Development of a knowledge management system for the NOMAD instrument onboard the ExoMars TGO spacecraft

Laszlo Hetey, Eddy Neefs and Ian Thomas
Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium

Joe Zender
Department of RSSD, European Space Agency, Nordwijk, The Netherlands

Ann-Carine Vandaele, Sophie Berkenbosch, Bojan Ristic, Sabrina Bonnewijn and Sofie Delanoye
Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium, and

Mark Leese, Jon Mason and Manish Patel
School of Physical Sciences, The Open University, Milton Keynes, UK

Abstract

Purpose – This paper aims to describe the development of a knowledge management system (KMS) for the Nadir and Occultation for Mars Discovery (NOMAD) instrument on board the ESA/Roscosmos 2016 ExoMars Trace Gas Orbiter (TGO) spacecraft. The KMS collects knowledge acquired during the engineering process that involved over 30 project partners. In addition to the documentation and technical data (explicit knowledge), a dedicated effort was made to collect the gained experience (tacit knowledge) that is crucial for the operational phase of the TGO mission and also for future projects. The system is now in service and provides valuable information for the scientists and engineers working with NOMAD.

Design/methodology/approach – The NOMAD KMS was built around six areas: official documentation, technical specifications and test results, lessons learned, management data (proposals, deliverables, progress reports and minutes of meetings), picture files and movie files. Today, the KMS contains 110 GB of data spread over 11,000 documents and more than 13,000 media files. A computer-aided design (CAD) library contains a model of the full instrument as well as exported sub-parts in different formats. A context search engine for both documents and media files was implemented.

Findings – The conceived KMS design is basic, flexible and very robust. It can be adapted to future projects of a similar size.

Practical implications – The paper provides practical guidelines on how to retain the knowledge from a larger aerospace project. The KMS tool presented here works offline, requires no maintenance and conforms to data protection standards.

Originality/value – This paper shows how knowledge management requirements for space missions can be fulfilled. The paper demonstrates how to transform the large collection of project data into a useful tool and how to address usability aspects.

Keywords Knowledge management, User interface, Lessons learned, Instrument development, Knowledge management systems, Space technology

Paper type Research paper

Introduction

In aeronautical technology development, it is considered vitally important to keep acquired knowledge at hand for consultation and also for future applications. Organisations must ensure that they have well-informed personnel to solve problems whenever they occur, and that the problem solver has the necessary information at their disposal during the problem-solving process. To avoid that knowledge disappears with personnel (e.g. through retirement, change of function, employees leaving the company, end of project), it should not only reside in the minds of staff members, but in an archived format. In the aerospace sector, a wide range of very specific knowledge domains (such as materials, mechanical and thermal design, electronics, spectroscopy, etc.) are used. Therefore, knowledge management (KM) systems are of particular importance in this industry branch (Rafiei *et al.*, 2016), both in private companies and in space agencies. A good overview of the

activities is given in the Aerospace Knowledge Management Toolkit (AGP – Aerospace Growth Partnership, 2016), covering processes, standards and metrics.

KM is an ancient technique of humankind, but the actual term came into use in the 1990s following several key publications. Nonaka and Takeuchi (1995) define tacit knowledge as a "non-linguistic, non-numerical form of knowledge that is highly personal and context specific" and explicit knowledge as formal and systematic. This distinction is still in use, sometimes using "implicit" instead of "tacit", thereby meaning knowledge that is difficult to write down or verbalise. Davenport and Prusak (1998) distinguish data, information and knowledge. Data are a set of signs and alphanumeric characters. Information results from analysing and interpreting data. Finally, "Knowledge is a mix of framed experiences, values, contextual information, and expert insight that provides a framework for evaluating and incorporating new experiences and information". Alavi and Leidner (2001) describe KM as "processes of creating, storing/retrieving, transferring and applying knowledge". They define a KM System as a "class of information systems applied to managing organisational knowledge. That is, they are IT-based systems developed to support and enhance the KM".

Dalkir (2017) provides a recent overview of KM, explaining the historical development and ongoing discussions. "The author identified over 100 published definitions of KM, and of these, at least 72 could be considered very good!" Since its inception, there were roughly three generations of KM. The first focused on the gathering of all information in containers of knowledge. The resulting information overload led to the second generation, which focused on people and communities of practice. Now, the third is about how to organise and share the content, with metadata and knowledge taxonomies. In the current debate, KM is often not even mentioned, as its basic principles have become common practice. Exemplary in 2015, the ISO 9001 quality management standard was updated to include KM principles. Still, news items such as "NASA admitted that knowledge how to put a man on the moon has been lost" point to the problem of today. Ever increasing amounts of data lead to constant information overload and difficulties to locate relevant information. "The third stage of KM brought about an awareness of the importance of content – how to describe and organise content so that intended end users are aware it exists, can easily access and apply this content."

Over the past decades, the National Aeronautics and Space Administration (NASA) and the European Space Agency (ESA) have invested in KM strategies to overcome a loss of knowledge due to decentralisation and the retirement of personnel. Knowledge reuse has some tradition at NASA, and the analysis of several approaches to knowledge transfer enabled the identification of key factors, described by Majchrzak et al. (2004). Hoffman et al. (2016) describe the current NASA standard, where each research centre has its own KM officer and provides online resources and tools. The KM situation at the European Space Operations Centre (ESA/ESOC) is described by Mugellesi Dow and Pallaschke (2010). Here, different KM systems are in place, e.g. for instrument development or spacecraft operation. Holm et al. (2006) provide an architectural approach to KM for space agencies

and industries, which was taken as a guideline during this project. The work by Linde (2006) provides an excellent overview of KM during the NASA Mars Exploration Rover mission. It also highlights the problem of preserving the knowledge between missions and even between generations. At the end of each space mission, a lessons-learned process should be carried out. Together with a KM plan, this is a standard requirement of mission management in space agencies (NASA, 2010, 2012).

The international Rosetta space probe was designed and implemented in the 1990s, launched in 2004 and started its operations in 2014, before it ended its mission with a soft landing on the comet 67 P/Churyumov-Gerasimenko in 2016. From this timeline, it is obvious that neither tacit knowledge nor technical documentation would be readily available some 10+ years after the launch. Several activities to ensure the availability of the documentation, as well as the capturing of tacit knowledge using video recordings, are described in Zender et al. (2006). Conclusions from this approach were used to design the Nadir and Occultation for Mars Discovery (NOMAD) knowledge management system (KMS).

