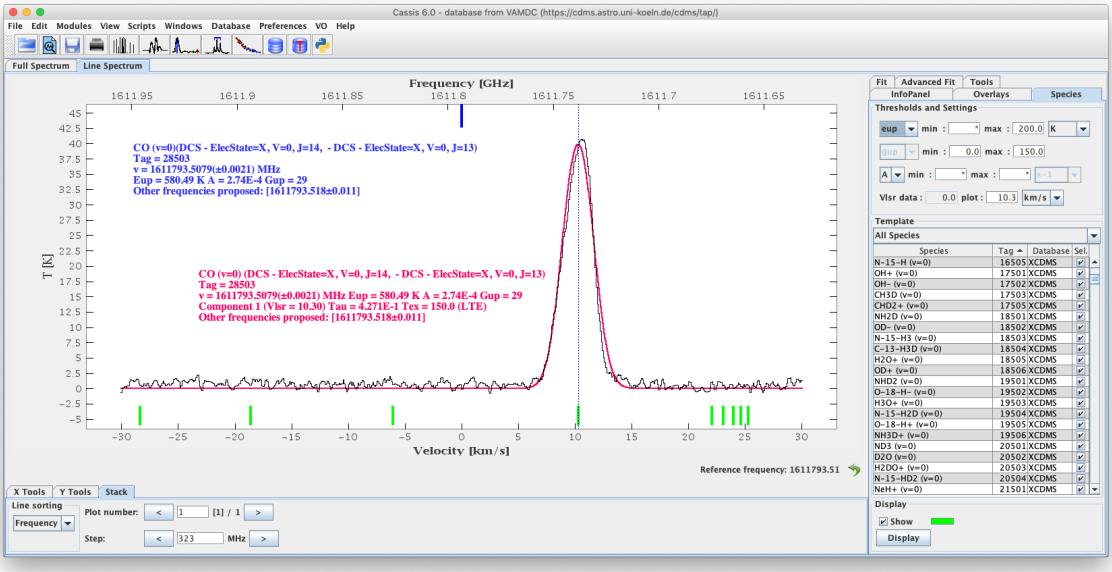

Figure 2. Example of a request made within CASSIS for an observed transition of CO (black) using the CDMS node of VAMDC for the line identification and LTE modelling (red) (see text). The right side of the figure shows the species available in the database and the possible thresholds on the Einstein coefficients and upper energy levels. The V_{LSR} is also given for the shift applied to the databases' frequencies, and the green bars below the spectrum correspond to the possible transitions within the frequency range.

Fig. 2 shows an example of an observed spectrum (in black) of a CO transition compared with LTE modelling (red) of the CO transition using the CDMS node in VAMDC. The vertical blue bar corresponds to the requested CO transition unshifted for the V_{LSR} (velocity in the standard of rest). The vertical green bars correspond to the transitions that appear in VAMDC with the thresholds on the upper energy levels and Einstein coefficients, and shift in V_{LSR} (right side of the Figure). The spectroscopic parameters are listed in blue for the observations and in red for the model with the

61 https://www.jython.org

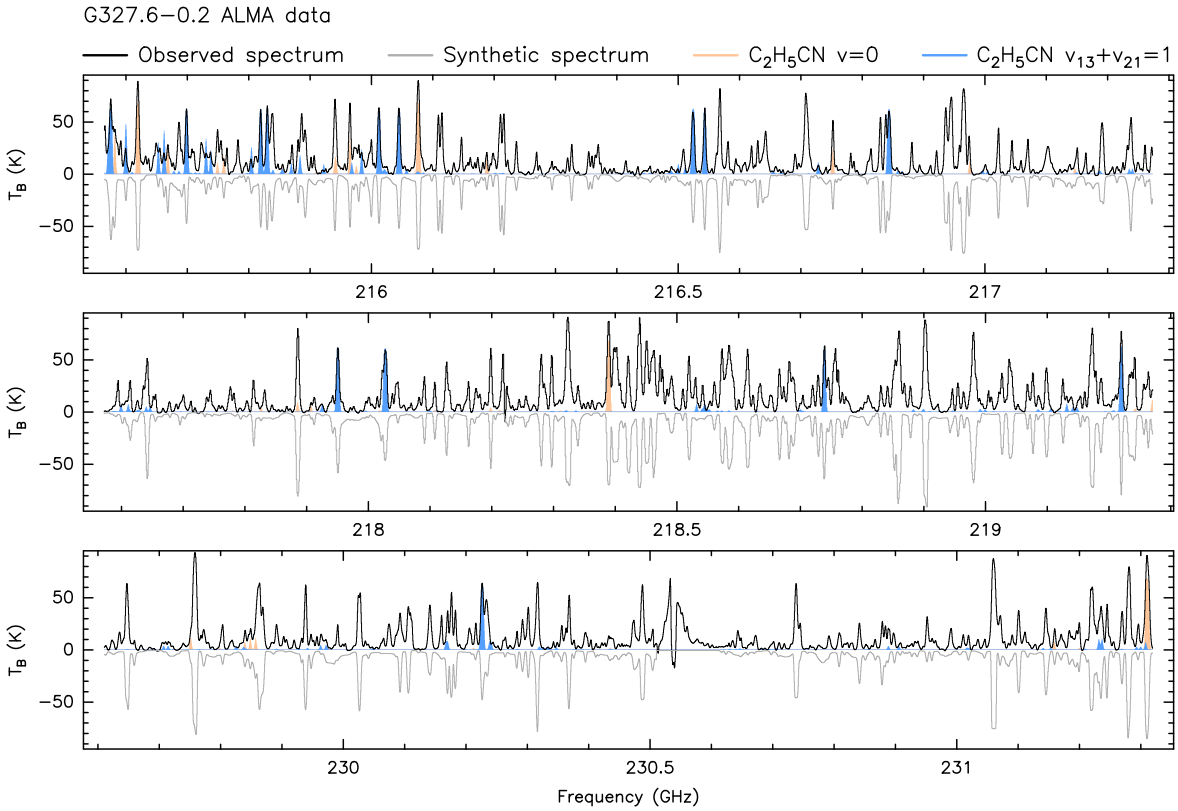

Figure 3. Spectra towards the high-mass star-forming region G327.3−0.6, with the intensity in brightness temperature unit (T_B) observed with the Atacama Large Millimeter/submillimeter Array (ALMA). The myCLASSFit function was used to model the observational data shown in black. Synthetic spectrum including all identified molecular species are shown in grey, and has the intensity multiplied by a factor −1. This study aimed at finding and quantifying ethylcyanide (C_2H_5CN) in these dense spectra. The contribution from C_2H_5CN $v = 0$ and C_2H_5CN $v_{13} + v_{21} = 1$ transitions are highlighted in orange and blue, respectively (Reproduced from [169], Copyright Elsevier (2020)).

resulting opacity and excitation temperature values. When VAMDC proposes different frequencies for a given transition (derived, predicted, experimental values) that may even originate from the same database, CASSIS selects the frequency with the minimum uncertainty.

4.2. XCLASS - eXtended CASA Line Analysis Software Suite

XCLASS[62] [170] is a full message passing interface (MPI) parallelised toolbox for the Common Astronomy Software Applications package (CASA), aimed at fitting spectral line data from astronomical sources observed both with interferometers or single dish telescopes. XCLASS models a synthetic spectrum that is automatically compared to the data with the aim of providing a measurement of physical quantities, such as the temperature, molecular abundance and velocity. Molecular data required by XCLASS are taken from an embedded SQLite3 database containing entries from CDMS [74] and JPL [15] that is populated and updated via the Python library VAMDCLIB, which queries the data directly from the VAMDC nodes. XCLASS offers the possibility of describing molecules and radio recombination lines (RRLs) in LTE and non-LTE conditions using the RADEX [63] formalism. Hereby, local overlap of lines can be taken into account as well. Furthermore, non-Gaussian line

[62] https://xclass.astro.uni-koeln.de/
[63] https://personal.sron.nl/~vdtak/radex/index.shtml

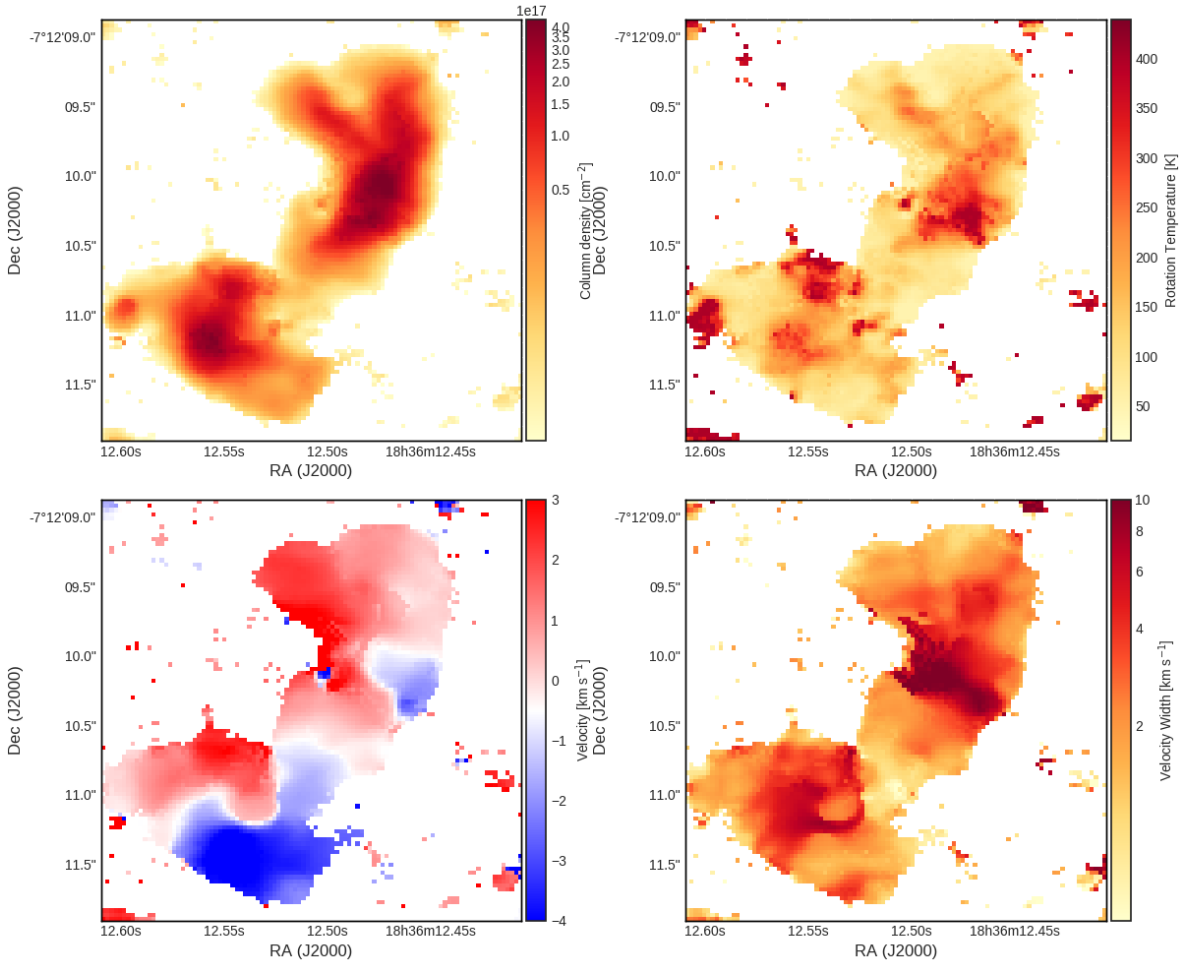

Figure 4. Example of parameter maps created by the myXCLASSMapFit function using nine transitions of CH_3OCHO simultaneously, taken from an ALMA data set of the core of G24.78 [172].

profiles such as Lorentzian, Voigt (including pressure broadening for RRLs) and Horn can be used. Different continuum contributions like dust, free–free and synchrotron emissions can also be modelled. Finally, complex source structures can be described by using sub-beam descriptions and component stacking. The toolbox contains an interface for the model optimiser package MAGIX (Modelling and Analysis Generic Interface for eXternal numerical codes) [171], which helps to find the best description of the data using a certain model; i.e. finding the parameter set that most closely reproduces the data. XCLASS can also automatically identify the molecules producing a given spectrum.

In XCLASS, the myXCLASSFit function can be used to fit multiple spectra, i.e. to fit multiple frequency ranges simultaneously in multiple files (see Fig. 3). The function returns the optimised model parameters and the corresponding modelled spectra.

In addition, the XCLASS interface contains the myXCLASSMapFit function that fits one or more complete (FITS) data cubes. For this the myXCLASSMapFit function reads in the data cube(s), extracts the spectra for each pixel and fits each spectrum separately. At the end of the whole fitting procedure, the myXCLASSMapFit function creates FITS images for each free parameter of the best fit, where each pixel corresponds to the value of the optimised parameter taken from the best fit for that pixel (see Fig. 4). Some applications of this include temperature maps, as well as first and second moment maps, which are based on the simultaneous fitting of many lines and are fairly robust against line confusion and blending of single lines—this can be a severe issue in many ALMA data sets with line-rich sources. The new CubeFit function is an extension to the myXCLASSMapFit function, and can be used to describe data cubes by physical models.

4.3. Use of Stark-B Data

From an analysis of applications of Stark broadening parameters [173], calculated with a semi-classical perturbation method [68], it was shown that the main users of Stark-B data are astronomers, who use them for investigation of A and B type stars, white dwarfs and hot stars in late evolution stages, as for example PG1159 type stars[64] [174]. The data most often used for this purpose primarily concern He I and Si II. Additional usage of these data is for plasma modelling in physics and technology, primarily for laser produced plasma with some emphasis on Stark broadening parameters of Zn I. An analysis of the citations to the Stark-B database shows that its data are also used for different analyses of regularities and systematic trends, as a source of a large number of homogeneous data obtained in the same way [175]. The Stark-B database is often cited in studies of spectral lines within laboratory plasma.

4.4. Examples of Concrete User Issues

Through the federation of resources and the adoption of common standards, the key features of the VAMDC e-infrastructure are its ability to query all resources with the same type of queries and to retrieve data identified with the same metadata. Theoretically this is a very big step towards speeding the retrieval of data in a safe and secure way. As an example, through VAMDC it is impossible to make a mistake when querying a given atomic species with its given ionisation stage: all databases will return data concerning that specific query. On the other hand, it is easy to query the wrong ionisation stage by going directly to the different databases. For example, a question from a user was: "I searched Ca^+ in the wavelength range from 4800 Å to 4950 Å directly from NIST, from VALD, and from VAMDC/NIST and VAMDC/VALD. None of the transition wavelengths seem to be identical, what shall I take?" It turned out that the user had in fact queried Ca^+ in NIST and Ca in VALD. The VAMDC query was obviously the same both for VAMDC/NIST and VAMDC/VALD, but the values seemed to differ between NIST and VAMDC/NIST. It turned out that VAMDC transfers by default wavelengths in vacuum while a direct request to NIST provides wavelengths in air and direct request to VALD provides wavelengths in vacuum. This simple example shows the difficulty that users might encounter.

