67th International Astronautical Congress (IAC), Guadalajara, Mexico, 26-30 September 2016.

IAC-16,D5,4,6,x32524

Mapping and Optimizing Big Space Data - An International, Interdisciplinary and Intercultural Perspective on the Space Related Data Processes

Daniel Brack[1], Barak Fishbain[2], Sanja Šćepanović [3]

[1] Technion, Israel Institute of Technology, dbrack@tx.technion.ac.il
[2] Technion, Israel Institute of Technology, fishbain@technion.ac.il
[3] Aalto University, sanja.scepanovic@aalto.fi

ISU SSP16 Big Space Data Project Team

Changgui Li, International Space University, China, changgui.li@community.isunet.edu
Deepa K. S., International Space University, India, krishnakumari.deepa@community.isunet.edu
Dimitra Stefoudi, International Space University, Greece, dimitra.stefoudi@community.isunet.edu
Francis Kudjoe, International Space University, Ghana, francis.kudjoe@community.isunet.edu
Guanghong Dong, International Space University, China, guanghong.dong@community.isunet.edu
Hendrik Kolvenbach, International Space University, Germany, hendrik.kolvenbach@community.isunet.edu
Jifeng Ma, International Space University, China, jifeng.ma@community.isunet.edu
Jillianne Pierce, International Space University, United States, jillianne.pierce@community.isunet.edu
Jingjing Peng, International Space University, China, jingjing.peng@community.isunet.edu
Josephine Gull, International Space University, Finland, j.josephine.gull@community.isunet.edu
Julia Heurtish, International Space University, Austria, julia.heuritsch@community.isunet.edu
Kai Zhang, International Space University, China, kai.zhang@community.isunet.edu
Kanika Anand, International Space University, Netherlands, kanika.sehdev@community.isunet.edu
Kartik Kumar, International Space University, Netherlands, kartik.kumar@community.isunet.edu
Laura Keogh, International Space University, Ireland, laura.keogh@community.isunet.edu
Mingfang Wang, International Space University, China, mingfang.wang@community.isunet.edu
Mitchell Scher, International Space University, United States, mitchell.scher@community.isunet.edu
Naga Sekhar Tammana, International Space University, India, nagasekhar.tammana@community.isunet.edu
Nini Berge, International Space University, Norway, nini.berge@community.isunet.edu
Ofra Boim, International Space University, Israel, ofra.boim@community.isunet.edu
Praskovia Milova, International Space University, Russia, praskovia.milova@community.isunet.edu
Robert Jacobson, International Space University, United States, robert.jacobson@community.isunet.edu
Ujjwal Gangele, International Space University, India, ujjwal.gangele@community.isunet.edu
Vincent Lochet, International Space University, France, vincent.lochet@community.isunet.edu
Wei Jiang, International Space University, China, wei.jiang@community.isunet.edu
Xiang Han, International Space University, China, xiang.han@community.isunet.edu
Yang Chen, International Space University, China, yang.chen@community.isunet.edu
Yongping Zou, International Space University, China, yongping.zou@community.isunet.edu
Yong Song, International Space University, China, yong.song@community.isunet.edu
Yue Yin, International Space University, China, yue.yin@community.isunet.edu
Zezhong Sun, International Space University, China, zezhong.sun@community.isunet.edu
Zhaoyong Ni, International Space University, China, zhaoyong.ni@community.isunet.edu
Zhiqi Zhou, International Space University, China, zhiqi.zhou@community.isunet.edu

In recent years the global data volume growth has been 40 percent each year, the number of bytes created daily is 10 to the power of 18 and the number of servers used by big data companies ranges up to millions. Space related research and space based observations are great contributors of big data. For example, satellite telemetry, sensor data, observation logs and manned space mission studies all produce enormous amounts of information that can be examined again and again to understand more about our universe, our solar system, our planet and our body.

The International Space University (ISU) Space Studies Program (SSP) "Big Space Data" Team Project (TP) held in summer of 2016 examined the world of big data through the eyes of the space community; the Team Project mapped the data creators, the data depositories, the data managers and data consumers related to space and space based activities. The TP discussed ways to better the data flow from sensor to processor to general public. The leading question for the TP was "how do we make new conclusions on different matters from all the data available to me", and by it the TP evaluated ways of using and reusing data acquired by space based and terrestrial sensors to learn as much as possible about the universe around us. The TP examined both the technological challenges in data management as well as the legal and ethical considerations of balancing privacy, data security and intellectual property with the socioeconomic benefits of global information sharing and openness for the general good. This paper summarizes the findings and conclusions of the ISU SSP 2016 Big Space Data Team Project.