This paper describes the KMS development for the NOMAD instrument on board the ESA/Roscosmos ExoMars Trace Gas Orbiter spacecraft. After several years of research and engineering work, there were around 300 gigabytes (GB) of data scattered over many different servers at the Royal Belgian Institute for Space Aeronomy (BIRA-IASB), which houses the instrument's principal investigator (PI) team, and across Europe at the various project's partner institutes. The data were merged, analysed and sorted, reducing the amount to roughly one-third of the initial volume. This effort did require active collaboration with different project members to understand and document their individual logic. But, the collection of lessons learned, the compilation of the CAD library and the classification of media files in particular generated significant new explicit knowledge, based on the experience (tacit knowledge) of the team. Without this dedicated KM effort, the tacit part would have likely been lost in future, or if not lost, would certainly have been much more difficult to exploit.

The NOMAD KMS is now operational at the main project partners' institutes and provides valuable information from the instrument development phase that can be used to interpret the measurements and understand the instrument in more detail. Examples of application areas are: the in-orbit calibration, analysing the influence of heating on instrument performance or the operation of the flight spare instrument that is maintained operational in the clean room at BIRA-IASB. In the following sections, we will describe the instrument, the KMS design and implementation, the included data and how the system works in practice.

NOMAD

NOMAD has been designed to investigate the composition of the atmosphere of Mars, with a particular focus on trace gases (especially methane that is an indicator of possible biological activity), clouds and dust. The detection sensitivity for trace gases has been improved considerably compared to previous instruments for measuring the Martian atmosphere, and it is

expected that the results will advance our understanding of the Martian atmosphere and the underlying chemical processes.

After the launch in March 2016, a seven-month journey to Mars, and 1.5 years of aerobraking, the scientific activities of TGO started in March 2018. Functional tests confirmed that the instrument was working as expected, and scientific data began to be downlinked to Earth shortly afterwards. Data analysis and the fine-tuning of the instrument are ongoing. The first scientific results show a surprising non-detection of methane in the atmosphere (Korablev *et al.*, 2019) and show what impact the recent planet-encompassing dust storm had on the transport of water in the atmosphere (Vandaele *et al.*, 2019).

NOMAD project

In 2010, NOMAD was selected to be part of the ESA/NASA ExoMars mission. The collaboration with NASA was cancelled in 2012 when the US federal administration ended its involvement due to budgetary limitations. ESA redirected and redefined the mission into a collaboration with the Russian space agency Roscosmos, which provided the launch vehicle and two science instruments on board the spacecraft. The ExoMars mission consists of two parts: the orbiter (TGO)/lander (Schiaparelli) mission of 2016 and the rover mission planned for 2020.

The NOMAD instrument development was led by a Belgian PI team (BIRA-IASB), assisted by co-PI teams from Spain (IAA/CSIC), the UK (Open University) and Italy (IAPS/INAF). The industrial efforts were coordinated by the Belgian prime contractor (OIP Sensor Systems, Oudernaarde). Altogether, more than 30 partners were involved in the project.

NOMAD instrument

The instrument conducts a spectroscopic survey of Mars' atmosphere in the ultraviolet (UV) (250-650 nm) and infrared (IR) (2.2-4.3 μm) wavelength domain. It comprises three channels: an infrared solar occultation (SO) channel, an infrared limb, nadir and occultation channel (LNO) an ultraviolet-visible channel (UVIS). Technical details are described in Vandaele *et al.* (2018).

During the development of NOMAD, several models of the instrument were built, as required by ESA: an electrical model (ELM) simulating the interfaces with the spacecraft, a structural and thermal model (STM) with a representative mechanical interface and thermal behaviour, a proto-flight model (PFM) now in orbit around Mars and an flight spare (FS) model that is an identical copy of the PFM.

After design, manufacturing and assembly, every channel and model required dedicated testing and documentation. As the development timelines were often overlapping, there was a risk of mixing up parts or test results. This was excluded by carefully labelling all parts and photographing every ingoing and outgoing inspection as well as all functional and environmental test setups. This resulted in a considerable number of media files at the end of the project.

Instrument development and knowledge transfer at BIRA-IASB

Over the past few decades, BIRA-IASB has been involved in several spectrometer developments for important planetary aeronomy missions, all of which contribute to the heritage of the NOMAD instrument:

- SPICAM-S (Mars 96, 1996): a two-channel solar spectrometer for UV (250-650 nm) and IR (1.8-4.8 μm) wavelengths (Neefs, 1995). It was placed on the Mars 96 orbiter, which failed half an hour after launch due to a rocket problem;
- SPICAM-Light (Mars-Express, 2003-today): an imaging spectrometer for the UV (118-320 nm) and IR (1-1.7 μm). The goal was to determine the Martian atmosphere composition and temperature as a function of altitude. The instrument is still active (Montmessin *et al.*, 2017); and
- SOIR (Venus-Express, 2005-2014): a compact high-resolution spectrometer operating in the IR (2.2-4.3 μm) aimed at the analysis of Venus' upper atmosphere (Nevejans *et al.*, 2006). It was operational during eight years in orbit.

As NOMAD builds on the legacy of these instruments, the NOMAD KMS includes important knowledge from these earlier missions, too. In turn, NOMAD and its KMS will serve as a source of knowledge and experience for new instrument developments at BIRA-IASB, such as ALTIUS, a spectral imager for the Earth's atmosphere to be launched in 2023 (Vanhamel *et al.*, 2015) and for VenSpec-H (former VEM-H) on board the EnVision mission to Venus, one of three projects selected by ESA in the framework of its M5 mission to enter phase A (Ghail *et al.*, 2017).

Obviously, for all of NOMAD's predecessors, documentation, test results and datasheets were preserved, but this did not occur in a very structured and transparent manner. Moreover, during these previous developments, it was particularly the long-serving engineers' and scientists' know-how (tacit knowledge) that was critical for success. Their continued involvement as consultants after retirement prevented a catastrophic loss of knowledge. At BIRA-IASB, it is also common practice to build teams with senior (experienced) and junior employees, which enables knowledge transfer. During the NOMAD project, it was understood (notably also by the funding agency) that in the long run, a more structured and permanent KM approach is needed. Future instruments will benefit from this effort as they will be able to build further on an existing KMS.

KMS design and implementation

At the start of the NOMAD KM project, the scope was narrowed down and the high-level requirements were formulated:

- retain the knowledge from the technical development of the instrument;
- deliver a system that contains all documentation in an easy-to-use searchable interface;
- financial details and personal information were to be filtered out; and
- cover the period from writing up the proposal to delivering the end product to ESA.

One of the first steps was to conduct a study of related literature and systems that had proved to be useful in the past. Access to

these programs is not straightforward, as KMSs often contain key technology information and confidential data.