In spite of major advances developed by the VAMDC federation, we have identified several user issues of such an e-infrastructure, which in turn might prevent its wider adoption by several communities. Users have their habits with queries and output formats; to a lesser extent, they question the data directly obtained from the long-established databases. Databases sometimes offer science services that facilitate the users' work; for example, helping the user to choose the right set of the database's data for their application. Other databases may include recommendations of their own data whenever, for identical processes, the database contains different values. For example, the BASECOL database keeps historical collisional sets even when new data sets are included for the same process, and data sets are labelled as recommended or not. This local feature is not currently transferred through VAMDC.

Some knowledgeable users would like that the VAMDC portal provide the functionality to choose the databases prior to specific queries on species and processes, thus avoiding unnecessary waiting time. Users deploying applications such as CASSIS (Section 4.1) would appreciate the option to make total or partial copies of the databases locally in order to speed up their application. Such copies would be coupled to a rapid search for the availability of updates. Some users would appreciate a lighter format than the XSAMS format.

We can therefore identify the following non-exhaustive and general issues to address in the future: (i) trust in a new numerical environment; (ii) the ability to quickly understand and compare data;

[64] It is the prototype of stars with a hydrogen-deficient atmosphere that is in transition between being the central star of a planetary nebula and a hot white dwarf, named after their prototype, PG1159-035.

(iii) the ability to choose the right data for its own usage and improvement/development of tools. Some of the future lines of work are outlined in Section 6. Closer collaborations between the VAMDC community and different user communities will be addressed through a VAMDC users committee whose representative will be part of our boards.

5. Current Impact of VAMDC on the Community

5.1. Node Data Standardisation and Quality Control

An interesting impact that VAMDC has on participating databases is the pressure to introduce accurate and complete descriptions of data elements stored on individual nodes. Such descriptions are necessary for uniform representation of A&M data in XSAMS. In practice the construction of the translation dictionary between the local node and XSAMS requires unambiguous interpretation of data elements. While this sounds trivial for physical parameters like energy and frequency, the task becomes increasingly more difficult when dealing with energy level classification, quantum numbers, mixed states, etc. For the nodes that store data compilations from different sources and authors (e.g. VALD) the connection to VAMDC provides a good motivation for setting internal standards. Without it, researchers maintaining the local node would be much less interested in spending time, for example, on converting or creating atomic energy level descriptions in the appropriate coupling scheme following the format that is unique and can be automatically analysed to get the main quantum numbers. Once this work is done, the XSAMS description of energy levels becomes straightforward. Standardisation also opens opportunities for verification of the basic QM selection rules, cross-comparison between data from different authors or/and data between nodes as well as trust assessment of expert data [176]. Mistakes, typos and misinterpretations discovered in this process significantly improve the quality of the data products and reduce the amounts of errors introduced between data production and its extraction through the VAMDC portal. Standards introduced by VAMDC play a decisive role in this improvement, while the practical work of course falls on the shoulders of people maintaining local nodes.

5.2. VAMDC and the FAIR Principles

VAMDC experts have been involved for several years in international data-sharing organisations (see Section 5.5) and have anchored *ante-litteram*[65] the FAIR Principles [177] into the design of the VAMDC infrastructure. The FAIR principles are implemented with a fine-grained granularity:

- Findable: data coming from the infrastructure are assigned persistently unique identifiers (Section 3.2.2), are described by rich metadata schemes and are indexed into public registries (Section 3.1.2).
- Accessible: the extraction query relies on open, documented standards (Section 3.1.1).
- Interoperable: the data extracted from VAMDC are formatted using the XSAMS standard (Section 3.1.1). Moreover VAMDC implements widely adopted international data-interoperability standards (Sections 3.1.3 and 3.2.2).
- Re-usable: the provenance and sources of all the data are documented in each data set extracted from VAMDC. Data tools are provided to convert VAMDC data into widely adopted community data formats (Section 3.1.3).

[65] An early definition of the Findable, Accessible, Interoperable and Reusable (FAIR) principles was given by Larry Lannon from CNRI in 2012 when he introduced the DAIR princpiples: the D standing for Discoverable will be replaced by the F for Findable.

5.3. Development of New Software, Databases and Portal

5.3.1. GrafOnto Collection of Scientific Plots

The construction of databases, extracted from tabular resources in spectroscopy, required VAMDC to pay attention to the spectral functions presented in the form of plots and figures. The basis for creating a database of spectral functions was a digital library of more than a thousand articles with plots and figures, containing spectral functions for the problem of continuum absorption [178], the properties of weakly bounded complexes, including atmospheric molecules [179] and the absorption cross sections in the near and far ultraviolet ranges. In these subject domains the amount of graphical information is much larger than the amount of tabular information. We identified typical examples of plots and figures under study and created ontologies [105] of graphical resources, tools for integrating plots and means for citation quality assessment [107].

At present the collection contains about 3500 original and 1055 cited primitive plots combined into 980 composite plots and 166 composite figures. The uploaded plots describe the properties of 19 molecules, 44 complexes and 66 mixtures. 2338 primitive plots characterise the properties of a water molecule and water dimer. The GrafOnto graphics system is available at its URL[66].

5.3.2. New On-Line Databases and Community Portal

The VAMDC project has directly stimulated the creation of new online databases such as TFSiCaSDa and UHeCaSDa [19], and the deployment of the theoretical PAH database within VAMDC [3,36] while motivating the development of an experimental databases for PAHs. The VAMDC infrastructure deployment also encouraged a similar strategy for experimental databases for solid-state data in SSHADE (Section 2.3.2). Collaborations with VAMDC directly led to a renewed effort to compile local data in India (see Section 2.3.5). The RADAM collaboration adopted the VAMDC standards and designed a RADAM portal[67] that offers access to the VAMDC connected databases in addition to some other databases to their community.

The Belgian repository of fundamental atomic data and stellar spectra (BRASS) [180]—which aims to provide the largest systematic and homogeneous quality assessment of atomic data to date in terms of wavelength, atomic and stellar parameter coverage—retrieved atomic data from various repositories and cross-matched these data. The majority of repositories were queried using the VAMDC e-infrastructure and the BRASS consortium thus acknowledged the efforts of the VAMDC team in homogenising the repositories as this has helped to expedite their comparisons and cross-match work.

5.4. Sustainability Issues

Today's science often seeks to treat complex applications such as the modelling of planetary atmospheres. The large amount of data needed to make realistic simulations is in sharp contrast to what a single researcher or institution is able to collect. Therefore, there is a common need and interest for large repositories of data and their accessibility by the community. Thus the need of sustainability of scientific data has been understood also by funding agencies and, on a larger scale, by science policy makers. All national and international calls for scientific projects now routinely demand that applicants provide a sustainable way to make validated scientific data from the project available to the general community for the long term, often specifying open access.

The sustainability issue occurs on different levels. At the lowest level, original data from laboratory measurements, computations or field measurements, e.g. astronomical observation data, are stored in a repository such that at least their validity can be checked after many years. The next level concerns

---

[66] http://wadis.saga.iao.ru/rdf/plot/plot.rdf/
[67] http://radamdb.mbnresearch.com/#

processed data being kept in a certain format to become available for later analysis to avoid unnecessary repetitions of work. All this requires a minimum level of documentation in order to retrieve the data. In that respect, association of data with a DOI or other identifiers is currently being developed and becoming available on national and international levels, but it requires infrastructure and organisation down to the level of individual institutions. Common standards such as the DOI are important means to reach a reasonable level of documentation and accessibility.

Databases connected to VAMDC offer a reliable and established way of supplying even further processed data, and therefore, are one of the most advanced answers to the concern of sustainability. The VAMDC consortium has agreed common standards and descriptions, which serve as a one-shop place for a plethora of atomic and molecular data. By defining common schemes for the data stored in the various databases, VAMDC has reached an invaluable level of interoperability and, therefore, a unique level of sophistication. VAMDC is an open platform that welcomes databases from all over the world and, as such, provides a focal point of atomic and molecular data. As a result, VAMDC serves as a role model to structure research data on international, national and local levels. We can cite, for instance, the DAT@OSU portal that gathers metadata for all research databases in the Bourgogne Franche-Comté region in France [181].

Sustainability issues are also related to human training and capabilities. In that respect, VAMDC has trained engineers and post-docs who are still involved in the fields of software developments, database management, development of standards for the benefit of the public sector in the VAMDC collaboration and in other domains, as well as in the private sector.

5.5. Impact on Open Science Initiatives and International Data Alliances

VAMDC delegates are active and have a high reputation in international data alliances, where they bring the VAMDC requirements and expertise about open-data sharing:

- VAMDC has been active in the Research Data Alliance since 2014 (Section 3.2.2), when VAMDC became an early pilot for the *data-citation recommendation*. Starting from 2016 VAMDC took a leading role in the RDA-Federated Identity management Interest Group: the RDA-recommendations produced by this interest Group [182] incorporate ideas and needs coming from the *Authentication, Authorisation and Accounting strategy* we developed for the VAMDC Consortium [183].
- Since 2016, we have worked with the IVOA on converging VAMDC and the IVOA atomic and molecular standards[68].

The current *Executive Director* of the Consortium represents VAMDC in the Group of European Data Experts in RDA (GEDE-RDA), and is one of the two co-chairs of this group. The aim of GEDE-RDA is to promote, foster and drive the discussions and consensus forming on creating guidelines, core components and concrete data fabric configuration building based on a bottom-up process. To achieve these goals, GEDE-RDA is composed of a group of European data professionals appointed by invitation from various European Research Infrastructures and some specialists from the Research Data Alliance. The core group includes delegates from 47 European research infrastructures. The chairs from the European Strategy Forum on Research Infrastructures (ESFRI) and the e-Infrastructures Reflection Group (eIRG) agreed to take an observation role. Through our participation in GEDE-RDA, the VAMDC experience and vision directly contribute to the definition of international standards [184] and to the definition of the European Open Science agenda [185]. Since November 2019, the VAMDC Portal, the Species Database and the Query Store (see, respectively, Sections 3.1.3, 3.2.1 and 3.2.2) have been registered as parts of the European Open Science Cloud (EOSC) hub and VAMDC is a service provider for the EOSC Market-Place.

[68] http://www.ivoa.net/documents/SLAP/, http://www.ivoa.net/documents/SLAP/

Our involvement in *Open Science* has been acknowledged at the French National level, where the VAMDC executive director has been invited to join the The Research Data College, which is responsible for defining and implementing the French Open Science data agenda.

6. Visions of the Future

When the VAMDC project was first conceived the thinking was that if we could introduce a strict and complete standard for describing atomic and molecular data, work out the transport protocol, the (standard) interface to the individual nodes and the common user portal, it would solve most of the problems limiting user access to the variety of data collections and data formats available on the web. VAMDC has made enormous progress in establishing the XSAMS standard for describing energy levels in atomic and molecular species, the node-software with its dictionaries for quick and reliable interfacing to the individual nodes and the portal. VAMDC introduced a registry for the regular indexing of node status and data content, publishing tools to encourage data providers to publish new data in VAMDC, query storage and many other tools. Essentially all the initial goals have been achieved and we have demonstrated all the functionality and reliability aspects of the new infrastructure.

Now, after several years of using and promoting VAMDC to other users, we see that our original goals, while ambitious, were not sufficient to make VAMDC the ultimate source of A&M data. The list of problems that an average VAMDC user is facing includes difficulties in automatically analysing XSAMS files, restriction on their size, complications in cross-identification between XSAMS generated by different nodes, and some missing functionality offered by the nodes through their native interfaces, etc. Thus, the work on VAMDC is far from complete, but it has a strong foundation and an extensive toolbox that can be built upon in the future.

In order to enhance the usefulness of the VAMDC system in the delivery of A&M data to a wider community of users, VAMDC needs a more "understanding" interface, much more flexible means for extracted data manipulation and new complex data selection constraints, for example, based on models of the environment. In the following sections we discuss our plans to develop VAMDC in the future.

6.1. User Interface

Our attempts to be very precise in describing atomic and molecular configurations with XML schema have had an unexpected adverse effect on the user interface implemented in the portal. At first sight the portal seems to have an amazing flexibility: one can start from specific energy levels or processes, but the flexibility is massively reduced by the rigidity of the XSAMS requirements and "restrictables" [69] supported by each node. The only field that has some rudimentary attempt in translating user input into alternative representations is the field "species" that allows auto-completion. This is made possible through the species database described in Section 3.2.1. In the future VAMDC will explore the possibility of creating/adopting an engine for interpretation of much more relaxed formulations of the user's request, using iterations gradually narrowing the requested data set to the user's needs. This more intuitive approach is complemented by the final query format available to the user (e.g. through the Query Store described in Section 3.2.2) allowing for quick automation of the process.

[69] Restrictables refer to the data types for which selection criteria can be included in the VAMDC query, https://standards.vamdc.eu/dictionary/restrictables.html

6.2. Complex Restrictions in VAMDC Request

Allowing for fuzzy definitions in the user interface through the portal may (and will) require additional levels of complexity in the request. For example, the interaction with the user may result in more than one complementary requests to be previewed before making a final decision. This can be further extended to the model-based selection, when constraints are set on properties sensitive to the environment, for example, spectral line strengths in the Earth atmosphere or in a given experimental or industrial setup. Estimating such derived properties requires additional information, such as the equation of state and partition functions, which at the moment are not offered through the VAMDC interface. Another example of such a complexity layer that could be useful is the ability to select multiple processes such as Raman transitions. Some molecular databases contain Raman transitions (e.g. MeCaSDa and SHeCaSDa).

6.3. Large Data Sets

As discussed above, the current VAMDC portal is not designed to easily handle databases with many millions or even billions of transitions. At the smaller end of these, it is possible to include the database but limit the amount of data that can be downloaded. At present really large databases are simply excluded. A possible solution could be to allow asynchronous queries followed by an e-mail sent to the user with a link to the result.