1. Introduction to Space Big Data

What could you build with three building blocks? Five? What if you got ten new blocks of different shapes and sizes every minute? This is the challenge that the big data community faces as technology has evolved.

The world of space data in 2016 is vastly different than just a few years in the past. The age of the public sector dictating space activities is in the past. We no longer operate in a world where the purpose and use of space technologies is limited to a specific mission or project. The increase in commercialization and private companies entering the space industry has led to a rapid increase in space data output, capabilities, and applications.

Effective big data utilization has changed from a niche field to the driver of new innovations in every industry, from agriculture to software manufacturing in both public and private enterprises. Every person plays a role in creating big data when they use the internet or make a phone call. Some companies such as Facebook and Google have successfully harnessed internet big data for their own purposes. However, space big data comes with its own technological, political, and economic risks and challenges that have prevented its parallel development.

The goal of the space big data (SBD) team project was to research, identify and map the key stakeholders in the space big data value chain, ranging from those who generate the data to those who use it. Through this research we found that each section of this value chain is facing its own challenges and has gaps in technologies, opportunities, and capabilities that have prevented them from moving forward.

Next, by focusing on these pieces of the value chain and their individual challenges, we identified the key common and general barriers to the overall development and effective expansion of the space big data industry. We then provided strategic recommendations for how government and industry can best understand and utilize space big data from a number of perspectives, ranging from legal to economic to engineering.

It is important to recognize that we did not approach this project attempting to find the best use of space big data, or solve the "problem" of space big data. In reality, we sought to understand the space big data ecosystem as a whole; the currently developing associated industries and disruptive technologies, and find opportunities to effectively expand its use. There is no "problem" with space big data. Space big data represents an opportunity that must be understood and analyzed before it can be realized.

2. Background

2.1 Defining Space Big Data

Space big data (SBD) represents just one branch of the newly evolving field of big data analytics. Over the last decade, this new concept has taken the world by storm, most visibly through its impact on generating meaningful insights in the internet industry. It is only more recently and can be expected in the next few years that space data will be effectively utilized. Space big data rests at the intersection of big data and the newly burgeoning space industry. To properly understand what space big data is, it is important to first understand the two fields it stems from. Space data encompasses all data gathered through activities that utilize space assets. These assets will be defined and expanded upon later on.

Big data is often defined as data with the following characteristics often described as the 5 V's: volume, velocity, variety, veracity, and variability (Marr, 2016).

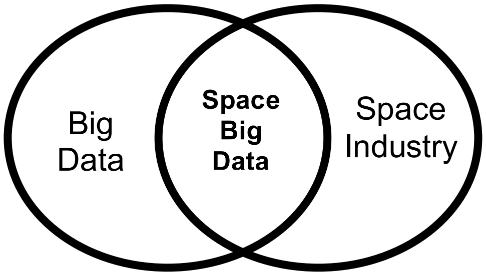

Fig. 1. Space Big Data is at the Intersection of the Space Industry and Big Data

Drawing from these properties, other definitions, and our analysis, our proposed unified definition for space big data is "a large variety (e.g., images, text, logs) and huge quantity of space data that is constantly generated.

2.2 Where Are We Now?

The space sector has long been involved in the big data discussion. Since the beginning of the 20th century, astronomers and engineers developed techniques to compile, handle, and analyze large data sets from observations of thousands of starts to establish relationships between luminosity and temperature. This formed the foundation of our understanding of stellar evolution (Hertzprung, 1908, 1911).

In the 1950s, the development of data management systems such as the World Data Centers (WDC) improved management and archiving capabilities for large data sets. Since then, there have been rapid improvements in storage capacities and technology. Nevertheless, the explosion of internet data in the last decade means that the space industry has an opportunity to take another step forward with new, cutting-edge tools and methodology to effectively handle this leap.

In addition to internet data in the last few years, the improvement of satellite and space-observation technologies has also led to a sharp increase in data collected that is not necessarily met with and equal increase in storage or processing abilities.

Space big data is currently at the threshold of exponential growth that will come in the next few years. This is a similar situation to where internet big data was just a few years ago. Due to the effective management and application of this data by the likes of Google and Facebook, industry is enjoying the results of this boom (Dale, 2016). Space big data has lagged behind and is only now on the cusp of expansion due to a combination of challenges and barriers, both common and unique to the space industry. In order to effectively utilize the opportunity this wealth of knowledge represents, we must now take steps to manage and harness the data being collected and processed and respond to these key barriers. These steps will rely on a complete understanding of not just the data itself, but the source and processes involved in its creation and movement.