Space related KMS

Lipusz *et al.* (2006) implemented a KMS on the basis of the Hungarian involvement in the Rosetta mission. A client-server architecture (web server with PHP/MySQL interface) was implemented for a collection of around 10 GB of data and 12,000 documents. The available documents were converted to an XML format to enable context search.

Schubert *et al.* (2010) explain the Software Platform for Organising and Capturing Knowledge (SPOCK) system for concurrent engineering during the concept phase of new spacecraft designs. In this system, teams of space engineers and scientists from different disciplines focus simultaneously on a range of design aspects. Their KMS is based on three pillars:

- archiving the decision-making process;
- capturing tacit knowledge, usually provided as PowerPoint presentations; and
- a search interface that allows the incorporation of internet-based databases and encyclopaedias such as Wikipedia or WolframAlpha.

Mugellesi Dow *et al.* (2016) describe the development of the Automatic Transfer Vehicle knowledge CApture Project (ATVCAP) KMS. This tool was designed to capture the knowledge from the five ATV supply ship missions to the International Space Station. Tacit knowledge was gathered with the help of recorded interviews and articles written in a wiki. The web-based user interface with search functionality was based on an operations database and includes a discussion forum.

The collection of lessons learned (LL) is the principal section of the KM tool for capturing tacit knowledge. The NASA (2019) public LL system was particularly useful as it is well-structured and has proven its value over decades. A recent LL system development is described by Ganopol *et al.* (2017). Their Continuous Improvement Method was aimed at improving the performance of the satellite operations team for a mission that lasted over four years. Lessons were initiated via a ticketing system of anomalies and included feedback through direct experiences.

ROsetta Knowledge SYstem (ROKSY) is the KMS from the Rosetta spacecraft, Zender *et al.* (2006). The system was aimed at the collection of all available mission, spacecraft and instrument documentation, including reports, figures, photos, drawings, etc. Different teams within the mission grouped the documents in different ways, e.g. the project management team followed the official agency's review format, while the instrument teams grouped the documents according to the originators' rules – the sub-contractors who had produced the documentation. The Rosetta KM team tried to capture all these different approaches and to translate them into meta-data that were used to tag the individual information items in the database. Another challenge was to tag all the information according to the model philosophy to allow future generations of engineers and scientists to distinguish clearly between information relating to instruments in space as opposed to those kept on Earth in the laboratories. The Rosetta KM team also collected tacit knowledge of instrument teams by

conducting video sessions with the entire team. Typically, a week of video interview preparation, execution and wrap-up was needed for each team. The videos were accompanied by XML-based content files that allowed a content search. Thanks to the involvement of BIRA-IASB in the Rosetta mission, the NOMAD KM team had access to ROKSY and was able to analyse its architecture and performance.

Holm *et al.* (2006) provide an overview of space-related KMS technology. The article is the result of the KM study group of the International Academy of Astronauts (IAA). The four critical success factors were identified as: cultural aspects, knowledge architecture (high accessibility, searchability, ease of use), IT infrastructure and supporting services. The outcome is a high-level guideline for the design of space-related KMSs.

However, since then, there is no common KMS approach, tool or library in place at the different space agencies, research institutions or companies. Most KM initiatives start developing from scratch and tailor solutions for their specific needs (Schubert *et al.*, 2010; Mugellesi Dow *et al.*, 2016). Besides the development of a practical tool for the NOMAD project, the aim was to design a flexible system that can be used for other missions. Documentation needs, instrument design data and project management aspects will be similar across agencies.

System requirements and implementation environment

Our high-level requirements had to be concretised and adapted to the available IT environment and resources. There were four main aspects for the design.

Which data to include? – The NOMAD project with more than 30 partner organisations, regular engineering meetings, a multitude of engineering interactions and hundreds of tests, had produced a huge amount of data (around 300 GB). Some of it was not relevant, with large parts being just duplicates or outdated document versions. The following information was defined as key knowledge and included in the KMS:

- all deliverable documents;
- all official documents from the project partners (i.e. carrying an official ESA ExoMars name and version number);
- all photos taken during the instrument development;
- CAD models and electrical part designs;
- technical data and test results; and
- collection of lessons learned.

How to capture tacit knowledge? – Ideally, solutions that solved major problems are documented immediately. This requires discipline and additional effort, which, in many cases, is not feasible due to project deliverable deadlines. Another good strategy to make up for a lack of immediate recording is to reflect when the instrument is finalised and the problems encountered are still fresh in the memory of the designers and manufacturers. An LL collection was compiled with the main development partners during a number of brainstorming sessions.

Who should be able to access the system? – A wide range of options can be adopted: from restricted access for the PI team only, to access for all participants of the NOMAD development, or even global open access. As the system contains confidential data, some access restrictions had to be implemented. As we wanted to accumulate the knowledge from all key project partners, a data sharing functionality was set up,

giving access to the KM contributors. The existing non-disclosure agreements between individual partners were then extended to all partners contributing to the KMS.

Which IT technology should be used for the implementation? – A client-server Web interface with different access levels and user database includes a security risk and requires intensive efforts to categorise the data. It was decided to implement a standalone application that runs locally and offline at the different contributing partners. This avoided online security risks and allowed the KMS to be integrated into the local IT infrastructure of the partners.

To provide a low maintenance solution that does not depend on operating system versions or runtime libraries, the decision was made to develop the NOMAD KMS using HTML and let users choose their preferred Web browser as the interface to their local KMS database. The functional requirements can be summarised as follows:

- easy installation and distribution of the system;
- search interface for documents and media files;
- high level of usability and adequate program documentation;
- option to extend the system with moderate effort; and
- low maintenance cost.

Following the definition of Alavi and Leidner (2001), a KMS is an IT-based system to support the KM. In our case, the KMS offers many functionalities of document and content management systems, and the CAD library has some features of a knowledge-based engineering tool. But, there is no version control and no interface for data entry, no wiki or discussion forum. The idea is not to publish or distribute the data to a larger audience. It is also not an anomaly tracking system nor a concurrent engineering tool. The aim of the NOMAD KMS was to collect the knowledge from the development of the instrument and transfer it to the operational phase and to any follow-up project. From the surveyed KMSs, the relevant bits were integrated into our design. Important architectural aspects, such as high accessibility, searchability and ease of use (Holm *et al.*, 2006), were addressed.

Program design and applied technology

After reviewing the existing technologies and establishing system requirements, the next step was to sort the data. This included filtering out irrelevant content and restructuring the data. The NOMAD KMS is condensed in six data folders and one additional folder (/HTML) which contains the interface functionality (see Figure 1). The data are well-structured and accessible even without the user interface. Table-of-content files in ASCII-format that explain the folder content were added to the main folders.