More concise data sets may require sacrificing the completeness of XSAMS. Formats that represent projections of XSAMS structure onto a single or multiple tables, independent tables such as FITS are not able to carry complex inter-relations of the data model. Indeed, FITS is an open standard mainly designed for image transport (as its name indicates). It may still be used in specific cases when requested data can be arranged in a table, for example a line list with data for computing opacity tables. In this case the FITS headers could provide the metadata explaining the meaning of numerical elements. In more general cases the complexity of relations in atomic and molecular physics calls for tree-based formats like XSAMS, that is a graph oriented data model.

Relational sets of tables could be the right compromise. The most concise but generally useful format may be an SQLite file containing the query results in the form of an independent, relational database. Supplying extracted data as miniature databases also provides an easy way to query within the data subset stored on the users' own computers. SQLite is a well-established tool with widespread support in scientific computing. Ultimately, this approach allows a user to obtain a database that is a cached copy of the data in a VAMDC node, and that can readily be refreshed from that node as data are added or changed in the node. The VAMDC infrastructure will be developed to support these cached copies and to embed them into applications, as has been demonstrated in the user case studies discussed above in Section 4.

6.4. XSAMS Manipulation

We have previously considered creating a library for XSAMS manipulations but we found this option to be insufficiently flexible. We will adopt a scripting language capable of handling one or several XSAMS data sets, allowing the creation of any simplified projection of XSAMS, unit conversion, cross-identification between XSAMS, formatting of the output, handling conditions, loops and data structures. The language should be simple enough for general users to learn (e.g. resembling Python or other popular programming/scripting languages). The language will be supported by a library of critical functions such as automatically analysing XSAMS, unit conversion, etc. This will provide a different level of flexibility for end users that could be integrated with asynchronous queries to the VAMDC request or with automated sequences of multiple requests.

6.5. Prototype of a New Bibliographic Service

We note that all the VAMDC data have a direct reference to the publications where the data have been originally presented, described and explained. The idea behind the new bibliographic service we have designed and are now prototyping is to use this bibliographic link "in the reverse direction" and discover data through publications. The resulting digital library portal, inspired by the *Astrophysics data system*[70], provides the community with a unique one-stop shop for the bibliographic information contained in VAMDC; the entry point is a tool where users may look for publications. The proposed filtering criterion for the search refinement is the author(s), title and year of publication. The results corresponding to the submitted bibliographic query are displayed in a tabular form where each line corresponds to a publication. For each publication:

- VAMDC provides the main bibliographic information complemented by the list of the VAMDC nodes containing data related to that particular publication;
- when supported by the node, VAMDC provides direct link(s) for extracting the XSAMS data related to the publication from the VAMDC node.

This new service may be used as a tool for identifying data-set overlaps between VAMDC nodes.

6.6. Semantic Search of Qualitative Tabular and Graphic Resources in Quantitative Spectroscopy

A semantic search for information resources is possible for systematic data, provided there is an explicit description of their properties in the form of ontologies (logical theories) [105]. In the Semantic Web approach [186] the rules for constructing such ontologies are defined in the languages of their specification [187]. At VAMDC the quality of spectral resources is determined by the data standardisation. This includes their analysis, validity control and trust assessment (see Section 5.1). A semantic search in molecular spectroscopy was implemented in 2007–2019 to find qualitative information resources in quantitative spectroscopy of atmospheric molecules [188], qualitative energy levels and molecular transitions [71][189,190], as well as spectral functions [191], used to solve the problems of continuum absorption and to describe the properties of weakly bounded complexes. The accumulated experience allowed systematic organisation of the semantic searches for high-quality spectral resources, states and transitions, as well as spectral functions in atomic, ionic and solid-state spectroscopy that describe the processes of absorption and emission.

A crucial task in the coming years is the organisation of a semantic search for spectral data in applied fields (astronomy and atmospheric optics of exoplanets), in which atoms, ions, molecules and complexes play an important role. The systematisation of these subjects will simplify the search for necessary spectral information. The main challenge that needs to be overcome is the construction of a template that allows the user to submit their requirements to be directly implemented as a set of queries to a semantic site. For this, it is necessary to build typical ontologies [105] of applied domains. The crucial task here is a reduction problem, defining a structure for each of the applications, which requires cooperation with experts in these applied domains.

6.7. Visualisation and Data Access from Python

Good data visualisation is important, particularly in data-set exploration. For larger data sets, dynamic display becomes essential, wherein the user scrolls through the data and zooms into areas of interest. Currently VAMDC provides some static displays of data (generated on the servers as web pages) and some dynamic displays (generated by a program on the user's computer using data downloaded from VAMDC). Neither of these approaches works well with large data sets; static pages are not sufficient and downloading large and rich data sets to plot only a small fraction of the data is frustratingly slow.

---

[70] https://ui.adsabs.harvard.edu
[71] http://wadis.saga.iao.ru/saga2/ontology-3-atomic-molecules/

Plotting large data sets is a common problem in scientific computing, and in recent years some solutions have been developed. The Python libraries Bokeh[72], Holoviews[73] and Datashader[74] can be combined to build plots that can be generated on a server, viewed in a web-browser and that are usefully interactive even with large data. The graphics packages also work well in Jupyter notebooks[75]. To exploit this, we should provide code that allows such a notebook to extract data from VAMDC, either via a downloaded XSAMS-file or directly from data nodes.

6.8. Link Data to Evaluation Ratings

We foresee handling the issue of linking data to evaluation ratings performed by review panels. The intention is to facilitate user choices among data sets provided by the various databases. This will involve a scientific organisation within VAMDC, collaborations with evaluation bodies and panels and, finally, technical developments in order to use the reviewed information.

6.9. Open Science for the Future

VAMDC is collaborating with the EOSC Secretariat to study an innovative *Flexible Semantic Mapping Framework* to achieve a scientific and interdisciplinary interoperability between the services within the European Open Science Cloud. The VAMDC consortium has explored the development of innovative services for automatic data quality assessment based on the interlinking between data and scientific papers (see Section 3.2.2) in combination with Artificial Intelligence and semantic techniques. In the context of Open Science, this approach may be adopted by any other discipline. The VAMDC consortium is evaluating the opportunity to put these ideas into effect as part of a future European project proposal.

6.10. Quality Status and Future

It should be stressed that the VAMDC infrastructure does not collectively address the issue of the data scientific quality. The scientific quality of the data taken from the e-infrastructure is that of the individual databases. The VAMDC infrastructure is currently a technical platform that ensures that the exchanged data are compliant with the standards, in particular with the data-format standards. As described in Section 5.1 the standardisation imposed quality/coherence checks at the individual nodes, and those checks were beneficial to the overall quality of the outputs. The data nodes must be compliant with an internal quality chart voted by the board[76]. Some specific quality improvements specify that references must be attached to the data, that the transferred data are timestamped and that the various queries are stored in the "Query Store" to ensure the tracing of the data. When uncertainties are available in the databases, this information is transferred through VAMDC.

In addition, as described above, it is planned to attach external rating reviews to data: rating reviews might concern comparison of several sets of data for a single range of applications or might concern a single set of data to provide its range of applicability. Obviously we should be aware that reviews are only a guide, they depend upon the group of people who performed the review, they can rapidly become out of date and should be updated. VAMDC offers an open infrastructure that allows third parties to establish some data assessment and to provide such services. VAMDC collects data and facilitates the comparison of data from different databases that might contain identical quantities: since all VAMDC nodes produce XSAMS formatted output data, it is easy for a user to identify overlaps between databases, e.g. by identifying when the same data units have different values.

---

[72] https://bokeh.org
[73] http://holoviews.org
[74] https://datashader.org
[75] https://jupyter.org/
[76] http://www.vamdc.org/structure/how-to-join-us/internal-regulations/

We would like to stress that VAMDC provides support to the work of scientists, but that users should be pro-active with VAMDC to enrich the scientific information. The system and service can be steadily improved and augmented by collecting and providing more central statistics about the content of the attached databases. Currently the species database allows the display of the species contained in the individual databases, but it could additionally list the processes and their range of applicability. The bibliographic service will display the papers attached to the data from the different databases, thus providing another type of central information system. Additional useful central information would be to display what fraction of data content in each database is accessible through VAMDC (sometimes it is 100% of data, sometimes it is not).

Finally some users would certainly like to have a "one-stop shop" for their application and such an effort can be made by users since VAMDC offers the technical possibility to easily build secondary data sets (see the discussion of impact on the BRASS database in Section 5.3.2). The VAMDC infrastructure could also provide the means to display such "user oriented secondary data collections" if the users wish to make them publicly available. The VAMDC consortium could also be the place where such demands are made and then collaborative projects could provide such "user oriented data sets". The communities should be aware that the VAMDC effort has already been huge, that the current people maintain efficiently the databases and the infrastructure with little manpower and financial resources, and that communities must get involved/engaged with VAMDC if they want to ensure the service they want.

6.11. Sustainability

The above technological and scientific innovations are driving forces for the sustainability of the VAMDC e-infrastructure. However, the VAMDC consortium must address the long-term issues of renewing the scientific and technical people involved in the curation of atomic and molecular data and of maintaining the leadership of these activities at/across institutes and in/across countries.

7. Conclusions

VAMDC is a sophisticated infrastructure that makes large sets of data publicly available in a common format.

A fundamental feature of VAMDC is its ability to connect more databases from atomic and molecular domains already described in VAMDC, but also from other domains. Therefore, it serves as an ideal platform for sustainability of scientific data.

A crucial key to the evolution of VAMDC is its usage in application software and the involvement of users in the orientation of the VAMDC e-infrastructure platform and services. Here, we have presented a few user case studies to illustrate this.

A striking point in the evolution of VAMDC is its involvement in data management at the highest international level. This places the VAMDC e-infrastructure as an international prototype and example, where general concepts are developed and applied, thus providing to the international and national communities diverse high-level expertise in data management and handling.

In this paper we have demonstrated once again the strength and coherence of the VAMDC concept and the quality and the innovation in the development of the standards and software solutions that the VAMDC consortium has delivered. It summarises the advancement of a ten-year effort of a large group of leading scientists. The paper has also shown how the VAMDC continues to question the current status and to identify current weaknesses, while outlining how it will upgrade its services (Sections 4.4 and 6) and make future improvements (see Section 6).

Finally, we provide help to newcomers and to users through detailed documentation and tutorials [77]. A 'helpdesk' [78] led by experts in the VAMDC portal and tools answers questions and links the user with appropriate organisers of relevant databases as well as supports the inclusion of new nodes into VAMDC.

Author Contributions: software and standards, D. Albert, Y.A. Ba, Y.L. Babikov, P. Bollard, M.L. Dubernet, M. Emoto, C. Endres, A. Fazliev, J.M. Glorian, I.E. Gordon, C. Hill, R. Kochanov, P.A. Loboda, A. Markwick, T. Marquart, C. Mendoza, N. Moreau, S.V. Morozov, T. Möller, G. Mulas, Y. Pakhomov, J. Penguen, J. Postler, A.I. Privezentsev, Y. Ralchenko, C. Richard, G. Rixon, T. Ryabchikova, P. Schilke, K. Smith, B. Schmitt, I.Y. Skobelev, C. Vastel, Veljko Vujčić, C.M. Zwölf;

data curation, B. Antony,V. Boudon, F. Delahaye, G. del Zanna, M.S. Dimitrijeviíc, B.J. Drouin, M.L. Dubernet, F. Duensing, A. Fazliev, I.E. Gordon, P. Gratier, C. Hill, D. Jevremović, C. Joblin, D.-H. Kwon, E. Krishnakumar, G. Leto, P.A. Loboda, A.A. Lukashevskaya, O.M. Lyulin, B.P. Marinković, A. Markwick, N.J. Mason, C. Mendoza,T.J. Millar, G. Mulas,I. Murakami, Y. Pakhomov, P. Palmeri, V.I. Perevalov, N. Piskunov, J. Postler, P. Quinet, Y. Ralchenko, Y.-J. Rhee, L.S. Rothman, E. Roueff, T. Ryabchikova, S. Sahal-Bréchot, P., Scheier, S. Schlemmer, B. Schmitt, V.A. Srecković, E. Stempels, S.A. Tashkun, J. Tennyson, V.G. Tyuterev, V. Wakelam, C. Zeippen;

writing–original draft preparation of section 2, all co-authors; writing–original draft preparation of section 3, T. Marquart, N. Moreau, G. Rixon, C.M. Zwölf ; writing–original draft preparation of section 4, M. S. Dimitrijevi, M.L. Dubernet, T. Möller, C. Vastel; writing–original draft preparation of section 5, V. Boudon, M.L. Dubernet, A. Fazliev, C. Joblin, N.J. Mason, N. Piskunov, S. Schlemmer, C.M. Zwölf; writing–original draft preparation of section 6, V. Boudon, M.L. Dubernet, A. Fazliev, N. Mason, N. Piskunov, G. Rixon, S. Schlemmer, J. Tennyson, N. Walton, C.M. Zwölf ; proof-reading; B.J. Drouin, C. Hill, N. Mason, G. Rixon, L.S. Rothman, S. Schlemmer, J. Tennyson; writing–review and editing, V. Boudon, M.L. Dubernet, C. Endres, C. Mendoza, N. Piskunov, G. Rixon, S. Schlemmer, J. Tennyson, C.M. Zwölf; funding acquisition, all co-authors; All authors have read and agreed to the published version of the manuscript.

Funding: The work on the ozone data was supported by the Russian Science Foundation (project RNF-19-12-00171). Yu.L.B. acknowledges the support from the Academic D.Mendeleev program of Tomsk State University. Collaboration with A.Barbe, S.Mikhailenko and E.Starikova in spectra analyses for S&MPO is acknowledged. VB and CR wish to thank the OSU THETA of Franche-Comté Bourgogne and DatOSU project for their support. The ExoMol project has been supported by ERC Advanced Investigator Project 267219, and STFC Projects No. ST/M001334/1 and ST/R000476/1.The HITRAN database is supported through the NASA Aura and PDART grants NNX17AI78G and NNX16AG51G, respectively. The implementation of the Query Store was partially funded by the European Project RDA EU3 (funded under H2020-EINFRA-2014-2, project ID: 653194). A portion of the research described in this paper was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA, US government sponsorship acknowledged. The work on the Spectr-W^3 update was supported by the Russian Foundation for Basic Research (project Nr. 14-07-00863) and the Russian Academy of Sciences Presidium Fundamental Research Program Nr. 13. Developments of the interface with the VAMDC portal to the NIFS database is partly supported by the NIFS program of Promoting Research by Networking among Institutions (grant No.01411702). EK acknowledges Raja Ramanna Fellowship from Dept. of Atomic Energy, Govt. of India. PEARL-KAERI database has been partly supported by the National Research Foundation of Korea (NRF-2019R1A2C1009529). Support for VAMDC has been provided by the UKRI/STFC under grant ST/M007731/1. The Paris team received fundings from Paris Observatory through its Paris Data Center (PADC), from CNRS through the overheads of the VAMDC project, from LERMA through the overheads of the SUP@VAMDC projects, from ASOV (Action Spécifique VO France). TJM is grateful to the STFC for support through grants ST/M007766/1 and ST/P000312/1. The SSHADE project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements No 654208 (Europlanet 2020 RI), No 871149 (Europlanet 2024 RI) and from the Observatoire des Sciences de l'Univers de Grenoble and INSU. Support for the VAMDC consortium has been provided through the VAMDC and the SUP@VAMDC projects funded under the "Combination of Collaborative Projects and Coordination and Support Actions" Funding Scheme of The Seventh Framework Program. Call topic: INFRA-2008-1.2.2 and INFRA-2012 Scientific Data Infrastructure. Grant Agreement numbers: 239108 and 313284.