2.3 The Value Chain

The value chain, or life cycle analysis, is based on the idea of surveying a series of activities that create and build value that is eventually delivered to customers (Porter,1980). In other words, data collected undergoes a series of transformations from raw to meaningful data (Curry, 2016) The value chain provides a clear breakdown of the core steps that range from generating new data, to understanding it, to utilizing it, and also storing it for future or ancillary use. We identified and investigate the following principal stakeholders:

1. Data Manufacturers
2. Data Processors
3. Data Applications
4. Data Repositories

Many of the challenges facing space big data are common throughout the value chain, but they also may have different resolutions depending on the section of the chain. The ability to address one or more of the core challenges effectively represents a clear competitive advantage. Therefore, understanding what they are is paramount to success in the future use of space big data.

2.4 End Users

On the other hand, one may argue that the value chain itself relies primarily on demand for results, whether public, private, academic, or otherwise. (Armbuster, Kearney, 2016) That is, the results and effectiveness of space big data is measured by our ability to respond to demand, need, or interest. These demands have changed over time, changing from primarily mission based to results supporting government agencies to the more recent trend of commercial application. Who uses the data answers the question of why we are doing the research and development at all.

2.4.1 Government End Users

Government end users focus on public service and assist in decision-making processes to support the economy, disaster management, and other fields. For time sensitive projects such as weather mapping for hurricane relief, special challenges are presented as well as a potentially different political framework from what other end users may experience (AIA, 2012) due to the nature of the work and parties impacted. Government end users are able to utilize space big data to yield more accurate prediction and statuses, and therefore save more lives.

2.4.2 Commercial End Users

Over the last decade, there has been a major switch from government to commercial use of space data. Commercial end-users pay for products and services provided by the industry or create profit for the industry itself using this data. These end users are generally blind to the full value chain of space big data, and instead focus on the meaningful results of the data itself and how it can be used to solve their problems.

These commercial end users can be divided into individual and organizational users. Individual users are general users of an application such as navigation or weather forecasting, regardless of whether they pay directly or indirectly for the service. On the other hand, organizational end users use SBD results to provide services to customers for a profit. Due to the more recent switch to commercialization of space and the use of space data, these organizational users will likely dictate the future success of space big data. Orbital Insight leverages space big data and fledgling artificial intelligence to analyze retail traffic, monitor global oil storage, and estimate harvest yields (Orbital Insight, 2016b). They are able to sell these results to other companies who can use it effectively in their business schemes.

2.4.3 Scientists and Institutions

End users who are interested in using space big data to prove or disprove their theories or to make new discoveries do not necessarily rely on practical outcomes for their work. These users often work with raw data from an academic perspective and focus on understanding rather than using the data. The Chinese space research project, the Five-hundred-meter Aperture Spherical Telescope (FAST) will conduct stellar, galactic, and extragalactic radio astronomy (Wall, 2016). Using FAST, scientists will be able to better understand the early universe. Projects or studies that do not have government support due to low-impact and that are not commercially viable are often found in these institutions.