The graphical user interface (Figure 1, left) is based on functionality located in the HTML folder. The KM data are sorted into six folders (right). The context search is based on Java Server Pages (JSP), the Java Runtime Environment (JRE) and Regain. The Lucene index file was created from four document folders. The picture/movie search was implemented with JavaScript and browses the pictures and movies folders.

The file index.html, located in the root folder, provides entry to the start page. The interface navigation is based on static HTML links, using the current HTML5 standard. There are

Figure 1 NOMAD KMS design
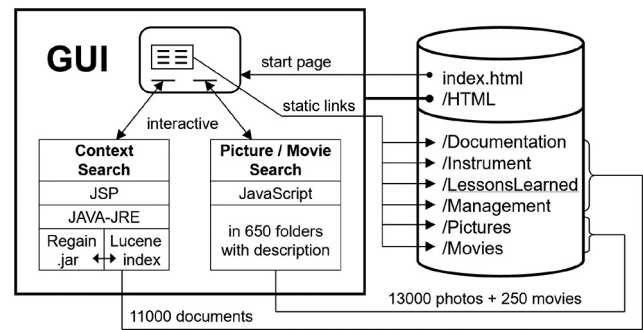
around 11,000 documents, 13,000 photos and 250 movies, which can all be accessed using the folder structure provided.

A search engine for pictures (photos) and movies was implemented with JavaScript, while for documents, a context search was implemented. This is done via JSP, the execution of Java code (Regain.jar) and a Lucene index file. To run this solution, users need to have the Java Runtime Environment (2019) installed, which is typically the case.

Apache Lucene (2019) is a free and open-source indexing and search library, renowned for its usefulness in the implementation of internet search engines. Large amounts of documents can be processed, and Lucene generates an index. Regain (2014), another open source project, provides a search engine to access the index file. The search results are generated very quickly and well-formatted. In summary, Lucene and Regain provide a convenient solution for local site searches that was perfectly suited to the project needs.

Photos and movies were sorted chronologically into approximately 650 folders. Some folders contain a large number of photos, which require considerable time to load and display in full resolution in the browser. To speed this up, thumbnails of photos and movies (screenshots of the initial frame) were generated using Image Magick (2019). The compilation of the HTML files that contain links to the thumbnails were automated with UNIX scripts.

The final program design could possibly be improved upon, but it meets the project requirements and is practical and robust. Version control during the KMS development was ensured via the BIRA-IASB standard archiving procedures.

NOMAD KMS – basic system

This section describes the KMS content and basic interface navigation from a user perspective. The system was delivered to the contributing partners on an external hard disk and has a data volume of 110 GB. If a partner signalled that folders contained sensitive data, these were not distributed to the other partners.

Interface description

The user interface can conveniently be accessed via a Web browser. Figure 2 shows the start page when loaded from a local drive. The six data folders (documentation, instrument, etc.) at the centre left and the context search and picture/movie search at the bottom form the basis of the KMS functionality.

Top left is a picture of NOMAD, with all documentation via text and four links to the right of the photo. In the middle section, links are provided to the six main data folders and to some "data highlights", which are important subfolders that facilitate the navigation. Further down is a timeline that shows milestones such as the initial proposal and important reviews (with links to their deliverables). At the very bottom of the page, there are links to the context and picture/movie search engines. In the following, each of the six NOMAD KMS folders is described in more detail.

Documentation

This folder contains all NOMAD documents that follow the official naming convention. They were sorted into subfolders by project partner and grouped by document type, e.g. technical reports (TRE), science notes (SNO), minutes of meeting (MOM), etc.

NOMAD adopted the ESA file naming convention: *EXM-NO-What-Who-Number-issXrevY-DocTitle-Date*, where "EXM" is the project ExoMars, "NO" is the instrument NOMAD, "What" is a three-letter document type (see examples above), "Who" is a three-letter code for the partner, e.g. AER for BIRA-IASB, "Number" is a five-digit sequence number, "issXrevY" contains the issue and revision numbers, "DocTitle" is a short content description, "Date" is the date of the last document change ("YYMMDD"). An example file name is EXM-NO-MOM-AER-00072-iss0rev3-AOTFtelecon-130111.doc.

All ESA deliverables conformed to this naming, including CAD models, figures and ZIP assembly files of test data. Documents in the LL section and a large part of the management folders also use the ESA naming. It should be noted that some partners used their own internal file naming convention for working documents. During the creation of the KMS folders, the focus was on ensuring there was a consistent logic that would allow locating the relevant documents during a search. Draft versions, accompanying test data or photos, often just duplicates from other folders, were reordered, deleted or sorted into subfolders. Otherwise, the documents' name, saving date and content were not altered. Where necessary as well as for the entire KMS, a table-of-content file was added to the folder to explain the particularities. This was carried out together with the document's main author and contributed to the tacit knowledge.

Instrument

The instrument folder collates technical documentation and testing results, often not captured in documents following the official file naming convention:

- Computer-Aided Design (CAD) data;
- input and result files for structural and thermal simulations (e.g. NASTRAN, ESATAN);
- parameter tables used to configure the instrument operation (uploaded to the spacecraft at regular times);
- datasheets with technical details of components;
- documentation for electronic board designs, electrical electronic electromechanical (EEE) components and cabling;
- software and firmware source codes embedded on board the instrument, together with their "git" version control repository, compiled executables and compiler versions; and
- processed engineering test data (vibration, thermal, power, etc.) sorted by system and subsystem.