Acknowledgments: We wish to acknowledge all people who contributed to the management of the databases over the years and to the VAMDC e-infrastructure.

Conflicts of Interest: "The authors declare no conflict of interest."

Abbreviations

[77] http://www.vamdc.org/activities/research/
[78] email: support@vamdc.eu

1267 The following abbreviations are used in this manuscript:
1268

A&M	Atomic and Molecular
AGB	Asymptotic Giant Branch
ALMA	Atacama Large Millimeter/submillimeter Array
AMBDAS	Atomic and Molecular Bibliographic Data System
AMDIS	Atomic and Molecular Data Interactive System
API	Application Programming Interface
ASCII	American Standard Code for Information Interchange
ASD	Atomic Spectra Database
ASD-1000	Acetylene Spectroscopic Databank
BASECOL	Rovibrational Collisional Database
BeamDB	Belgrade Electron-Atom/Molecule DataBase
1269 CAS	Chemical Abstracts Service
CASSIS	Centre d'Analyse Scientifique de Spectres Instrumentaux et Synthétiques
CDMS	Cologne Database for Molecular Spectroscopy
CDSD-x	Carbon Dioxide Spectroscopic Databank
CNRS	Centre National de la Recherche Scientifique
CRM	Collisional–Radiative Model
DESIRE	Database on Sixth Row Elements
DOI	Digital Object Identifier
DR	Dielectronic Recombination
DREAM	Database of Rare Earths At Mons University
EII	Electron Impact Ionisation
EOSC	European Open Science Cloud

FAIR	Findable, Accessible, Interoperable, Reusable
FITS	Flexible Image Transport System
GEDE-RDA	Group of European Data Expert in the Research Data Alliance
HAPI	HITRAN Application Programming Interface
HITEMP	High-Temperature Molecular Database
HITRAN	High-resolution Transmission Molecular Absorption Database
HTML	Hypertext Markup Language
IAMDB	Indian Atomic and Molecular DataBase
IDEADB	Innsbruck Dissociative Electron Attachment DataBase
IDP	Interplanetary Dust Particle
INAF	Istituto Nazionale di Astrofisica
InChI	International Chemical Identifier
ISM	Interstellar medium
IP	Iron Project
IPPJ	Institute of Plasma Physics, Nagoya University, Nagoya, Japan
IRAP	Institut de Recherche en Astrophysique et Planétologie
IVOA	International Virtual Observatory Alliance
JIHT RAS	Joint Institute for High Temperatures of the Russian Academy of Sciences
JPL	Jet Propulsion Laboratory
KAERI	Korea Atomic Energy Research Institute
KIDA	KInetic Database for Astrochemistry
LASp	Laboratorio di Astrofisica Sperimentale
LIBS	Laser-Induced Breakdown Spectroscopy
MolLIST	Molecular Line Lists, Intensities and Spectra
NDSD-1000	Nitrogen Dioxide Spectroscopic Databank
NIFS	National Institute for Fusion Science
NIST	National Institute of Standards and Technology
NOEMA	Northern Extended Millimeter Array
NOSD-1000	Nitrous Oxide Spectroscopic Databank
OAC	Osservatorio Astrofisico di Catania
OP	Opacity Project
OVGSO	Observatoire Virtuel du Grand Sud-Ouest
PAH	Polycyclic Aromatic Hydrocarbon
PEARL	Photonic Electronic Atomic Reaction Laboratory
QM	Quantum Mechanics
RADEX	Statistical Equilibrium Radiative Transfer Code
RADMC-3D	Multi-Purpose Radiative Transfer Tool
RDA	Research Data Alliance
RATRAN	Radiative Transfer and Excitation of Molecular Lines Code
RDA	Research Data Alliance
REST	Representational State Transfer
RFNC	Russian Federal Nuclear Centre
SAMP	Simple Application Messaging Protocol
SESAM	Spectroscopy Database Dedicated to Electronic Spectra of Diatomic Molecules
SSHADE	Solid Spectroscopy Hosting Architecture of Databases and Expertise
S&MPO	Spectroscopy and Molecular Properties of Ozone
SSAP	Simple Spectral Access Protocol, an IVOA standard
SQL	Structured Query Language
TAP	Table Access Protocol, an IVOA standard
UDfA	UMIST Database for Astrochemistry
UMIST	University of Manchester Institute of Science and Technology
URL	Uniform Resource Locator
VALD	Vienna Atomic Line Database
VAMDC	Virtual Atomic and Molecular Data Centre
VNIITF	All-Russian Institute of Technical Physics
VO	Virtual Observatory
VSS2	VAMDC SQL Sub-set 2
XSAMS	XML Schema for Atoms, Molecules and Solids

References

1. 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.; Mason, H.E.; Mason, N.J.; Mendoza, C.; Mulas, G.; Millar, T.J.; Nuñez, L.A.; Perevalov, V.I.; Piskunov, N.; Ralchenko, Y.; Rixon, G.; Rothman, L.S.; Roueff, E.; Ryabchikova, T.A.; Ryabtsev, A.; Sahal-Bréchot, S.; Schmitt, B.; Schlemmer, S.; Tennyson, J.; Tyuterev, V.G.; Walton, N.A.; Wakelam, V.; Zeippen, C.J. Virtual atomic and molecular data centre. *J. Quant. Spectrosc. Radiat. Transfer* **2010**, *111*, 2151–2159. doi:10.1016/j.jqsrt.2010.05.004.
2. Zwolf, C.; Dubernet, M.L.; Ba, Y.A.; Moreau, N. Experience and feedbacks from the sustainability for the virtual atomic and molecular data centre E-infrastructure. IST-Africa Conference Proceedings, 2014, 2014, pp. 1–9. doi:10.1109/ISTAFRICA.2014.6880621.
3. Dubernet, M.L.; Antony, B.; Ba, Y.A.; Babikov, Y.; Bartschat, K.; Boudon, V.; Braams, B.; Chung, H.K.; Daniel, F.; Delahaye, F.; Del Zanna, G.; de Urquijo, J.; Dimitrijevic, M.; Domaracka, A.; Doronin, M.; Drouin, B.; Endres, C.; Fazliev, A.; Gagarin, S.; Gordon, I.; Gratier, P.; Heiter, U.; Hill, C.; Jevremovic, D.; Joblin, C.; Karsprzak, A.; Krishnakumar, E.; Leto, G.; Loboda, P.A.; Louge, T.; Maclot, S.; Marinkovic, B.; Markwick Kemper, A.; Marquart, T.; Mason, H.; Mason, N.; Mendoza, C.; Mihajlov, A.; Millar, T.; Moreau, N.; Mulas, G.; Pakhomov, Y.; Palmeri, P.; Pancheshnyi, S.; Perevalov, V.I.; Piskunov, N.; Postler, J.; Quinet, Pascaland Quintas SÃ¡nchez, E.L.; Ralchenko, Y.; Rhee, Y.J.; Rixon, G.; Rothman, L.; Roueff, E.; Ryabchikova, T.; Sahal-Brechot, S.; Scheier, P.; Schlemmer, S.; Schmitt, B.; Stempels, E.; Tashkun, S.; Tennyson, J.; Tyuterev, V.; Vujcic, V.; Wakelam, V.; Walton, N.; Zatsarinny, O.; Zeippen, C.; Zwölf, C.M. The Virtual Atomic and Molecular Data Centre (VAMDC) Consortium. *Journal of Physics B: Atomic, Molecular and Optical Physics* **2016**, *49*, 074003. doi:10.1088/ 0953-4075/49/7/074003.
4. Emoto, M.; Murakami, I.; Kato, D.; Yoshida, M.; Kato, M.; Imazu, S. Improvement of the NIFS Atom and Molecular Database. *Atoms* **2019**, *7*, 91. doi:10.3390/atoms7030091.
5. Murakami, I.; Kato, D.; Kato, M.; Sakaue, H.A. Atomic and molecular database and data evaluation activities at the National Institute for Fusion Science. *Fusion Sci. Technol.* **2013**, *63*, 400–405. doi:10.13182/FST13-A16448.
6. Ryabchikova, T.; Piskunov, N.; Kurucz, R.L.; Stempels, H.C.; Heiter, U.; Pakhomov, Y.; Barklem, P.S. A major upgrade of the VALD database. *Physica Scripta* **2015**, *90*, 054005. doi:10.1088/0031-8949/90/5/054005.
7. Kramida, A.; Yu. Ralchenko.; Reader, J.; and NIST ASD Team. NIST Atomic Spectra Database (ver. 5.7.1), [Online]. Available: `https://physics.nist.gov/asd` [2020, July 9]. National Institute of Standards and Technology, Gaithersburg, MD., 2019.
8. Skobelev, I.Y.; Loboda, P.A.; Gagarin, S.V.; Ivliev, S.V.; Kozlov, A.I.; Morozov, S.V.; S. A. Pikuz, J.; Pikuz, T.A.; Popova, V.V.; Faenov, A.Y. The Spectr-W^3 Database on the Spectroscopic Properties of Atoms and Ions. *Optics and Spectroscopy* **2016**, *120*, 507–514.
9. Del Zanna, G.; Young, P.R. Atomic data for plasma spectroscopy: the CHIANTI database, improvements and challenges. *Submitted to Atoms* **2020**.
10. Hummer, D.G.; Berrington, K.A.; Eissner, W.; Pradhan, A.K.; Saraph, H.E.; Tully, J.A. Atomic data from the IRON project. I. Goals and methods. *A&A* **1993**, *279*, 298–309.
11. Cunto, W.; Mendoza, C.; Ochsenbein, F.; Zeippen, C.J. TOPbase at the CDS. *A&A* **1993**, *275*, L5–L8.
12. Sahal-Bréchot, S.; Dimitrijević, M.S.; Moreau, N. Virtual Laboratory Astrophysics and the Stark-B database VAMDC node: A resource for electron and ion impact widths and shifts of isolated lines. *Journal of Physics: Conference Series* **2020**, *1412*, 132052.
13. Roueff, E.L.; Sahal-Bréchot, S.; M., D.; Moreau, N.; Abgrall, H. The spectroscopic atomic and molecular databases at Paris Observatory. *Atoms* **2020**, *8*, 36–48.
14. Endres, C.P.; Schlemmer, S.; Schilke, P.; Stutzki, J.; Mueller, H.S.P. The Cologne Database for Molecular Spectroscopy, CDMS, in the Virtual Atomic and Molecular Data Centre, VAMDC. *J. Mol. Spect.* **2016**, *327*, 95–104.
15. Pickett, H.M.; Poynter, R.L.; Cohen, E.A.; Delitsky, M.L.; Pearson, J.C.; Müller, H.S.P. Submillimeter, millimeter and microwave spectral line catalog. *J. Quant. Spectrosc. Radiat. Transfer* **1998**, *60*, 883–890. doi:10.1016/S0022-4073(98)00091-0.
16. 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.; Flaud, J.M.; Gamache, R.R.; Hodges, J.T.; Jacquemart, D.; Perevalov, V.I.;

Perrin, A.; Shine, K.P.; Smith, M.A.H.; Tennyson, J.; Toon, G.C.; Tran, H.; Tyuterev, V.G.; Barbe, A.; Császár, A.G.; Devi, V.M.; Furtenbacher, T.; Harrison, J.J.; Hartmann, J.M.; Jolly, A.; Johnson, T.J.; Karman, T.; Kleiner, I.; Kyuberis, A.A.; Loos, J.; Lyulin, O.M.; Massie, S.T.; Mikhailenko, S.N.; Moazzen-Ahmadi, N.; Müller, H.S.P.; Naumenko, O.V.; Nikitin, A.V.; Polyansky, O.L.; Rey, M.; Rotger, M.; Sharpe, S.W.; Sung, K.; Starikova, E.; Tashkun, S.A.; Vander Auwera, J.; Wagner, G.; Wilzewski, J.; Wcisło, P.; Yu, S.; Zak, E.J. The *HITRAN* 2016 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transfer* **2017**, *203*, 3–69. doi:10.1016/j.jqsrt.2017.06.038.