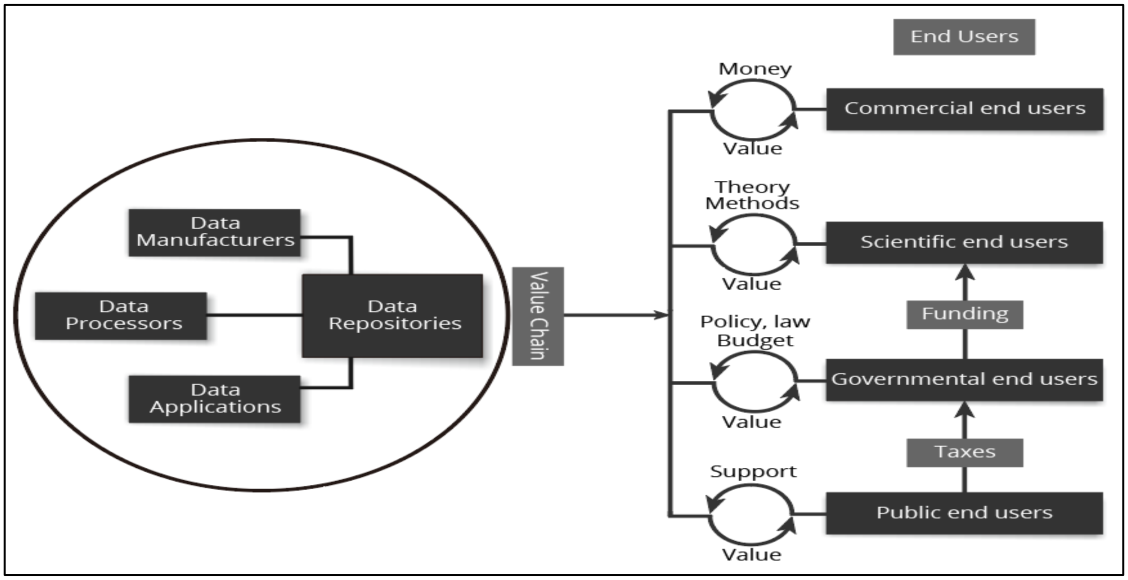

Fig. 2. Value Chain and End Users of Space Big Data

As these demands are in a constant state of evolution, and development based on non-space industry needs changes, the pieces of data and associated processing is also not static. That is, processing of an individual piece data is not a one-time event. As processing technology evolves, new information can be extracted from the same raw data. Additionally, the same piece of raw data can be combined with different data to yield new results such as governments' use of weather mapping for immediate crisis relief and long term weather trending and mapping (AIA, 2012). This added layer of complexity lends itself to the reality that space big data is truly not an issue to be solved, as there will always be more questions and problems to solve. However, understanding the landscape and development of space big data will allow us to effectively utilize its potential.

3. Manufacturers

Data creation is the first step in the SBD lifecycle. Raw data is data that has been collected but not yet analyzed or processed (TechTarget, 2009). This data can include imaging data, non-visual measurements such as temperature, location, or magnitude of any geophysical parameter, or even human data from manned missions.

3.1 Manufacturing Systems

Just as there are many types of data, there are many types of systems that collect the data itself. Space data manufacturing systems can be separated into ground-based and space-based systems.

3.1.1 Ground-Based Systems

Ground-based SBD manufacturing systems include various terrestrial radio and optical telescopes and space tracking radars. These systems are able to generate an immense volume of data about space from Earth. The Square Kilometer Array (SKA), for example (SKA, 2016) will produce about 700TB of data per second at full capacity when it becomes fully operational (JPL, 2016). This volume represents ten times the entire output of the Internet. (Gilliland, 2015).

3.1.2 Space-Based Systems

With the advancement of modern sensors and computing technologies, the ability to generate big data from space assets has increased dramatically (Pixelytics, 2014). The most common payloads that produce big data are sensors aboard Earth observation satellites, space telescopes, planetary rovers and orbiters, and several kinds of in-space experimental instruments - including many on the ISS. This data has traditionally been collected by governments; however, there has been a recent trend toward commercialization in the space industry and, accordingly, toward the growth of commercial manufacturers (Bargmeyer, 2016).

3.2 Data Downloading

For SBD, the downloading process involves the transfer of data from where it was generated to the potential processors, repositories, and/or directly to end-users if no processing is needed first.

Generally, there are two ways of downloading SBD. The first way is by using direct space-to-ground communication stations. For most Earth observation satellites that adopt inclined or sun-synchronized low Earth orbits (LEOs), the orbital period is approximately 90 minutes and the time window to connect to a ground station is typically around 10 minutes. For this reason, data must be stored onboard the satellite and transmitted to the ground station over the course a period of several hours. For satellites in geostationary orbit (GEO), one ground antenna pointing to the specific satellite provides the wider line of sight needed to allow continuous data downloading.

The second way to transmit data to the ground is through space-based data relay. In these systems, the onboard data is transmitted to a relay satellite, and then from the relay satellite to the ground. For example, all the experimental instruments on the International Space Station (ISS) generate huge amounts of data. Several communication upgrades on the ISS are carried out to boost the scientific output. The U.S. Tracking and Data Relay Satellite System (TDRSS) in geostationary orbit provide the ISS almost continuous real-time communications to the ground (Nguyen, Hadjitheodosiou, and Baras, 2004). The Ku band download and upload link on the ISS is now 300 Mbps and 25 Mbps respectively (Cecil et al, 2014).

The latest data relay systems have reached up to an 800 Mbps download link on Ku/Ka bands (Yang, H., 2016). It is very hard to further increase the data rate by radio technology. Laser communication experiments are being carried out as the next breakthrough (Yang, H., 2016).

Data manufacturing is an important process representing the foundation of the SBD supply chain that enables the subsequent levels. Space-based and ground-based data manufacturers have been producing an increasing amount of space data through improved technologies. Nonetheless, there are many open challenges remaining and areas for improvement in both data manufacturing and the connections between data manufacturers and the rest of the data value chain.

3.3 Manufacturing Challenges

3.3.1 Data Transmission Rate, and Time

Data manufacturers carry the burden of setting the tone for the data to be used later in the value chain, but also reconciling that with requirements from the entity controlling the manufacturing mission itself. In addition to the specific task of generating data, the manufacturers must also help provide the link between the data collection and those users and processors that will utilize it. Obstacles include clear insufficiencies in current technologies as well as ancillary legal and business concerns.