Figure 2 The KMS start page provides easy navigation via links
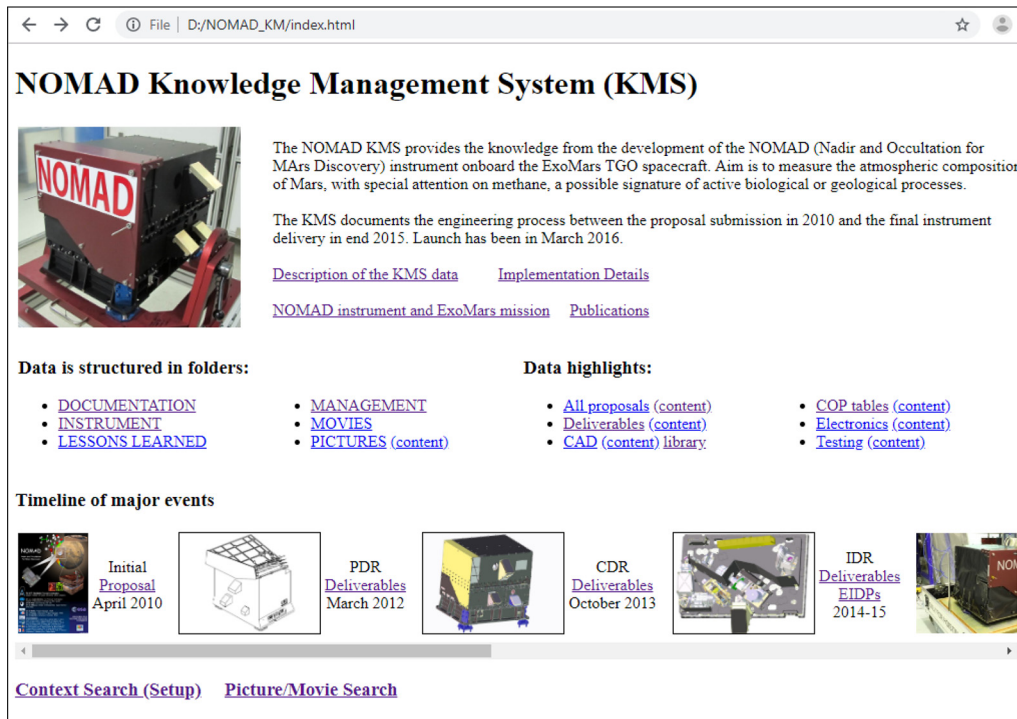
Cleaning this folder after the instrument delivery was a challenge, as it contained the media files, unsorted test data, deliverable documentation from subcontractors, temporary document versions, say all non-official data that were shared in the development team at BIRA-IASB. Here again, several discussions and iterations allowed the involved engineers and scientists to verbalise and document their experience, e.g. via table-of-content files.

Lessons learned

The NOMAD LL effort was guided by a methodology (NASA, 2010) used by the NASA (2019) public LL system. Adapted to the project's needs, a document template containing four mandatory and two optional sections was defined:

- subject/title;
- problem/driving event;
- lessons learned;
- recommendations/good practice;
- references (optional); and
- pictures (optional).

Several LL workshops were called in to discuss and identify with the development partners the weaknesses of existing documentation, problems/solutions encountered and best practices. Some lessons were associated with concrete problems, such as an anomaly due to a technical error. Others were connected to working procedures that at some point became suboptimal, e.g. sharing documents between partners on a common server without clearly defined rules. This latter category involves more tacit aspects and the main additional knowledge often stems from the identification of the cause of the problem. The lessons should be formulated in a way that does not attribute blame to specific persons and contains ideas for future improvements.

Each of the lessons provided was assigned to one of the following types: management, mechanics, optics, electronics, verification, software, assembly. After the initial compilation, lessons were reviewed by the involved specialists and also edited by a member of a different team to have the descriptions in a more generally understandable format. Relevant emails were searched in the archive, and sensible content was deleted, where required. A final document for every LL was saved together with supporting documents (related reports, presentations, emails, and photos). Today, the NOMAD KMS contains 90 LL from 14 contributors. To have an easily searchable and distributable format, a summary PDF file was compiled.

Management

Documents containing management information are often quite diverse. The folder contains the following types of documents (financial details were not considered):

- the initial proposal and preparatory inputs;
- statements of work and proposals (for Phases A/B/C/D);
- deliverables, such as preliminary design review (PDR), critical design review (CDR), end item data packages (EIDP);
- presentations and MOMs (technical and scientific); and
- progress reports (monthly and quarterly).

The sequential collection of proposals and deliverables is potentially useful for future projects. Adding MOMs and progress reports is extremely valuable, as these documents reveal hidden knowledge via the content search functionality of the KMS interface.

Pictures

From an initial number of about 30,000 pictures, 13,000 were retained and sorted chronologically into 650 folders. Duplicates, screenshots of results, blurred images, repetitions and photos mainly containing people were discarded. The folders follow the naming convention: *Date-Location-Model-ShortDescription*, where "Date" is in YYYY-MM-DD format, "Location" is the place where the photos were taken, "Model" can be ELM, STM, PFM, FS (see Section NOMAD) and "ShortDescription" is a short description of the content (example: 2015-09-11_CSL_FS_TV_tests).

The photo collection is one of the KMS's key components, as it documents many project stages, such as the engineering progress, the assembly, the testing activities, etc. Media files are assigned a name by the camera used, usually containing a serial number (e.g. IMG_0001.jpg, IMG_0002.jpg, etc.). They are kept in the KMS. Only important photos received a dedicated title. The saving date is another important information, but it is often overwritten when copying files. Depending on the camera, additional information can be found in the file properties. Best practice is to take pictures of an activity, save them in a dedicated folder with relevant naming and add a short description in textual format, e.g. in an Excel file. There can be several activities the same day or at different locations, and there can be activities that were executed over a longer period. This classification may require minor compromises (such as splitting events into several folders), but the distinction of the different instrument models was an important feature.

After finalisation of the instrument, the media files were spread over different servers. The photo classification was a larger process of knowledge compilation involving a large number of sessions with participating team members. The first stage of finding duplicates (roughly 15 percent) was automated (find files with same name and size), but the majority of the selection work was done manually and required "practical judgement". Other classification strategies, such as naming the photos or adding keywords to file properties, were determined to be unfeasible for the amount of files. Suitable pattern recognition tools may be developed in the future, but the distinction between similar parts from different instrument model versions remains a complex problem. The derived classification model seems most suitable for this kind of project.

Movies

The movie collection uses the same structure and naming convention as the pictures folder. While the folder required the largest amount of space (32 GB), it contains the lowest number of files (250 movies).

NOMAD KMS – advanced functionality

Up to this point, the KMS has become a well-organised data collection. This in itself is of great value as duplicates were identified and for every file only the latest version was retained.

Search functionality

Finding documents directly in the directory tree requires knowledge of the project history, as information can be scattered over different locations. Similarly, although photos were sorted chronologically, which would allow persons that were present at a specific event to find them, it is impossible for people unfamiliar with an event/activity to locate relevant media files. Therefore, a search functionality was incorporated in the KMS.

Regain (2014) is a search engine on top of Apache Lucene (2019). It allows to find specific documents by their name, but also facilitates a text search within the document. Regain can be configured for Web servers or, like in our case, as a desktop search, which provides a search in local documents or intranets. The work consists of two parts: the creation of the index file and the actual search. Large amounts of documents and different file formats (Microsoft Word/Excel/PowerPoint, PDF, ZIP, etc.) can be analysed, and Lucene generates an index file based on the occurrence of words within documents. The index was created for the folders "Documentation", "Instrument", "Lessons Learned" and "Management". Files that mainly contain raw data (*.raw, *.step, *.bdf, *.f06, etc.) were not included. From over 10,000 documents or 10 GB of data, an index file size of 825 MB was generated.