17. Babikov, Y.; Mikhailenko, S.; Barbe, A.; Tyuterev, V. S&MPO - An information system for ozone spectroscopy on the WEB. *J. Quant. Spectrosc. Radiat. Transfer* **2014**, *145*, 169–196. doi:10.1016/j.jqsrt.2014.04.024.
18. Richard, C.; Boudon, V.; Rotger, M. Calculated spectroscopic databases for the VAMDC portal: New molecules and improvements. *J. Quant. Spectrosc. Radiat Transfer* **2020**, *251*, 107096–1–107096–13.
19. Boudon, V.; Manceron, L.; Richard, C. High-resolution spectroscopy and analysis of the ν_3, ν_4 and $2\nu_4$ bands of SiF_4 in natural isotopic abundance. *J. Quant. Spectrosc. Radiat Transfer* **2020**, *253*, 107114–1–107114–20.
20. Tashkun, S.A.; Perevalov, V.I.; Gamache, R.R.; Lamouroux, J. CDSD-296, high-resolution carbon dioxide spectroscopic databank: An update. *J. Quant. Spectrosc. Radiat. Transfer* **2019**, *228*, 124–131. doi:10.1016/j.jqsrt.2019.03.001.
21. Tashkun, S.A.; Perevalov, V.I. CDSD-4000: High-resolution, high-temperature carbon dioxide spectroscopic databank. *J. Quant. Spectrosc. Radiat. Transfer* **2011**, *112*, 1403–1410. doi:10.1016/j.jqsrt.2011.03.005.
22. Tashkun, S.A.; Perevalov, V.I.; Lavrentieva, N.N. NOSD-1000, the high-temperature nitrous oxide spectroscopic databank. *J. Quant. Spectrosc. Radiat. Transfer* **2016**, *177*, 43–48. doi:10.1016/j.jqsrt.2015.11.014.
23. Lukashevskaya, A.A.; Lavrentieva, N.N.; Dudaryonok, A.C.; Perevalov, V.I. NDSD-1000: High-resolution, high-temperature Nitrogen Dioxide Spectroscopic Databank. *J. Quant. Spectrosc. Radiat. Transfer* **2016**, *184*, 205–217. doi:10.1016/j.jqsrt.2016.07.014.
24. Lyulin, O.M.; Perevalov, V.I. ASD-1000: High-resolution, high-temperature acetylene spectroscopic databank. *J. Quant. Spectrosc. Radiat. Transfer* **2017**, *201*, 94–103. doi:10.1016/j.jqsrt.2017.06.032.
25. Akhlyostin, A.; Apanovich, Z.; Fazliev, A.; Kozodoev, A.; Lavrentiev, N.; Privezentsev, A.; Rodimova, O.; Voronina, S.; Császár, A.; Tennyson, J. The current status of the W@DIS information system. 22-nd International Symposium Atmospheric and Ocean Optics: Atmospheric Physics, 2016. doi:10.1117/12.2249235.
26. Wakelam, V.; Herbst, E.; Loison, J.C.; Smith, I.W.M.; Chandrasekaran, V.; Pavone, B.; Adams, N.G.; Bacchus-Montabonel, M.C.; Bergeat, A.; Béroff, K.; Bierbaum, V.M.; Chabot, M.; Dalgarno, A.; van Dishoeck, E.F.; Faure, A.; Geppert, W.D.; Gerlich, D.; Galli, D.; Hébrard, E.; Hersant, F.; Hickson, K.M.; Honvault, P.; Klippenstein, S.J.; Le Picard, S.; Nyman, G.; Pernot, P.; Schlemmer, S.; Selsis, F.; Sims, I.R.; Talbi, D.; Tennyson, J.; Troe, J.; Wester, R.; Wiesenfeld, L. A KInetic Database for Astrochemistry (KIDA). *Astrophys. J. Suppl.* **2012**, *199*, 21, [1201.5887]. doi:10.1088/0067-0049/199/1/21.
27. Wakelam, V.; Loison, J.C.; Herbst, E.; Pavone, B.; Bergeat, A.; Béroff, K.; Chabot, M.; Faure, A.; Galli, D.; Geppert, W.D.; Gerlich, D.; Gratier, P.; Harada, N.; Hickson, K.M.; Honvault, P.; Klippenstein, S.J.; Le Picard, S.D.; Nyman, G.; Ruaud, M.; Schlemmer, S.; Sims, I.R.; Talbi, D.; Tennyson, J.; Wester, R. The 2014 KIDA Network for Interstellar Chemistry. *Astrophys. J. Suppl.* **2015**, *217*, 20, [1503.01594]. doi:10.1088/0067-0049/217/2/20.
28. McElroy, D.; Walsh, C.; Markwick, A.J.; Cordiner, M.A.; Smith, K.; Millar, T.J. The UMIST database for astrochemistry 2012. *A&A* **2013**, *550*, A36, [arXiv:astro-ph.SR/1212.6362]. doi:10.1051/0004-6361/201220465.
29. Dubernet, M.L.; Alexander, M.H.; Ba, Y.A.; Balakrishnan, N.; Balança, C.; Ceccarelli, C.; Cernicharo, J.; Daniel, F.; Dayou, F.; Doronin, M.; Dumouchel, F.; Faure, A.; Feautrier, N.; Flower, D.R.; Grosjean, A.; Halvick, P.; Kłos, J.; Lique, F.; McBane, G.C.; Marinakis, S.; Moreau, N.; Moszynski, R.; Neufeld, D.A.; Roueff, E.; Schilke, P.; Spielfiedel, A.; Stancil, P.C.; Stoecklin, T.; Tennyson, J.; Yang, B.; Vasserot, A.M.; Wiesenfeld, L. BASECOL2012: A collisional database repository and web service within the Virtual Atomic and Molecular Data Centre (VAMDC). *A&A* **2013**, *553*, A50. doi:10.1051/0004-6361/201220630.
30. Ba, Y.; Dubernet, M.; Moreau, N.; Zwolf, C. BASECOL (basecol.vamdc.org) new technical design. *Submitted to Atoms* **2020**.

31. Vujčič, V.; Jevremović, D.; Mihajlov, A.; Ignjatović, L.M.; Srećković, V.; Dimitrijević, M.; Malović, M. MOL-D: A Collisional Database and Web Service within the Virtual Atomic and Molecular Data Center. *J. Astrophys. Astron.* **2015**, *36*, 0.
32. Jevremović, D.; Srećković, V.A.; Marinković, B.P.; Vujčić, V. Databases for collisional and radiative processes in small molecules needed for spectroscopy use in astrophysics. *Contrib. Astron. Obs. Skalnate Pleso* **2020**, *50*, 44–54.
33. Marinković, B.P.; Vujčić, V.; Sushko, G.; Vudragović, D.; Marinković, D.B.; Đorđević, S.; Ivanović, S.; Nešić, M.; Jevremović, D.; Solov'yov, A.V.; Mason, N.J. Development of collisional data base for elementary processes of electron scattering by atoms and molecules. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **2015**, *354*, 90 – 95. 26th International Conference on Atomic Collisions in Solids, doi:https://doi.org/10.1016/j.nimb.2014.12.039.
34. Gorfinkiel, J.D.; Ptasinska, S. Electron scattering from molecules and molecular aggregates of biological relevance. *Journal of Physics B: Atomic, Molecular and Optical Physics* **2017**, *50*, 182001. doi:10.1088/1361-6455/aa8572.
35. Schmitt, B.; Bollard, P.; Damien, A.; Garenne, A.; Bonal, L.; Gorbacheva, M.; the SSHADE partner's consortium. SSHADE: Solid Spectroscopy Hosting Architecture of Databases and Expertise and its databases. online, 2018. database Infrastructure, publisher: OSUG Data Center, https://www.sshade.eu, doi:10.26302/SSHADE.
36. Malloci, G.; Joblin, C.; Mulas, G. On-line database of the spectral properties of polycyclic aromatic hydrocarbons. *Chem. Phys.* **2007**, *332*, 353–359, [arXiv:astro-ph/astro-ph/0701254]. doi:10.1016/j.chemphys.2007.01.001.
37. Chubb, K.L.; Rocchetto, M.; Yurchenko, S.N.; Min, M.; Waldmann, I.; Barstow, J.K.; Molliere, P.; Al-Refaie, A.F.; Phillips, M.; Tennyson, J. The ExoMolOP Database: Cross-sections and K-tables for Molecules of Interest in High-Temperature Exoplanet Atmospheres. *A&A* **2020**.
38. Fivet, V.; Quinet, P.; Palmeri, P.; Biémont, E.; Xu, H. Transition probabilities and lifetimes for atoms and ions from the sixth row of the periodic table and the database DESIRE. *J. Electr. Spectrosc. Rel. Phen.* **2007**, *250*, 156–158.
39. Quinet, P.; Palmeri, P. Current status and developments of the Atomic Database on Rare-Earths at Mons University (DREAM). *Atoms* **2020**, *8*, 18.
40. Kwon, D.H.; K.-B. Chai.; and KAERI Atomic Data Center Team. Photonic Electronic Atomic Reaction Laboratory database, [Online]. Available: `http://pearl.kaeri.re.kr`.
41. Takayanagi, K.; Suzuki, H. Cross sections for Atomic Processes Vol.1. Processes Involving Hydrogen Isotopes, Their Ions, Electrons and Photons. Technical Report IPPJ-DT-48, Institute of Plasma Physics, Nagoya University, Nagoya, Japan, 1975.
42. Takayanagi, K.; Suzuki, H. Cross sections for Atomic Processes Vol.2. Processes Involving Hydrogen and Helium Isotopes, Their Ions, Electrons and Photons. Technical Report IPPJ-DT-50, Institute of Plasma Physics, Nagoya University, Nagoya, Japan, 1976.
43. Kato, T.; Itikawa, Y.; Kanada, Y.; Watanabe, R. Database and Retrieval-Display System of Atomic Data for Fusion. *Physica Scripta* **1981**, *23*, 198–201. doi:10.1088/0031-8949/23/2/027.
44. Murakami, I.; Kato, M.; Emoto, M.; Kato, D.; Sakaue, H.A.; Kawate, T. NIFS Atomic and Molecular Database for collision processes. *submitted to Atoms* **2020**.
45. Biémont, E.; Palmeri, P.; Quinet, P. A New Database of Astrophysical Interest. *Astrophys. Space Sci.* **1999**, *269*, 635–637. doi:10.1023/A:1017049314691.
46. Pakhomov, Y.V.; Ryabchikova, T.A.; Piskunov, N.E. Hyperfine Splitting in the VALD Database of Spectral-line Parameters. *Astronomy Reports* **2019**, *63*, 1010–1021, [arXiv:astro-ph.IM/1911.03189]. doi:10.1134/S1063772919120047.
47. Ryabchikova, T.; Piskunov, N.; Sitnova, T. Evaluation of the Recent Atomic Data for Fe uc(i) Lines Based on Solar and Stellar Spectra. *Solar Physics* **2019**, *294*, 156. doi:10.1007/s11207-019-1543-2.
48. Pakhomov, Y.V.; Ryabchikova, T.A. Evolution of spectral line parameters database VALD. *INASAN Science Reports* **2019**, *4*, 79–84. doi:10.26087/INASAN.2019.4.2.014.
49. Ryabchikova, T.; Pakhomov, Y.; Piskunov, N. VALD: The Meeting Point of Data Producers and Data Users. *Galaxies* **2018**, *5*, 93. doi:10.3390/galaxies6030093.

50. Pakhomov, Y.V.; Ryabchikova, T.A. Virtual atomic and molecular data centre VAMDC. Stars and Satellites, Proceedings of the Memorial Conference Devoted to A.G. Masevich 100th Anniversary; Shustov, B.M.; Wiebe, D.S., Eds., 2018, pp. 402–408. doi:10.26087/INASAN.2018.2.2.065.
51. Pakhomov, Y.V. Method of spectral wavelength calibration using the VALD database. Stars and Satellites, Proceedings of the Memorial Conference Devoted to A.G. Masevich 100th Anniversary; Shustov, B.M.; Wiebe, D.S., Eds., 2018, pp. 374–379. doi:10.26087/INASAN.2018.2.2.060.
52. Pakhomov, Y.; Piskunov, N.; Ryabchikova, T. VALD3: Current Developments. Stars: From Collapse to Collapse; Balega, Y.Y.; Kudryavtsev, D.O.; Romanyuk, I.I.; Yakunin, I.A., Eds., 2017, Vol. 510, *Astronomical Society of the Pacific Conference Series*, p. 518, [arXiv:astro-ph.IM/1710.10854].
53. Kramida, A. NIST Atomic Spectroscopy Bibliographic Databases, [Online]. Available: `URL https://www.nist.gov/pml/nist-atomic-spectra-bibliographic-databases` [2020, July 9]. National Institute of Standards and Technology, Gaithersburg, MD., 2019.
54. Kramida, A.; Olsen, K.; Yu. Ralchenko. NIST LIBS Database, [Online]. Available: `https://physics.nist.gov/PhysRefData/ASD/LIBS/libs-form.html` [2020, July 9]. National Institute of Standards and Technology, Gaithersburg, MD., 2019.
55. Dere, K.P.; Landi, E.; Mason, H.E.; Monsignori Fossi, B.C.; Young, P.R. CHIANTI - an atomic database for emission lines. *Astrophys. J. Suppl.* **1997**, *125*, 149.
56. Del Zanna, G.; Dere, K.P.; Young, P.R.; Landi, E.; Mason, H.E. CHIANTI - An atomic database for emission lines. Version 8. *A&A* **2015**, *582*, A56. doi:10.1051/0004-6361/201526827.
57. Dere, K.P.; Del Zanna, G.; Young, P.R.; Landi, E.; Sutherland, R.S. CHIANTI—An Atomic Database for Emission Lines. XV. Version 9, Improvements for the X-Ray Satellite Lines. *Astrophys. J. Suppl.* **2019**, *241*, 22. doi:10.3847/1538-4365/ab05cf.
58. Seaton, M.J. Atomic data for opacity calculations. I. General description. *J. of Phys. B Atom. Molec. Phys.* **1987**, *20*, 6363–6378. doi:10.1088/0022-3700/20/23/026.
59. Cunto, W.; Mendoza, C. The Opacity Project - the Topbase Atomic Database. *Revista Mexicana de Astronomía y Astrofísica* **1992**, *23*, 107.
60. Berrington, K.A.; Eissner, W.B.; Norrington, P.H. RMATRX1: Belfast atomic R-matrix codes. *Computer Physics Communications* **1995**, *92*, 290–420.
61. Ballance, C.P.; Griffin, D.G. Relativistic radiatively damped R-matrix calculation of the electron-impact excitation of W46+. *J. Phys. B* **2006**, *39*, 3617. doi:10.1088/0953-4075/39/17/017.
62. Badnell, N.R. A Breit-Pauli distorted wave implementation for AUTOSTRUCTURE. *Computer Physics Communications* **2011**, *182*, 1528–1535.
63. Badnell, N.R.; Bautista, M.A.; Butler, K.; Delahaye, F.; Mendoza, C.; Palmeri, P.; Zeippen, C.J.; Seaton, M.J. Updated opacities from the Opacity Project. *M. N. R. A. S.* **2005**, *360*, 458–464, [astro-ph/0410744]. doi:10.1111/j.1365-2966.2005.08991.x.
64. Delahaye, F.; Zwölf, C.M.; Zeippen, C.J.; Mendoza, C. IPOPv2 online service for the generation of opacity tables. *J. Quant. Spectrosc. Radiat. Transfer* **2016**, *171*, 66–72, [arXiv:astro-ph.SR/1511.07260]. doi:10.1016/j.jqsrt.2015.11.010.
65. Sahal-Bréchot, S.; Dimitrijević, M.S.; Moreau, N.; Ben Nessib, N. The Stark-B Database as a Resource for "Stark" Widths and Shifts Data: State of Advancement and Program of Development. *Advances in Space Research* **2014**, *54*, 1148–1151.
66. Sahal-Bréchot, S.; Dimitrijević, M.S.; Moreau, N.; Ben Nessib, N. The Stark-B Database VAMDC node: a Repository for Spectral Line Broadening and Shifts Due to Collisions with Charged Particles. *Physica Scripta* **2015**, *50*, 054008.
67. Dimitrijević, M.S.; Sahal-Bréchot, S.; Moreau, N. The Stark-B Database, A Node of Virtual Atomic and Molecular Data Center (VAMDC). *Publications of the Astronomical Observatory of Belgrade* **2018**, *98*, 285–288.
68. Sahal-Brćhot.; Dimitrijević, M.S.; Ben Nessib, N. Widths and Shifts of Isolated Lines of Neutral and Ionized Atoms Perturbed by Collisions With Electrons and Ions: An Outline of the Semiclassical Perturbation (SCP) Method and of the Approximations Used for the Calculations. *Atoms* **2014**, *2*, 225–252.
69. Dimitrijević, M.S.; Konjević, N. Stark broadening of doubly and triply ionized atoms. *J. Quant. Spectrosc. Radiat. Transfer* **1980**, *24*, 451–459.
70. Sahal-Bréchot, S.; Dimitrijević, M.S.; Ben Nessib, N. Comparisons and comments on electron and ion impact profiles of spectral lines. *Baltic Astronomy* **2011**, *20*, 523–530.