One of the most significant challenges that manufacturers face regarding SBD is related to downloading process limitations. For ground-based data manufacturers, this downloading process is achieved mainly through fiber networks (IRA, 2016). Some of the ground manufacturers that don't have access to terrestrial networks use telecommunication satellites to download the data. For space-based manufacturing systems, onboard processing and storage capability are constrained by power and space limits (Van Den Hof, 2003). Data generated in space must be downloaded to the ground for initial or further processing and analysis to provide meaningful results. Once data has been downloaded to the ground, the various repository and dissemination methods will be employed to store and share it.

With the latest advances in sensor technology, the real-time data generation on spacecraft has become faster and faster. For example, data generated from the Charge-Coupled Device (CCD) on some Earth observation satellites can reach up to 1.14 Gbps per CCD chip (Zhang and Shuyan, 2009). The downloading capacity, however, is restricted by spectrum, bandwidth, gain, antenna size, and transmission power (Tozer, 2016). These restrictions will continue to exist as long as there is the need to separate data manufacturing from processing and storage, especially because the data generation rate has increased at a faster rate than the processing and downloading rates. To bear the sharply increasing data transmission burden, the data manufacturers must optimize spectrum use, increase antenna size, and develop more efficient power supplies, while controlling costs or reducing the data volume needing to be downloaded.

3.3.2 Format and Standards

Because data manufacturers generate data to comply with specific requirements set by customers or missions, they may end up generating very different types of data that are not compatible with other data available. Such data may then be useless for customers who plan for large-scale merged analysis. Thus, if the data owner expects to sell its product to a wide range of customers, it should focus on standardization of its data requirements across the value chain, which will impact the manufacturing standards; however, it is imperative that the entire chain adapt simultaneously.

Conclusion

Raw data is manufactured from a variety of different sources, often each with its own unique format. Manufacturers must be able to provide customers with timely, high-quality data to generate valuable products. Manufacturers face a growing need to improve the data downlink from space. More specifically, if SBD is to be useful, it needs to be almost accessible in real time. There are many steps to take to resolve these issues, which will be discussed in later chapters. However, the first step is to understand both the nuances of data manufacturing and the origin of space data as well as how it connects with the rest of the value chain.

4. Processors

Before SBD can be utilized, it must be run through software with algorithms to distill relevant information from bulk datasets. The processed data can then be used in decision-making. Data processing is "the collection and manipulation of items of data to produce meaningful information" (French, 1996). Processing data involves "exploring, transforming, and modeling data with the goal of highlighting relevant data, synthesizing and extracting useful hidden information with high potential from a business point of view. Related areas include data mining, business intelligence, and machine learning" (Curry, 2016). Because of the high demand for data processing of SBD, many stakeholders provide data processing services. Each of these different processing sources has different priorities, and as a result, each has different methods for processing data and distributing findings.

4.1.1 Space Agencies as Processors

Space agencies often manufacture and process SBD from their own assets. One of the primary SBD processing foci for government agencies is mission data, because a mission may collect a varied and voluminous amount of data at a high velocity, relating to both real-time mission analytics and gathered data meant for future use.

NASA's Curiosity Mars Rover is a semi-autonomous robot that gathers and sends raw data to Earth (Taylor, 2012). NASA's Mission Data Processing and Control System (MPCS) interfaces with NASA's network and the Mars Reconnaissance Orbiter (MRO) in orbit around Mars. The MRO then relays data to and from the Curiosity rover and processes the raw data in real time to generate results usable by both Curiosity and flight operations. In short, MPCS can take data points related to weather condition, rover position, and external forces, and make sense of them in real time. Prior to MPCS, processing this data could take days to accomplish, so the enhanced processing ability has greatly improved mission efficiency and general capabilities.

Human space mission data also needs to be processed so that value can be extracted from the SBD bulk dataset. The value hidden within can then be collated in a meaningful form (human readable). Human readable data allows researchers to understand the raw data as well as permits data to be interfaced and read by other systems and languages.

4.1.2 Commercial Processors

The versatility of SBD is such that there is a high demand for customized data analysis that might not be regarded as necessary or worthwhile by or for an agency because of budget, resource, and/or political constraints. In the commercial sector, the main driver of a private company's actions is profitability, and many companies either sell processed versions of already-collected data or take orders for new data and analytics. Other companies provide software for customers to analyze their own data through Application Programming Interfaces (APIs), which may be used depending on what type of data will be analyzed. The wide-ranging impact of this software is that researchers and customers are able to independently process data collected from other sources.