The search index is built to enable a very fast search. The time required for a full text search is moved from the actual search to the index creation. Search queries can be formulated using regular expressions. Depending on the query complexity, the result is generated in fractions of a second (see Figure 3). The search results are rated according to the relative frequency of the search terms in the document, relative to the number of times the term appears in all the documents in the collection

(Apache Lucene, 2019). Another advantage of Regain is that it can be combined with static HTML. Search requests and result visualisation are handled via browser-based JSP.

The context search can be launched from the start page (Figure 2). It performs an offline search using the functionality of a Web search engine. Figure 3 shows the search results for the query "2014 pfm lno oip", i.e. "all documents containing (the number/year) 2014, discussing the PFM model of the LNO channel and involving the Prime Contractor OIP". The underlined filenames in blue are links to the files and the bottom lines in green show the file location. Matches are sorted by relevance and strings matching the search terms are highlighted in bold in the text.

The picture/movie search works in a similar way but uses a different logic. Files usually have a random name (e.g. IMG_5784.jpg) and the content cannot (easily) be derived from the structure depicted in the image. The search is performed on the folder name and on an associated description of its content. Figure 4 (left pane) shows the outcome of a search for "OIP Pfm coalignment", i.e. "all coalignment pictures of the PFM model while residing at the company OIP". The results in blue are links to the folders. When an additional description for a folder is found, it is shown as text below the link. Clicking on a link opens the thumbnail preview for all files, resized to a uniform height of 120 pixels, see Figure 4 (right pane). Clicking on a thumbnail opens the real photo/movie.

CAD library

The project partners used different CAD software for the mechanical design, and the assembly was made by incorporating parts using the STandard for the Exchange of

Figure 3 Regain context search; first page search results corresponding to "all documents created in 2014, discussing the PFM model of the LNO channel and involving the Prime Contractor OIP"
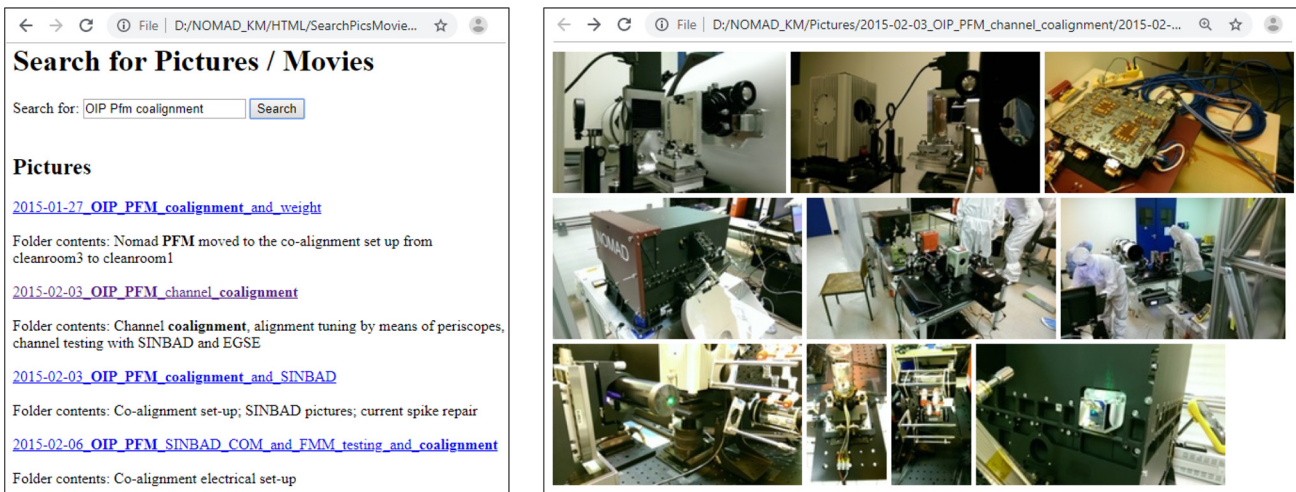
Figure 4 Picture/movie search; selecting the second link from the top in the pane on the left hand-side displays a set of pictures in the pane on the right

Product (STEP) model data format. The final assembly has a file size of over 200 MB, which can be challenging to open and modify, depending on the software used and the hardware environment. In many cases, only small parts are of interest, e.g. for the reuse in a future design (Choi *et al.*, 2007).

Figure 5 shows the top part of the CAD library. All files are available in the STEP format. The first row contains two versions of the entire NOMAD instrument. The middle row shows three important sub-assemblies. The bottom row contains the three instrument channels (LNO, SO and UVIS). Additional rows contain sub-parts of the channels such as detectors, gratings, periscopes and mirrors, with file sizes of less than 10 MB each. To enhance usability, these parts are also provided in the stereolithography (STL) format that can be displayed with general visualisation software.

An explanation of the different part naming conventions used by the partners and a description of encountered problems were documented with table-of-content files for the related folders. This allows to a better understanding of the final design, including its evolution. An archive of previous versions, official part and material lists, the annotated two-dimensional (2D) drawings and a collection of free CAD programs are available. Part of the KMS effort was to compile a final version of the CAD assembly and to export the parts from there. The CAD library offers easy access to the final three-dimensional (3D) mechanical and geometrical details that can be imported into simulation software.

Testing and validation

The NOMAD KMS was developed iteratively. The immediate feedback of different reviewers was used to influence the final program structure. A usability test that contained eight specific tasks and four more general questions about user-friendliness was defined at a later stage. Several senior engineers evaluated the program, and the feedback was generally positive. Particularly, the LL and search functionality were considered to be very useful. The NOMAD KMS was also subjected to integration testing with all common browsers and operating systems.

Discussion

The NOMAD KMS has been delivered and is in use at different locations. The following points underline the benefit:

- Regain context search: The functionality is based on a pre-processed index file that proved to be very efficient and well-suited to the actual amount of data. Search results are generated instantly.
- Photo/movie collection: The aim is to sort the photos by content with descriptive directory names, which can then be searched. For larger collections, this seems the best practical solution. Again, a consistent effort to take and archive photos from the main partners is required over the entire project duration.
- Lessons Learned: Capture the experience gained from overcoming problems. A substantial part of which is admitting errors and documenting improvements. Looking up relevant emails, suitable photos and peer reviewing the descriptions require great effort. As LL systems often fail in practice, because their usage is conceived as too complicated, we addressed usability. Having only four mandatory sections helped to focus on the most relevant. The final PDF document can be searched easily and is clearly structured. The exercise was a first-time experience for the partners and was perceived to add significant value.
- CAD part library: This is important for verification and validation activities. Engineering and scientific work is often clearly separated, which makes it difficult at a later stage to interpret measurements. The library provides easy access to information about parts and will facilitate the mechanical design of future instruments.
- KMS design: The program is user-friendly, very robust, requires no installation nor maintenance nor internet and has no security issues (compared to a Web solution). Reading the provided documentation (top part in Figure 2) should allow anyone to use the program within an hour.