71. Pearson, J.C.; Mueller, H.S.P.; Pickett, H.M.; Cohen, E.A.; Drouin, B.J. Introduction to submillimeter, millimeter and microwave spectral line catalog. *J. Quant. Spectrosc. Radiat. Transfer* **2010**, *111*, 1614–1616. doi:{10.1016/j.jqsrt.2010.02.002}.
72. Endres, C.; Schlemmer, S.; Drouin, B.; Pearson, J.; Müller, H.S.P.; Schilke, P.; Stutzki, J. Improved Infrastucture for Cdms and JPL Molecular Spectroscopy Catalogues. 69th International Symposium on Molecular Spectroscopy, 2014.
73. Müller, H.S.P.; Thorwirth, S.; Roth, D.A.; Winnewisser, G. The Cologne Database for Molecular Spectroscopy, CDMS. *A&A* **2001**, *370*, L49–L52. doi:10.1051/0004-6361:20010367.
74. Müller, H.S.P.; Schlöder, F.; Stutzki, J.; Winnewisser, G. The Cologne Database for Molecular Spectroscopy, CDMS: a useful tool for astronomers and spectroscopists. *Journal of Molecular Structure* **2005**, *742*, 215–227. doi:10.1016/j.molstruc.2005.01.027.
75. McClatchey, R.; Benedict, W.; Clough, S.; Burch, D.; Calfee, R.; Fox, K.; Rothman, L.; Garing, J. AFCRL Atmospheric Absorption Line Parameters Compilation. Technical Report 434, Optical Physics Laboratory, 1973.
76. Rothman, L.S. Atmospheric absorption-line parameters: the HITRAN data base (A). *Journal of the Optical Society of America (1917-1983)* **1983**, *73*.
77. Rothman, L.S.; Gamache, R.R.; Goldman, A.; Brown, L.R.; Toth, R.A.; Pickett, H.M.; Poynter, R.L.; Flaud, J.M.; Camy-Peyret, C.; Barbe, A.; Husson, N.; Rinsland, C.P.; Smith, M.A.H. The HITRAN database: 1986 edition. *Applied Optics* **1987**, *26*, 4058. doi:10.1364/AO.26.004058.
78. Rothman, L.; Gamache, R.; Tipping, R.; Rinsland, C.; Smith, M.; Benner, D.; Devi, V.; Flaud, J.M.; Camy-Peyret, C.; Perrin, A.; Goldman, A.; Massie, S.; Brown, L.; Toth, R. The HITRAN molecular database: Editions of 1991 and 1992. *Journal of Quantitative Spectroscopy and Radiative Transfer* **1992**, *48*, 469–507. doi:10.1016/0022-4073(92)90115-K.
79. Rothman, L.; Rinsland, C.; Goldman, A.; Massie, S.T.; D.P., E.; Flaud, J.M.; Perrin, A.; Camy-Peyret, C.; Dana, V.; Mandin, J.; J., S.; McCann, A.; Gamache, R.; Wattson, R.; Yoshino, K.; Chance, K.; Jucks, K.; Brown, L.; Nemtchinov, V.; Varanasi, P. The HITRAN Molecular Spectroscopic Database and HAWKS (HITRAN Atmospheric Workstation): 1996 Edition. *Journal of Quantitative Spectroscopy and Radiative Transfer* **1998**, *60*, 665–710. doi:10.1016/S0022-4073(98)00078-8.
80. Rothman, L.; Barbe, A.; Chris Benner, D.; Brown, L.; Camy-Peyret, C.; Carleer, M.; Chance, K.; Clerbaux, C.; Dana, V.; Devi, V.; Fayt, A.; Flaud, J.M.; Gamache, R.; Goldman, A.; Jacquemart, D.; Jucks, K.; Lafferty, W.; Mandin, J.Y.; Massie, S.; Nemtchinov, V.; Newnham, D.; Perrin, A.; Rinsland, C.; Schroeder, J.; Smith, K.; Smith, M.; Tang, K.; Toth, R.; Vander Auwera, J.; Varanasi, P.; Yoshino, K. The HITRAN molecular spectroscopic database: edition of 2000 including updates through 2001. *Journal of Quantitative Spectroscopy and Radiative Transfer* **2003**, *82*, 5–44. doi:10.1016/S0022-4073(03)00146-8.
81. Rothman, L.; Jacquemart, D.; Barbe, A.; Chris Benner, D.; Birk, M.; Brown, L.; Carleer, M.; Chackerian, C.; Chance, K.; Coudert, L.; Dana, V.; Devi, V.; Flaud, J.M.; Gamache, R.; Goldman, A.; Hartmann, J.M.; Jucks, K.; Maki, A.; Mandin, J.Y.; Massie, S.; Orphal, J.; Perrin, A.; Rinsland, C.; Smith, M.; Tennyson, J.; Tolchenov, R.; Toth, R.; Vander Auwera, J.; Varanasi, P.; Wagner, G. The HITRAN 2004 molecular spectroscopic database. *Journal of Quantitative Spectroscopy and Radiative Transfer* **2005**, *96*, 139–204. doi:10.1016/j.jqsrt.2004.10.008.
82. Rothman, L.S.; Gordon, I.E.; Barbe, A.; Benner, D.C.; Bernath, P.F.; Birk, M.; Boudon, V.; Brown, L.R.; Campargue, A.; Champion, J.P.; Chance, K.; Coudert, L.H.; Dana, V.; Devi, V.M.; Fally, S.; Flaud, J.M.; Gamache, R.R.; Goldman, A.; Jacquemart, D.; Kleiner, I.; Lacome, N.; Lafferty, W.J.; Mandin, J.Y.; Massie, S.T.; Mikhailenko, S.N.; Miller, C.E.; Moazzen-Ahmadi, N.; Naumenko, O.V.; Nikitin, A.V.; Orphal, J.; Perevalov, V.I.; Perrin, A.; Predoi-Cross, A.; Rinsland, C.P.; Rotger, M.; Šimečková, M.; Smith, M.A.H.; Sung, K.; Tashkun, S.A.; Tennyson, J.; Toth, R.A.; Vandaele, A.C.; Vander Auwera, J. The HITRAN 2008 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transfer* **2009**, *110*, 533–572. doi:10.1016/j.jqsrt.2009.02.013.
83. Rothman, L.S.; Gordon, I.E.; Babikov, Y.; Barbe, A.; Chris Benner, D.; Bernath, P.F.; Birk, M.; Bizzocchi, L.; Boudon, V.; Brown, L.R.; Campargue, A.; Chance, K.; Cohen, E.A.; Coudert, L.H.; Devi, V.M.; Drouin, B.J.; Fayt, A.; Flaud, J.M.; Gamache, R.R.; Harrison, J.J.; Hartmann, J.M.; Hill, C.; Hodges, J.T.; Jacquemart, D.; Jolly, A.; Lamouroux, J.; Le Roy, R.J.; Li, G.; Long, D.A.; Lyulin, O.M.; Mackie, C.J.; Massie, S.T.; Mikhailenko, S.; Müller, H.S.P.; Naumenko, O.V.; Nikitin, A.V.; Orphal, J.; Perevalov, V.; Perrin, A.;

Polovtseva, E.R.; Richard, C.; Smith, M.A.H.; Starikova, E.; Sung, K.; Tashkun, S.; Tennyson, J.; Toon, G.C.; Tyuterev, V.G.; Wagner, G. The HITRAN2012 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transfer* **2013**, *130*, 4–50. doi:10.1016/j.jqsrt.2013.07.002.

84. Hill, C.; Gordon, I.E.; Kochanov, R.V.; Barrett, L.; Wilzewski, J.S.; Rothman, L.S. HITRANonline: An online interface and the flexible representation of spectroscopic data in the HITRAN database. *Journal of Quantitative Spectroscopy and Radiative Transfer* **2016**, *177*, 4–14. doi:10.1016/j.jqsrt.2015.12.012.
85. Skinner, F.M.; Gordon, I.E.; Hill, C.; Hargreaves, R.J.; Lockhart, K.E.; Rothman, L.S. Referencing Sources of Molecular Spectroscopic Data in the Era of Data Science: Application to the HITRAN and AMBDAS Databases. *Atoms* **2020**, *8*, 16. doi:10.3390/atoms8020016.
86. Rothman, L.; Gordon, I.; Barber, R.; Dothe, H.; Gamache, R.; Goldman, A.; Perevalov, V.; Tashkun, S.; Tennyson, J. HITEMP, the high-temperature molecular spectroscopic database. *Journal of Quantitative Spectroscopy and Radiative Transfer* **2010**, *111*, 2139–2150. doi:10.1016/j.jqsrt.2010.05.001.
87. Li, G.; Gordon, I.E.; Rothman, L.S.; Tan, Y.; Hu, S.M.; Kassi, S.; Campargue, A.; Medvedev, E.S. Rovibrational Line Lists for Nine Isotopologues of the CO Molecule in the X $^1\Sigma^+$ Ground Electronic State. *The Astrophysical Journal Supplement Series* **2015**, *216*, 15. doi:10.1088/0067-0049/216/1/15.
88. Hargreaves, R.J.; Gordon, I.E.; Rothman, L.S.; Tashkun, S.A.; Perevalov, V.I.; Lukashevskaya, A.A.; Yurchenko, S.N.; Tennyson, J.; Müller, H.S. Spectroscopic line parameters of NO, NO_2, and N_2O for the HITEMP database. *Journal of Quantitative Spectroscopy and Radiative Transfer* **2019**, *232*, 35–53. doi:10.1016/J.JQSRT.2019.04.040.
89. Hargreaves, R.J.; Gordon, I.E.; Rey, M.; Nikitin, A.V.; Tyuterev, V.G.; Kochanov, R.V.; Rothman, L.S. An Accurate, Extensive, and Practical Line List of Methane for the HITEMP Database. *Astrophys. J. Suppl.* **2020**, *247*, 55. doi:10.3847/1538-4365/ab7a1a.
90. Kochanov, R.; Gordon, I.; Rothman, L.; Shine, K.; Sharpe, S.; Johnson, T.; Wallington, T.; Harrison, J.; Bernath, P.; Birk, M.; Wagner, G.; Le Bris, K.; Bravo, I.; Hill, C. Infrared absorption cross-sections in HITRAN2016 and beyond: Expansion for climate, environment, and atmospheric applications. *Journal of Quantitative Spectroscopy and Radiative Transfer* **2019**, *230*, 172–221. doi:10.1016/J.JQSRT.2019.04.001.
91. Karman, T.; Gordon, I.E.; van der Avoird, A.; Baranov, Y.I.; Boulet, C.; Drouin, B.J.; Groenenboom, G.C.; Gustafsson, M.; Hartmann, J.M.; Kurucz, R.L.; Rothman, L.S.; Sun, K.; Sung, K.; Thalman, R.; Tran, H.; Wishnow, E.H.; Wordsworth, R.; Vigasin, A.A.; Volkamer, R.; van der Zande, W.J. Update of the HITRAN collision-induced absorption section. *Icarus* **2019**, *328*, 160–175. doi:10.1016/J.ICARUS.2019.02.034.
92. Kochanov, R.; Gordon, I.; Rothman, L.; Wcisło, P.; Hill, C.; Wilzewski, J. HITRAN Application Programming Interface (HAPI): A comprehensive approach to working with spectroscopic data. *Journal of Quantitative Spectroscopy and Radiative Transfer* **2016**, *177*, 15–30. Symposium on High Resolution Molecular Spectroscopy (HighRus-2015), Tomsk, Russia, doi:https://doi.org/10.1016/j.jqsrt.2016.03.005.
93. Mikhailenko, S.; Barbe, A. High resolution infrared spectrum of $^{16}O_3$: The 3600-4300 cm^{-1} range reinvestigated. *J. Quant. Spectrosc. Radiat. Transfer* **2020**, *244*, 106823. doi:https://doi.org/10.1016/j.jqsrt.2019.106823.
94. Starikova, E.; Barbe, A.; Tyuterev, V. The v3 bands of $^{17}O^{17}O^{18}O$ and $^{17}O^{18}O^{17}O$ ozone isotopomers. *J. Quant. Spectrosc. Radiat. Transfer* **2019**, *232*, 87–92. doi:10.1016/j.jqsrt.2019.05.002.
95. Barbe, A.; Starikova, E.; De Backer, M.; Tyuterev, V. Analyses of infrared FT spectra of asymmetric ozone isotopologue $^{16}O^{16}O^{18}O$ in the range 950-3850 cm^{-1}. *J. Quant. Spectrosc. Radiat. Transfer* **2018**, *218*, 231–247. doi:10.1016/j.jqsrt.2018.06.022.
96. Tyuterev, V.; Kochanov, R.; Tashkun, S. Accurate ab initio dipole moment surfaces of ozone: First principle intensity predictions for rotationally resolved spectra in a large range of overtone and combination bands. *J. Chem. Phys.* **2017**, *146*, 064304. doi:10.1063/1.4973977.
97. Tyuterev, V.; Barbe, A.; Jacquemart, D.; Janssen, C.; Mikhailenko, S.; Starikova, E. Ab initio predictions and laboratory validation for consistent ozone intensities in the MW, 10 and 5μm ranges. *J. Chem. Phys.* **2019**, *150*. doi:10.1063/1.5089134.
98. Boudon, V.; Champion, J.P.; Gabard, T.; Loëte, M.; Rotger, M.; Wenger, C. Spherical Top Theory and Molecular Spectra. In *Handbook of High-Resolution Spectroscopy*; Quack, M.; Merkt, F., Eds.; Wiley: Chichester, West Sussex, United Kingdom, 2011; Vol. 3, pp. 1437–1460.