4.1.3 Academia

Students and professors have the advantage of being free from the need for profit and from much of the political turmoil involved in agencies' data processing, so they often use this opportunity to conduct higher risk or lower-profitability studies. Many of the risk studies are geared towards improving existing technology, and lead to the advancement of data processing. The Massachusetts Institute of Technology (MIT) developed the JULIA programming language (Bezanson, 2016) to streamline programming languages. JULIA was meant to be easier to use to analyze data, but it is also suitable for extremely complex datasets. It has been used in areas ranging from spaceflight mapping to bioinformatics, to environmental initiatives.

4.1.4 Crowd Sourced Processing

Crowdsourcing is a practice that harnesses contributions from a large group of people rather than from traditional employees or suppliers to accomplish a goal. The interest for SBD is that crowd sourced processing can bridge the gap between human and machine computations (Anhai, Raghu, and Alon, 2011). The human intellect has absolute advantages in understanding images and signals, which are the main form of SBD. In this way, the crowd is known as the Human Processing Unit (HPU) and works to solve high level challenges together. Zooniverse is an example of an online platform that tasks volunteers from all around the world to help analyze various pieces of information more quickly and accurately than computers. Crowdsourcing has been introduced into database, data mining, social media research, and applications (Lei, Dongwon, and Tova, 2015).

4.2 Challenges for Processors

The overarching challenge of big data processing stems from its definition; the high volume, velocity, variety, and veracity of the data requires processing to extract more value from the data. The most significant challenges facing data manufacturers are identified as standardization, accessibility, interfaces, and data processing methodology. To generate even more value that better reflects the entire scope of the data, better technological tools for processing are needed. Navigating the political arena to define what is and what is not legal and accepted practice has also been a challenge (NASA Science, 2016b).

4.2.1 Standardization and Accessibility

Just as with data manufacturers, there is a standardization problem for data processors. Adopting common approaches to data standardization increases the consistency and credibility of data. Data processors should be able to clearly identify what type and quality of raw data is available to them that will allow them to make effective use and to extract higher-level information.

Accessibility can be divided into two categories, Open Access and Restricted Access, depending also on the processing level. Open Access means that the general public can access data while Restricted Access is defined as a limitation of access on data to a certain group of people such as scientists, governments and/or agencies. Accessibility is also defined by the availability of data itself. Not all data is available for an unlimited amount of time, as the storage capabilities are limited. In the example of the EO data from the EU's Copernicus program, data is stored for one month in a rolling archive. This creates two dilemmas: first is that scientists are unable to use future processing techniques on older data, which has in the past proven to be successful and to yield new discoveries. The second is that since data is only stored for a period, it is unable to be used to generate reliable trends. This can create a problem if a high data output space telescope is operational and the amount of data produced cannot be stored long enough for second order scientists.

4.2.2 Data Processing Method

Data mining algorithms limit computer-based processing (Microsoft, 2016) and determine which data is referenced, how it is applied, and the models that result from it (Visa, 2013). The use of data mining algorithms has resulted in significantly faster processing speeds when compared to human processing or non-algorithm based computer processing. Currently, the technology that enables computer-based processing cannot effectively process and understand images and signals and how they relate to each other, which represents much of SBD . This represents a "semantic gap" between what computers can process and what human experts can extrapolate from the same data (Gançarski, 2014a). Other forms of big data can be processed and referenced through other methods that do not apply to Space Big Data. For example, internet big data is processed and understood by Google, using available technology, because internet big data is largely text-based. Even images searched through Google are referenced using text-based tags associated with the image. The same framework does not exist in SBD, creating a dilemma of still-undeveloped technology.

Another challenge pertaining to processing is the limited amount of influence on the onboard processing. To optimize data usage downstream, satellites use a set of algorithms that analyze the data before transmitting a higher level of the processed data to Earth. For example, clouds automatically mask land areas for Earth observation satellites that operate in the visible spectrum. Better imaging results for the scientists and users processing them are obtained after almost all the clouds have been removed with the help of the set of algorithms onboard the satellites. On the other hand, while it may not be the primary goal of the mission and the satellite, big data revolves around the core principle of using data for different purposes. As a result of the onboard processing tailored to the needs of one primary data user, the raw data that could be useful to other users is not readily available.

4.2.3 Interfaces

Data interfacing is the process by which data from different sources or formats is combined or compared effectively. Current interface systems are not set up to effectively limit the number of standard formats and sharing data is often a difficult task. At the moment there is no universal application program interface (API), which could be a great solution for this issue. Unfortunately, for a universal API to be successful in the current data environment, it would need not only to be a confluence of every existing standard, but also be adaptable to any future changes.

4.3 Concluding Comments

Different groups, companies, and institutions all serve as data processors in different capacities, representing the conflicting and overlapping interests of different bodies in this segment of the supply chain. However, each type of processor is part of the larger picture of how the industry can draw meaningful results from raw, or even previously processed, data. Data processing capabilities and variety lead to possible downstream applications, and dictate, to a certain extent, the scope of the Space Big Data industry.