There are many ideas about how to extend or improve the functionality. It could be helpful to set up a forum where the entire project team could discuss the content (e.g. using a wiki),

Figure 5 The CAD library, with direct links (".STP"/".STL") to parts in the STEP/STL format
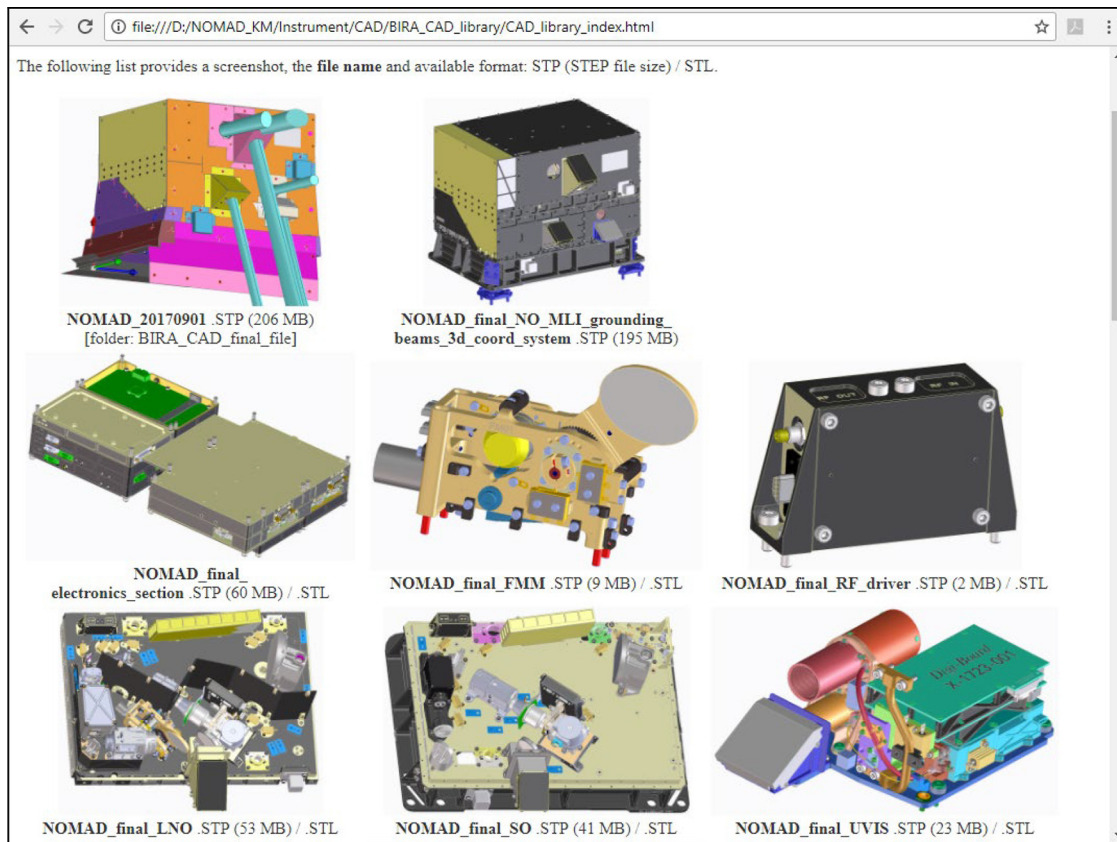
to add or delete files under version control and including synchronisation. Direct access to the Lucene index with the JavaScript interface of the browser or the HTML-embedded visualisation of STEP files are interesting possibilities as well. But, this should be considered carefully, as each added feature increases complexity and the need for maintenance.

Conclusions

A KMS for the NOMAD instrument has been developed. Its successful completion relied on the availability of data, the given IT infrastructure, strong support from management and the willingness of knowledgeable persons to share their experience. The objective was the creation of a system that can easily be adapted for future projects of a similar size. The KMS is currently used to interpret the measurements during the in-orbit calibration, study thermal effects and to operate the flight spare instrument. The context and media search engines proved to be very efficient and well-suited to the actual amount of data. A good compromise has been found between the level of detail and the amount of information provided, functionality, portability and maintainability. The system can be copied onto a laptop or transferred to a different operating system while keeping full functionality. The IT design was kept simple allowing for low development and maintenance costs. Two-thirds of the KMS development time was spent with organising the data and collecting and processing the tacit

knowledge into a format that is useful not only for the ongoing project, but also for future missions.

The KMS now contains data about the NOMAD instrument development; however, an extension for financial management or spacecraft operations would be possible. A practical approach would be to capture the additional tacit knowledge, organise the data in new folders, update the Lucene index and finally adapt the user interface. The scientific operations of NOMAD, foreseen for a duration of at least two years, delivered already interesting measurements. Data and analysis methods from the science phase could also be added in the future to create a complete start-to-finish KMS.

Acknowledgements

References

AGP – Aerospace Growth Partnership (2016), "Aerospace knowledge management toolkit: a practical guide supporting knowledge and skills retention in the aerospace sector", available at: www.adsgroup.org.uk/wp-content/uploads/sites/21/2016/06/AGP-Knowledge-Management-Toolkit.pdf (accessed June 2019).

Alavi, M. and Leidner, D.E. (2001), "Knowledge management and knowledge management systems: conceptual foundations and research issues", *MIS Quarterly*, Vol. 25 No. 1, pp. 107-136.

Apache Lucene, version 7.7.2 (2019), "Information retrieval software library", available at: https://lucene.apache.org (accessed June 2019).

Choi, J.W., Kelly, D. and Raju, J. (2007), "A knowledge-based engineering tool to estimate cost and weight of composite aerospace structures at the conceptual stage of the design process", *Aircraft Engineering and Aerospace Technology*, Vol. 79 No. 5, pp. 459-468.

Dalkir, K. (2017), *Knowledge Management in Theory and Practice*, 3rd ed., The MIT Press, Boston, MA.

Davenport, T.H. and Prusak, L. (1998), *Working Knowledge: How Organizations Manage What They Know*, Harvard Business Press, Boston, MA.