99. Wenger, C.; Boudon, V.; Rotger, M.; Sanzharov, J.P.; Champion, J.P. XTDS and SPVIEW: Graphical tools for the analysis and simulation of high-resolution molecular spectra. *Journal of Molecular Spectroscopy* **2008**, *251*, 102–113.
100. Boudon, V.; Champion, J.P.; Gabard, T.; Loëte, M.; Michelot, F.; Pierre, G.; Rotger, M.; Wenger, C.; Rey, M. Symmetry-adapted tensorial formalism to model rovibrational and rovibronic spectra of molecules pertaining to various point groups. *J. Mol. Spectrosc.* **2004**, *228*, 620–634.
101. Ba, Y.A.; Wenger, C.; Surleau, R.; Boudon, V.; Rotger, M.; Daumont, L.; Bonhommeau, D.A.; Tyuterev, V.G.; Dubernet, M.L. MeCaSDa and ECaSDa: Methane and ethene calculated spectroscopic databases for the virtual atomic and molecular data centre. *Journal of Quantitative Spectroscopy and Radiative Transfer* **2013**, *130*, 62–68.
102. Person, W.; Krohn, B. Coriolis intensity perturbations of the ν_4 band of SF_6. *J. Mol. Spectrosc.* **1983**, *98*, 229–257.
103. Aldridge, J.; Brock, E.; Filip, H.; Flicker, H.; Fox, K.; Galbraith, H.; Holland, R.; Kim, K.; Krohn, B.; Magnuson, D.; II, W.M.; owell, R.; Patterson, C.; Person, W.; Smith, D.; Werner, G. Measurement and analysis of the infrared-active stretching fundamental (ν_3) of UF_6. *J. Chem. Phys.* **1985**, *83*, 34–48.
104. De Roure, D.; Jennings, N.R.; Shadbolt, N.R. The Semantic Grid: A Future e-Science Infrastructure. In *Grid Computing: Making the Global Infrastructure a Reality*; EPSRC/DTI Core e-Science Programme, 2001; pp. 437 – 470. doi:10.1002/0470867167.
105. Gruber, T. Toward principles for the design of ontologies used for knowledge sharing? *International Journal of Human-Computer Studies* **1995**, *43*, 907–928. doi:10.1006/ijhc.1995.1081.
106. Lavrentiev, N.; Privezentsev, A.; Fazliev, A. Tabular and Graphic Resources in Quantitative Spectroscopy. In *Data Analytics and Management in Data Intensive Domains. DAMDID/RCDL 2018. Communications in Computer and Information Science*; Springer, 2019; Vol. 1003, pp. 55–69. doi:10.1007/978-3-030-23584-0_4.
107. Lavrentiev, N.; Privezentsev, A.; Fazliev, A. Applied Ontologies for Managing Graphic Resources in Quantitative Spectroscopy. In *Data Analytics and Management in Data Intensive Domains DAMDID/RCDL 2019. Communications in Computer and Information Science*; Springer, 2020; Vol. 1223, chapter 6. doi:10.1007/978-3-030-51913-1_6.
108. Millar, T.J.; Bennett, A.; Rawlings, J.M.C.; Brown, P.D.; Charnley, S.B. Gas phase reactions and rate coefficients for use in astrochemistry - The UMIST ratefile. *Astron. Astrophys. Sup.* **1991**, *87*, 585–619.
109. Dubernet, M.L.; Nenadovic, L. SPECTCOL: Spectroscopic and Collisional Data Retrieval. Astrophysics Source Code Library, record ascl:1111.005, 2011, p. 11005.
110. Ba, Y.A.; Dubernet, M.L. The VAMDC SPECTCOL tool. *to be submitted to Atoms* **2020**.
111. Srećković, V.A.; Ignjatović, L.M.; Jevremović, D.; Vujčić, V.; Dimitrijević, M.S. Radiative and Collisional Molecular Data and Virtual Laboratory Astrophysics. *Atoms* **2017**, *5*, 31.
112. Ignjatovic, L.M.; Sreckovic, V.; Dimitrijevic, M. Photoionization of the alkali molecular ions in geo-cosmical plasmas. *Contrib. Astron. Obs. Skalnaté Pleso* **2020**, *50*, 187–192.
113. Marinković, B.P.; Srećković, V.A.; Vujčić, V.; Ivanović, S.; Uskoković, N.; Nešić, M.; Ignjatović, L.M.; Jevremović, D.; Dimitrijević, M.S.; Mason, N.J. BEAMDB and MOLD—Databases at the Serbian Virtual Observatory for Collisional and Radiative Processes. *Atoms* **2019**, *7*, 11.
114. Marinković, B.P.; Jevremović, D.; Srećković, V.A.; Vujčić, V.; Ignjatović, L.M.; Dimitrijević, M.S.; Mason, N.J. BEAMDB and MolD – databases for atomic and molecular collisional and radiative processes: Belgrade nodes of VAMDC. *The European Physical Journal D* **2017**, *71*. doi:10.1140/epjd/e2017-70814-6.
115. Predojević, B.; Šević, D.; Marinković, B.P.; McEachran, R.P.; Blanco, F.; García, G.; Brunger, M.J. Joint theoretical and experimental study on elastic electron scattering from bismuth. *Phys. Rev. A* **2020**, *101*, 032704. doi:10.1103/PhysRevA.101.032704.
116. McEachran, R.P.; Marinković, B.P.; García, G.; White, R.D.; Stokes, P.W.; Jones, D.B.; Brunger, M.J. Integral Cross Sections for Electron–Zinc Scattering over a Broad Energy Range (0.01–5000 eV). *Journal of Physical and Chemical Reference Data* **2020**, *49*, 013102. doi:10.1063/1.5135573.
117. Marinković, B.; Bredehöft, J.; Vujčić, V.; Jevremović, D.; Mason, N. Rosetta Mission: Electron Scattering Cross Sections—Data Needs and Coverage in BEAMDB Database. *Atoms* **2017**, *5*, 46. doi:10.3390/atoms5040046.

118. Marinković, B.P.; Pejčev, V.; Filipović, D.M.; Šević, D.; Milosavljević, A.R.; Milisavljević, S.; Rabasović, M.S.; Pavlović, D.; Maljković, J.B. Cross section data for electron collisions in plasma physics. *Journal of Physics: Conference Series* **2007**, *86*, 012006. doi:10.1088/1742-6596/86/1/012006.
119. Denifl, S.; Garcia, G.; Huber, B.A.; Marinković, B.P.; Mason, N.; Postler, J.; Rabus, H.; Rixon, G.; Solovyov, A.V.; Suraud, E.; Yakubovich, A.V. Radiation damage of biomolecules (RADAM) database development: current status. *Journal of Physics: Conference Series* **2013**, *438*, 012016. doi:10.1088/1742-6596/438/1/012016.
120. Bartschat, K.; Zatsarinny, O. Benchmark calculations of atomic data for plasma and lighting applications. *Plasma Sources Science and Technology* **2011**, *20*, 024012. doi:10.1088/0963-0252/20/2/024012.
121. Nina, A.; Radovanovic, M.; Srećković, V. *Integrations of satellite and ground-based observations and multi-disciplinarity in research and prediction of different types of hazards in Solar system, Book of abstracts*; Geographical Institute "Jovan Cvijić" of Serbian Academy of Sciences and Arts, 2019, Belgrade, 2019.
122. Baratta, G.A.; Domingo, M.; G., F.; Leto, G.; Palumbo, M.E.; Satorre, M.A.; Strazzulla, G. Ion irradiation of CH4-containing icy mixtures. *Nuclear Instruments and Methods in Physics Section B* **2003**, *209*, 283.
123. Leto, G.; Baratta, G.A. Ly-alpha photon induced amorphization of Ic water ice at 16 Kelvin. Effects and quantitative comparison with ion irradiation. *A&A* **2003**, *397*.
124. Tennyson, J.; Yurchenko, S.N. ExoMol: molecular line lists for exoplanet and other atmospheres. *M. N. R. A. S.* **2012**, *425*, 21–33. doi:10.1111/j.1365-2966.2012.21440.x.
125. Tennyson, J.; Hill, C.; Yurchenko, S.N. Data structures for ExoMol: Molecular line lists for exoplanet and oth er atmospheres. 6^{th} international conference on atomic and molecular data and their applications ICAMDATA-2012. AIP, New York, 2013, Vol. 1545, *AIP Conference Proceedings*, pp. 186–195. doi:10.1063/1.4815853.
126. Tennyson, J.; Yurchenko, S.N.; Al-Refaie, A.F.; Barton, E.J.; Chubb, K.L.; Coles, P.A.; Diamantopoulou, S.; Gorman, M.N.; Hill, C.; Lam, A.Z.; Lodi, L.; McKemmish, L.K.; Na, Y.; Owens, A.; Polyansky, O.L.; Rivlin, T.; Sousa-Silva, C.; Underwood, D.S.; Yachmenev, A.; Zak, E. The ExoMol database: molecular line lists for exoplanet and other hot atmospheres. *J. Mol. Spectrosc.* **2016**, *327*, 73–94. doi:10.1016/j.jms.2016.05.002.
127. Tennyson, J.; Yurchenko, S.N.; abd V. H. J. Clark a nd K. L. Chubb, A.F.A.R.; Conway, E.K.; Dewan, A.; Gorman, M.N.; Hill, C.; Lynas-Gray, A.E.; Mellor, T.; McKemmish, L.K.; Owens, A.; Polyansky, O.L.; Semenov, M.; Somogyi, W.; Tinetti, G.; y, A.U.; Waldmann, I.; Wang, Y.; Wright, S.; Yurchenko, O.P. The 2020 release of the ExoMol database: molecular line lists for exoplanet and other hot atmospheres. *J. Quant. Spectrosc. Radiat. Transfer* **2020**, *255*, 107228. doi:10.1016/j.jqsrt.2020.107228.
128. Lacis, A.A.; Oinas, V. *J. Geophys. Res.* **1991**, *96*, 9027.
129. Bernath, P.F. MoLLIST: Molecular Line Lists, Intensities and Spectra. *J. Quant. Spectrosc. Radiat. Transfer* **2020**, *240*, 106687. doi:10.1016/j.jqsrt.2019.106687.
130. Hill, C.; Dubernet, M.L.; Endres, C.; Karwasz, G.; Marinković, B.; Marquart, T.; Heinola, K.; Zwolf, C.M.; Moreau, N. Classification of Processes in Plasma Physics. Technical Report INDC(NDS)-0812, IAEA, 2020.
131. Biémont, E.; Palmeri, P.; Quinet, P. A new database of astrophysical interest. *Astrophys. Space Sci.* **1999**, *635*, 269–270.
132. Quinet, P.; Palmeri, P.; Biémont, E.; McCurdy, M.; Rieger, G.; Pinnington, E.; Wickliffe, M.; Lawler, J. Experimental and theoretical lifetimes, branching fractions and oscillator strengths in Lu II. *M. N. R. A. S.* **1999**, *307*, 934.
133. Quinet, P.; Palmeri, P.; Biémont, E.; Li, Z.; Zhang, Z.; Svanberg, S. Radiative lifetime measurements and transition probability calculations in lanthanide ions. *J. Alloy. Compd.* **2002**, *344*.
134. Rahman, M.; Krishnakumar, E. Absolute partial and total electron ionization cross sections of uracil. *Int. J. Mass Spetrom* **2015**, *392*, 145–153. doi:10.1016/j.ijms.2015.10.003.
135. Rahman, M.; Krishnakumar, E. Electron ionization of DNA bases. *J. Chem. Phys.* **2016**, *144*, 161102. doi:10.1063/1.4948412.
136. Tadsare, V. Dissociative electron attachment to organic molecules. PhD thesis, Tata Institute of Fundamental Research, TIFR, Mumbai, India, 2018.
137. Sinha, N.; Gupta, D.; Antony, B. Electron impact ionisation cross sections for complex molecules. *J. Phys. B: At. Mol. Opt. Phys.* **2019**, *52*, 145202. doi:10.1088/1361-6455/ab215a.
138. Modak, P.; Antony, B. Electron scattering from FO. *J. Phys. B: At. Mol. Opt. Phys.* **2019**, *52*, 095202. doi:10.1088/1361-6455/ab13d2F.