5. Applications

5.1 What are SBD Applications?

Space data has the power to revolutionize whole industries by providing accurate, reliable, and consistent information to drive decision-making. Diverse fields such as agriculture, transportation, fishing, and retail all stand to benefit from big data. Applications of SBD inform government decision-making and improve commercial use of data. Space applications generate added value from space big data (ESRI, 2016; Buczkowski, 2016) for public and private sectors. Applications that harness space data generally fall into one of the following categories (Tan, 2016):

- remote sensing and GIS
- satellite navigation
- satellite telecommunications
- astronomical observation and space microgravity science

5.1.1 Stakeholders

The key stakeholders in SBD applications are private and public, with the market including governments, private companies, non-profit organizations, and individual end users. Each stakeholder interacts with the other entities based on its background, goals, and resources.

Governments can use SBD to drive decision-making particularly in forestry and land use sectors. Applications developed for weather forecasting are among the most visible examples of SBD use (Sala, 2016). The example of weather satellite to forecast the path of Hurricane Sandy described in Section 2.1 allowed the U.S. government to accurately plan its natural disaster response, protecting millions of lives and saving billions of dollars (AIA, 2012).

Governments also share SBD to encourage economic growth. A common trend is for governments to build open access platforms that present space data to the general public. This provides opportunity for commercial entities to extract value from SBD.

Private companies and non-profit organizations are using space data applications to address social, economic, and humanitarian needs. Growth of these companies stems from a number of recent trends, including increases in computing speeds, advances in distributed computing, decreases in cost of access to space, and miniaturization of electronics. Over recent years, hundreds of millions of U.S. dollars from private sources have been invested in private, for-profit organizations to create new markets and disrupting existing ones.

5.1.2. Market trends

The growth of launch vehicle, satellite manufacturer, and spacecraft operator markets has led to an explosion of downstream activities. The space downstream market refers to the space applications and products provided to the end-users and form an integral part of space economy. These include direct-to-home satellite television services, satellite navigation consumer equipment, and value-added services, as well as small terminal providers for data handling and banking. The UK downstream market is expected to grow to GBP40 billion by 2030 (Technology Strategy Board, 2014).

5.2 Challenges for Applications

As more public and private funds flow into commercial applications, the downstream boom in application markets is expected to quadruple in size over the next three decades (Technology Strategy Board, 2014) with new entrants to the market, referred to as NewSpace, further increasing possible areas of growth. This makes it even more important to identify the challenges facing SBD applications. The next sections address financial, market identification, transparency, legal, and policy considerations.

Conclusion

It is difficult to approach the issue of SBD from a holistic view point as there are many different interests related to data applications including scientific, legal, human performance in space, and engineering. SBD applications have the potential to create new scientific and business opportunities. It is an exciting time for space industry applications and the future has great potential for an industrial boom, but there are barriers and challenges to this expansion. Business climate thrives on efficient and effective policy framework and knowledge of market trends, but investors in space applications are most often constrained by data transfer restrictions and prohibitions from one country to the other. Developers and manufacturers of space applications face scrutiny for producing high quality technologies that could optimize space data use. This poses a threat to space application technology innovation and big space data use.

6. Data Repositories

6.1 What are data repositories?

A data repository is a central location in which data is stored and managed and is essential to the lifecycle of SBD. Manufacturers, processors, and applications all need data repositories because they enable high performance, easy-to-use, manageable SBD to be accessed and used. We need to understand how data repositories function, who owns them, where they are located, how are they accessed, what they provide, and how they are funded.

6.1.1 Ownership and Operation

Governments, organizations, and commercial entities that develop space missions and activities are usually the data owners of the resulting SBD from these activities. Stakeholders who are data owners usually build their own repositories due to concerns about security, data protection, and competitive advantage. In recent years, space data has trended toward an open model. ESA announced that it may diverge from the traditional closed model and share common space data with the public through a private cloud (Red hat, 2016). In China, observing data belongs to the Chinese Academy of Sciences (CAS), and some of the data could be shared using the National Earth System Science Data Sharing Infrastructure (GeoData, 2016). We expect more movement from private repositories to more open data access in the future.

6.1.2 Location

The data repositories of national agencies tend to be physically located within the country's borders, while economic and political factors influence the location of private data archives or archives run by cooperative efforts. For instance, several data centers of China's high resolution Earth observation data (CHEOS) were built in provinces of China, with each center taking responsibility for storing EO data (CHEOS, 2016). The NASA space science data coordinated archive (NASA, 2016b) is located in the U.S., but the cooperative effort SIMBAD astronomical database is located in Strasbourg, France.