Ganopol, A., Oglietti, M., Ambrosino, A., Patt, F., Scott, A., Hong, L. and Feldman, G. (2017), "Lessons learned: an effective approach to avoid repeating the same old mistakes", *Journal of Aerospace Information Systems*, Vol. 14 No. 9, pp. 483-492.

Ghail, R., Wilson, C., Widemann, T., Bruzzone, L., Dumoulin, C., Helbert, J., Herrick, R., Marcq, E., Mason, P., Rosenblatt, P. and Vandaele, A.C. (2017), "EnVision: understanding why our most earth-like neighbour is so different", *arXiv Preprint arXiv:1703.09010*.

Hoffman, E., Boyle, J. and Rogers, E. (2016), "REAL knowledge and the James Webb Space Telescope: success and failure coexisting in NASA", in Liebowitz, J. (Ed.), *Successes and Failures of Knowledge Management*, Morgan Kaufman, Burlington, MA, pp. 35-58.

Holm, J., Olla, P., Moura, D. and Warhaut, M. (2006), "Creating architectural approaches to knowledge management: an example from the space industry", *Journal of Knowledge Management*, Vol. 10 No. 2, pp. 36-51.

Image Magick, version 7.0.8-14 (2019), "Software to display, convert and edit bitmap images", available at: www.imagemagick.org (accessed June 2019).

Java Runtime Environment, version 8 update 191 (2019), "Set of software tools for development of java applications", available at: https://java.com/en/download (accessed June 2019).

Korablev, O., Vandaele, A.C., Montmessin, F., Fedorova, A.A., Trokhimovskiy, A., Forget, F., Lefèvre, F., Daerden, F., Thomas, I.R., Trompet, L. and Erwin, J.T. (2019), "No detection of methane on mars from early ExoMars trace gas orbiter observations", *Nature*, Vol. 568 No. 7753, pp. 517-520.

Linde, C. (2006), "Learning from the mars rover mission: scientific discovery, learning and memory", *Journal of Knowledge Management*, Vol. 10 No. 2, pp. 90-102.

Lipusz, S.C., Tróznai, G., Bogdány, J. and Szalai, S. (2006), "The Hungarian space research knowledge management project: a focus on the Rosetta mission", *Journal of Knowledge Management*, Vol. 10 No. 2, pp. 76-89.

Majchrzak, A., Cooper, L.P. and Neece, O.E. (2004), "Knowledge reuse for innovation", *Management Science*, Vol. 50 No. 2, pp. 174-188.

Montmessin, F., Korablev, O., Lefèvre, F., Bertaux, J.L., Fedorova, A., Trokhimovskiy, A., Chaufray, J.Y., Lacombe, G., Reberac, A., Maltagliati, L. and Willame, Y. (2017), "SPICAM on mars express: a 10 year in-depth survey of the Martian atmosphere", *Icarus*, Vol. 297, pp. 195-216.

Mugellesi Dow, R. and Pallaschke, S. (2010), "Managing knowledge for spacecraft operations at ESOC", *Journal of Knowledge Management*, Vol. 14 No. 5, pp. 659-677.

Mugellesi Dow, R., Cano Argamasilla, R., Bois, J.M., Steinkopf, M. and Maree, H. (2016), "The challenge of knowledge preservation: the case of the ATV control Centre", *14th International Conference on Space Operations*, *AIAA 2016-2460*, p. 2460.

NASA (2010), "Lessons learned process", *NASA Procedural Requirement 7120.6*, NASA, Washington, DC.

NASA (2012), "NASA space flight program and project management requirements", *NASA Procedural Requirement 7120.5E*, NASA, Washington, DC.

NASA (2019), "NASA public lessons learned system", available at: https://llis.nasa.gov (accessed June 2019).

Neefs, E. (1995), "Martian atmosphere spectrometry from 0.25 to 5 μm in the MARS-96 mission", *Nouvelles de la Science et des Technologies*, Vol. 13, pp. 185-190.

Nevejans, D., Neefs, E., Van Ransbeeck, E., Berkenbosch, S., Clairquin, R., De Vos, L., Moelans, W., Glorieux, S., Baeke, A., Korablev, O. and Vinogradov, I. (2006), "Compact high-resolution spaceborne echelle grating spectrometer with acousto-optical tunable filter based order sorting for the infrared domain from 2.2 to 4.3 μm", *Applied Optics*, Vol. 45 No. 21, pp. 5191-5206.

Nonaka, I. and Takeuchi, H. (1995), *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*, Oxford University Press, Oxford.

Rafiei, A., Akhavan, P. and Hayati, S. (2016), "Knowledge management in successful technology transfer (case study: Iranian aerospace industries and knowledge-based centers)", *Aircraft Engineering and Aerospace Technology*, Vol. 88 No. 1, pp. 178-188.

Regain, version 2.1.0 (2014), "Search engine that can work offline", available at: http://regain.sourceforge.net (accessed June 2019).

Schubert, D., Romberg, O., Kurowski, S., Gurtuna, O., Prévot, A. and Savedra-Criado, G. (2010), "A new knowledge management system for concurrent engineering facilities", *4th International Workshop on System & Concurrent Engineering for Space Applications, SECESA*, pp. 13-15.

Vandaele, A.C., Lopez-Moreno, J.J., Patel, M.R., Bellucci, G., Daerden, F., Ristic, B., Robert, S., Thomas, I.R., Wilquet, V., Allen, M. and Alonso-Rodrigo, G. (2018), "NOMAD, an integrated suite of three spectrometers for the ExoMars trace gas mission: technical description, science objectives and expected performance", *Space Science Reviews*, Vol. 214 No. 5, pp. 80-127.

Vandaele, A.C., Korablev, O., Daerden, F., Aoki, S., Thomas, I.R., Altieri, F., López-Valverde, M., Villanueva, G., Liuzzi, G., Smith, M.D. and Erwin, J.T. (2019), "Martian dust storm impact on atmospheric H_2O and D/H observed by ExoMars trace gas orbiter", *Nature*, Vol. 568 No. 7753, pp. 521-525.

Vanhamel, J., Fussen, D., Dekemper, E., Neefs, E., Van Opstal, B., Pieroux, D., Maes, J., Van Lil, E. and Leroux, P. (2015), "RF-driving of acoustic-optical tunable filters; design, realization and qualification of analog and digital modules for ESA", *Microelectronics Reliability*, Vol. 55 No. 9-10, pp. 2103-2107.

Zender, J., Schwehm, G. and Wilke, M. (2006), "The Rosetta video approach: an overview and lessons learned so far", *Journal of Knowledge Management*, Vol. 10 No. 2, pp. 66-75.

Corresponding author

Laszlo Hetey can be contacted at: laci.hetey@gmx.de