139. Nahar, S.; Antony, B. Review: Positron scattering from atoms and molecules. *Atoms* **2020**, *8*, 29. doi:10.3390/atoms8020029.
140. Sinha, N.; Sahoo, A.; Antony, B. Positron scattering from pyridine and pyrimidine. *J. Phys. Chem. A* **2020**, *124*, 5147–5156. doi:10.1021/acs.jpca.0c02575.
141. Modak, P.; Antony, B. Photoionization of CO using R-matrix theory. *Astrophys. J.* **2019**, *887*, 262. doi:10.3847/1538-4357/ab5583.
142. Modak, P.; Antony, B. Probing photon interaction with H2O and D2O. *J. Phys. B: At. Mol. Opt. Phys.* **2020**, *53*, 045202. doi:10.1088/1361-6455/ab5904.
143. Kwon, D.H.; Savin, D.W. Electron-impact ionization of P-like ions forming Si-like ions. *ApJ* **2014**, *784:13*, 1–7. doi:10.1088/0004-637X/784/1/13.
144. Kwon, D.H.; Savin, D.W. Effects of configuration interaction for dielectronic recombination of Na-like ions forming Mg-like ions. *ApJ* **2011**, *734:2*, 1–7. doi:10.1088/0004-637X/734/1/2.
145. Kim, D.S.; Kwon, D.H. Theoretical photoionization spectra for Mg-isoelectronic Cl^{5+} and Ar^{6+} ions. *J. Phys. B: At. Mol. Opt. Phys.* **2015**, *48:105004*, 1–11. doi:10.1088/0953-4075/48/10/105004.
146. Zhang, D.H.; Kwon, D.H. Theoretical electron-impact ionization of W^{17+} forming W^{18+}. *J. Phys. B: At. Mol. Opt. Phys.* **2014**, *47:075202*, 1–6. doi:10.1088/0953-4075/47/7/075202.
147. Kwon, D.H. Dielectronic recombination of lowly charged tungsten ions W^{q+} ($q = 5 - 10$). *J. Quant. Spectrosc. Radiat. Transfer* **2018**, *208*, 64–70. doi:10.1016/j.jqsrt.2018.01.005.
148. Chai, K.B.; Kwon, D.H. Optical emission spectroscopy and collisional-radiative modeling for low temperature Ar plasmas. *J. Quant. Spectrosc. Radiat. Transfer* **2019**, *227*, 136–144. doi:10.1016/j.jqsrt.2019.02.015.
149. Ralser, S.; Postler, J.; Harnisch, M.; Ellis, A.M.; Scheier, P. Extracting cluster distributions from mass spectra: IsotopeFit. *International journal of mass spectrometry* **2015**, *379*, 194–199. doi:10.1016/j.ijms.2015.01.004.
150. Scheier, P.; Märk, T. Doubly charged argon clusters and their critical size. *The Journal of chemical physics* **1987**, *86*, 3056–3057. doi:10.1063/1.452013.
151. Schöbel, H.; Bartl, P.; Leidlmair, C.; Denifl, S.; Echt, O.; Märk, T.; Scheier, P. High-resolution mass spectrometric study of pure helium droplets, and droplets doped with krypton. *The European Physical Journal D* **2011**, *63*, 209–214. doi:10.1140/epjd/e2011-10619-1.
152. Regandell, S.; Marquart, T.; Piskunov, N. Inside a VAMDC data node - putting standards into practical software. *Physica Scripta* **2018**, *93*, 035001, [arXiv:astro-ph.IM/1803.09217]. doi:10.1088/1402-4896/aaa268.
153. Zwölf, C.M.; Moreau, N.; Dubernet, M.L. New model for datasets citation and extraction reproducibility in VAMDC. *Journal of Molecular Spectroscopy* **2016**, *327*, 122–137, [arXiv:cs.DL/1606.00405]. doi:10.1016/j.jms.2016.04.009.
154. Walton, N.; Astrogrid Consortium. Meeting the User Science Challenge for a Virtual Universe. Toward an International Virtual Observatory; Quinn, P.J.; Górski, K.M., Eds., 2004, p. 187. doi:10.1007/10857598_29.
155. Moreau, N.; Zwolf, C.M.; Ba, Y.A.; Richard, C.; Boudon, V.; Dubernet, M.L. The VAMDC Portal as a Major Enabler of Atomic and Molecular Data Citation. *Galaxies* **2018**, *6*, 105. doi:10.3390/galaxies6040105.
156. Taylor, M.B.; Boch, T.; Taylor, J. SAMP, the Simple Application Messaging Protocol: Letting applications talk to each other. *Astronomy and Computing* **2015**, *11*, 81–90, [arXiv:astro-ph.IM/1501.01139]. doi:10.1016/j.ascom.2014.12.007.
157. Heller, S.; McNaught, A.; Pletnev, I.; Stein, S.; Tchekhovskoi, D. InChI, the IUPAC International Chemical Identifier. *J Cheminform.* **2015**, pp. 7–23. doi:10.1186/s13321-015-0068-4.
158. Zwölf, C.; Moreau, N.; Yaya-Awa, B.; Dubernet, M.L. Implementing in the VAMDC the New Paradigms for Data Citation from the Research Data Alliance. *Data Science Journal* **2018**, *1*, 4. doi:10.5334/dsj-2019-004.
159. Rey, M.; Nikitin, A.V.; Babikov, Y.L.; Tyuterev, V.G. TheoReTS – An information system for theoretical spectra based on va riational predictions from molecular potential energy and dipole moment surfaces. *J. Mol. Spectrosc.* **2016**, *327*, 138–158. doi:10.1016/j.jms.2016.04.006.
160. Huang, X.; Schwenke, D.W.; Lee, T.J. Quantitative validation of Ames IR intensity and new line lists for $^{32/33/34}S^{16}O_2$, $^{32}S^{18}O_2$ and $^{16}O^{32}S^{18}O$. *J. Quant. Spectrosc. Radiat. Transfer* **2019**, *225*, 327 – 336. doi:10.1016/j.jqsrt.2018.11.039.
161. Owens, A.; Yachmenev, A.; Tennyson, J.; Thiel, W.; Yurchenko, S.N. ExoMol Molecular line lists XXIX: The rotation-vibration spectrum of methyl chloride up to 1200 K. *M. N. R. A. S.* **2018**, *479*, 3002–3010. doi:10.1093/mnras/sty1542.

162. Rey, M.; Chizhmakova, I.; Nikitin, A.; Tyuterev, V. Understanding global infrared opacity and hot bands of greenhouse molecules with low vibrational modes from first-principles calculations: The case of CF_4. *Phys. Chem. Chem. Phys.* **2018**, *20*, 21008–21033. doi:10.1039/c8cp03252a.
163. Egorov, O.; Nikitin, A.; Rey, M.; Rodina, A.; Tashkun, S.; Tyuterev, V. Global modeling of NF3 line positions and intensities from far to mid-infrared up to 2200 cm-1. *J. Quant. Spectrosc. Radiat. Transfer* **2019**, *239*, 106668. doi:10.1016/j.jqsrt.2019.106668.
164. Rey, M.; Nikitin, A.V.; Tyuterev, V.G. Accurate Theoretical Methane Line Lists in the Infrared up to 3000 K and Quasi-continuum Abs orption/Emission Modeling for Astrophysical Applications. *ApJ* **2017**, *847*, 105.
165. Yurchenko, S.N.; Amundsen, D.S.; Tennyson, J.; Waldmann, I.P. A hybrid line list for CH_4 and hot methane continuum . *A&A* **2017**, *605*, A95. doi:10.1051/0004-6361/201731026.
166. Wong, A.; Bernath, P.; Rey, M.; Nikitin, A.; Tyuterev, V. Atlas of Experimental and Theoretical High-temperature Methane Cross Sections from T = 295 to 1000 K in the Near-infrared. *Astrophys. J. Suppl.* **2019**, *240*. doi:10.3847/1538-4365/aaed39.
167. Erard, S.; Cecconi, B.; Le Sidaner, P.; Berthier, J.; Henry, F.; Molinaro, M.; Giardino, M.; Bourrel, N.; André, N.; Gangloff, M.; et al.. The EPN-TAP protocol for the Planetary Science Virtual Observatory. *Astronomy and Computing* **2014**, *7-8*, 52–61. doi:10.1016/j.ascom.2014.07.008.
168. van der Tak, F.F.S.; Lique, F.; Faure, A.; Black, J.H.; van Dishoeck, E.F. The Leiden Atomic and Molecular Database (LAMDA): Current Status, Recent Updates, and Future Plans. *Atoms* **2020**, *8*, 15. doi:10.3390/atoms8020015.
169. Endres, C.P.; Martin-Drumel, M.A.; Zingsheim, O.; Bonah, L.; Pirali, O.; Zhang, T.; Sánchez-Monge, Á.; Möller, T.; Wehres, N.; Schilke, P.; McCarthy, M.C.; Schlemmer, S.; Caselli, P.; Thorwirth, S. SOLEIL and ALMA views on prototypical organic nitriles: C_2H_5CN. *J. Mol. Spectrosc.* **2020**, *submitted*.
170. Möller, T.; Endres, C.; Schilke, P. eXtended CASA Line Analysis Software Suite (XCLASS). *A&A* **2017**, *598*, A7. doi:10.1051/0004-6361/201527203.
171. Möller, T.; Bernst, I.; Panoglou, D.; Muders, D.; Ossenkopf, V.; Röllig, M.; Schilke, P. Modeling and Analysis Generic Interface for eXternal numerical codes (MAGIX). *A&A* **2013**, *549*, A21. doi:10.1051/0004-6361/201220063.
172. Moscadelli, L.; Rivilla, V.M.; Cesaroni, R.; Beltrán, M.T.; Sánchez-Monge, Á.; Schilke, P.; Mottram, J.C.; Ahmadi, A.; Allen, V.; Beuther, H.; Csengeri, T.; Etoka, S.; Galli, D.; Goddi, C.; Johnston, K.G.; Klaassen, P.D.; Kuiper, R.; Kumar, M.S.N.; Maud, L.T.; Möller, T.; Peters, T.; Van der Tak, F.; Vig, S. The feedback of an HC HII region on its parental molecular core. The case of core A1 in the star-forming region G24.78+0.08. *A&A* **2018**, *616*, A66. doi:10.1051/0004-6361/201832680.
173. Dimitrijević, M.S.; Sahal-Bréchot, S. On the Application of Stark Broadening Data Determined with a Semiclassical Perturbation Approach. *Atoms* **2014**, *2*, 357–377.
174. Dreizler, S. HST Spectroscopy of PG 1159 Stars: the Boundary of the GW VIR Instability Strip. *Baltic Astronomy* **1998**, *7*, 71–82.
175. Dojčinović, I.P.; Trklja, N.; Tapalaga, I.; Purić, J. Investigation of Stark line broadening within spectral series of potassium and copper isoelectronic sequences. *M. N. R. A. S.* **2019**, *489*, 2997–3002. doi:10.1093/mnras/stz2367.
176. Lavrentyev, N.; Makogon, M.; Fazliev, A. Comparison of the HITRAN and GEISA Spectral Databases Taking into Account the Restriction on Publication of Spectral Data. *Atmospheric and Oceanic Optics* **2011**, *24*, 436–451. doi:10.1134/S1024856011050113.
177. Wilkinson, M.D.; Dumontier, M.; Aalbersberg, I.J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.W.; da Silva Santos, L.B.; Bourne, P.E.; Bouwman, J.; Brookes, A.J.; Clark, T.; Crosas, M.; Dillo, I.; Dumon, O.; Edmunds, S.; Evelo, C.T.; Finkers, R.; Gonzalez-Beltran, A.; Gray, A.J.G.; Groth, P.; Goble, C.; Grethe, J.S.; Heringa, J.; 't Hoen, P.A.C.; Hooft, R.; Kuhn, T.; Kok, R.; Kok, J.; Lusher, S.J.; Martone, M.E.; Mons, A.; Packer, A.L.; Persson, B.; Rocca-Serra, P.; Roos, M.; van Schaik, R.; Sansone, S.A.; Schultes, E.; Sengstag, T.; Slater, T.; Strawn, G.; Swertz, M.A.; Thompson, M.; van der Lei, J.; van Mulligen, E.; Velterop, J.; Waagmeester, A.; Wittenburg, P.; Wolstencroft, K.; Zhao, J.; Mons, B. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* **2016**, *3*, 160018. doi:10.1038/sdata.2016.18.
178. Lavrentiev, N.; Rodimova, O.; Fazliev, A. Systematization of published scientific graphics characterizing the water vapor continuum absorption: I. Publications of 1898-1980. 24-th International Symposium on Atmosperic and Oceanic Optics: Atmopheric Physics, 2018. doi:10.1117/12.2504325.

179. Lavrentiev, N.; Rodimova, O.; Fazliev, A. Systematization of graphically plotted published spectral functions of weakly bound water complexes. 22-nd International Symposium on Atmosperic and Oceanic Optics: Atmopheric Physics, 2016. doi:10.1117/12.2249159.
180. Laverick, M.; Lobel, A.; Merle, T.; Royer, P.; Martayan, C.; David, M.; Hensberge, H.; Thienpont, E. The Belgian repository of fundamental atomic data and stellar spectra (BRASS). I. Cross-matching atomic databases of astrophysical interest. *A&A* **2018**, *612*, A60. doi:10.1051/0004-6361/201731933.
181. See http://dataosu.obs-besancon.fr.
182. Atherton, C.J.; Barton, T.; Basney, J.; Broeder, D.; Costa, A.; van Daalen, M.; Dyke, S.; Elbers, W.; Enell, C.F.; Fasanelli, E.M.V.; Fernandes, J.; Florio, L.; Gietz, P.; Groep, D.L.; Junker, M.B.; Kanellopoulos, C.; Kelsey, D.; Kershaw, P.; Knapic, C.; Kollegger, T.; Koranda, S.; Linden, M.; Marinic, F.; Matyska, L.; Nyrönen, T.H.; Paetow, S.; Paglione, L.A.D.; Parlati, S.; Phillips, C.; Prochazka, M.; Rees, N.; Short, H.; Stevanovic, U.; Tartakovsky, M.; Venekamp, G.; Vitez, T.; Wartel, R.; Whalen, C.; White, J.; Zwölf, C.M. Federated Identity Management for Research Collaborations. Standard, Research Data Alliance, 2019. doi:10.5281/zenodo.1307551.
183. Zwölf, C.M.; Rixon, G. Authentication, Authorisation and Accounting strategy. Technical delivrable, VAMDC Consortium, 2015. doi:10.5281/zenodo.3936606.
184. Wittenburg, P.; Hellström, M.; Zwölf, C.M.; Abroshan, H.; Asmi, A.; Di Bernardo, G.; Couvreur, D.; Gaizer, T.; Holub, P.; Hooft, R.; Häggström, I.; Kohler, M.; Koureas, D.; Kuchinke, W.; Milanesi, L.; Padfield, J.; Rosato, A.; Staiger, C.; van Uytvanck, D.; Weigel, T. Persistent identifiers: Consolidated assertions. Standard, Research Data Alliance, 2017. doi:10.5281/zenodo.1116189.
185. Collins, S.; Genova, F.; Harrower, N.; Hodson, S.; Jones, S.; Laaksonen, L.; Mietchen, D.; Petrauskaité, R.; Wittenburg, P. Turning FAIR into reality - Final report and action plan from the European Commission expert group on FAIR data. Report, Publications Office of the European Union, 2018. doi:10.2777/1524.
186. Berners-Lee, T.; Hendler, J.; Lassila, O. Toward principles for the design of ontologies used for knowledge sharing. *The Semantic Web* **2001**, *43*, 907–928.
187. Group, W.O.W. *OWL 2 Web Ontology Language Document Overview, W3C Recommendation 27 October 2009*. The World Wide Web Consortium (W3C), https://www.w3.org/, 1 ed., 2009. http://www.w3.org/TR/2009/REC-owl2-overview-20091027/.
188. Privezentsev, A.; Tsarkov, D.; Fazliev, A.; Tennyson, J. Computed Knowledge Base for Description of Information Resources of Water Spectroscopy. 7th International Workshop on OWL: Experiences and Directions (OWLED 2010), 2010.
189. Fazliev, A.; Privezentsev, A.; Tsarkov, D.; Tennyson, J. Ontology-Based Content Trust Support of Expert Information Resources in Quantitative Spectroscopy. In *Knowledge Engineering and the Semantic Web, Communications in Computer and Information Science*; Springer, 2013; Vol. 394, pp. 15–28. doi:10.1007/978-3-642-41360-5_2.
190. Voronina, V.; Privezentsev, A.; Tsarkov, D.; Fazliev, A. An Ontological Description of States and Transitions in Quantitative Spectroscopy. XX-th International Symposium on Atmospheric and Ocean Optics: Atmospheric Physics, 2014. doi:10.1117/12.2075512.
191. Akhlestin, A.; Lavrentiev, N.; Rodimova, O.; Fazliev, A. The continuum absorption: trust assessment of published graphical information. 25-th International Symposium on Atmospheric and Ocean Optics: Atmospheric Physics, 2019. doi:10.1117/12.2541741.