6.1.3 Accessibility

Open data is the concept of free access to and use of data. It is similar to open access, but the latter refers to the free use of scientific results, such as papers containing scientific knowledge, rather than data specifically.

Each repository has its own policy related to data accessibility. For instance, NASA must keep all its data open, regardless of whether data repositories want their data to be shared. Some data repositories offer paid access, while others provide free access to the data. Owners may choose to keep their data proprietary, due to national security concerns, or for commercial advantage. However, the accessibility of data repositories within the space industry is unique due to the high volume of open data within the industry.

Most data repositories rely on web-based services to provide access. Government funded entities often require that data be freely distributed. National space agencies dominate the publicly available archives. In the case of NASA, the free distribution requirement is dictated by the U.S. Freedom of Information Act (NASA , 2015). The U.S. has the largest space budget in the world (OECD, 2011), and is able to provide the bulk of easily accessible data in online archives, mainly through agencies such as the NASA, the National Oceanic and Atmospheric Administration (NOAA), and the Geological Survey (USGS). These agencies produce a wide variety of datasets such as astronomical data and climate data. On the European side, European Southern Observatory (ESO) archives astronomical data while EUMETSAT provides climate data similar to that at NOAA (Eumetsat, 2016a).

6.1.4 Data types

SBD can be classified into three types: structured, semi-structured, and unstructured.

Due to the rapid proliferation of SBD, it is very difficult to collect, store, and analyze SBD using commonly available databases or data analysis applications. The differences in data formatting are an additional challenge to the processing, analysis, and distribution of SBD to end-users.

6.1.5 Functions

In addition to storing SBD, repositories also provide metadata. Metadata is structured information about the data in a data archive that describe content, format, location, access privileges, and keywords. Metadata allows datasets to be screened and identified, as well as combined with other datasets when needed (NISO, 2004).

Many data repositories also provide support tools to help the user easily find, process, or understand data. An example is virtual observatories (VO), web-based services for scientific research that combine different software and data archives (IVOA, 2016). For example, EuroVO (Euro-vo.org, 2016) is a European collection of applications that can be used for working with and visualizing data from sources such as the ESO. Another example is ESA Near Earth Objects (NEO) coordination center, which provides tools for astronomers (ESA SSA NEO, 2016). NASA also provides some tools for educational purposes (NASA Science, 2016a). If end users want to use data to produce more specific or specialized results, they may need to conduct their own processing.

6.2 Challenges for Data Repositories

The challenges facing data repositories relate to both the space data and metadata. Data repositories face a sharply increasing volume of data, even beyond their own processing ability. This issue can be divided into five challenges: standardization, accessibility, organization, security, and financing.

Conclusion

Data repositories provide data storage and maintenance, as well as information about the data itself. We conclude SBD repositories are not fundamentally different from any other non-space sector in terms of functionality. However, space data repositories are unique in terms of accessibility and funding.

With the emergence of SBD, challenges for data repositories include standardization, storage organization, accessibility, and financing issues. It is interesting to note that these challenges may likewise not be unique to the space industry; they also appear in other unique terrestrial industries, such as data regarding the high seas.

7. Core Challenges

From researching and understanding the main stakeholders in the SBD value chain, we found that there were key themes regarding the challenges the face. Due to the broad nature of these themes, they also represent challenges associated with the different core disciplines of business, engineering, policy, human performance, and the humanities. Therefore, just as we approached the value chain from an interdisciplinary perspective, so too, must we approach the challenges and solutions.

7.1 SBD Market

The current space industry is marked by high costs and questionable demand as applications of space-data are still in their early stages.

7.1.1 Costs of Market Entry

Because new commercial data manufacturers and repositories have a vested interest in keeping their data private, there is an additional barrier to accessing data to utilize. There is also the need for new or more precise data. As a result, companies without certain data must still look to launching new satellites.

The advent of CubeSats and small satellites has reduced the costs of launching satellites in recent years, but private industry has not yet hit critical mass.[11] (Vecchi and Brennan, 2015). Since the satellite technology itself only accounts for a portion of the total cost, and because the average SpaceX Falcon 9 costs $60M to produce (Grush, 2016), the total costs have only been slightly diminished.

7.2 Engineering Challenges

As a highly complex industry, space big data's stakeholders each face their own engineering and technical challenges, ranging from receiving raw data to finding stored data.

7.2.1 Data Downloading

As mentioned in section 3.2 above, the increasing volume of data produced must find a way to the ground to be utilized. Due to the limited amount of time a satellite is visible and can have its data downloaded, this downloadable volume has a limit. Therefore, even if we are able to produce more important data, there is a limit related to our ability to download it. The ISS downloads its information either directly to the ground or through the Tracking and Data Relay Satellite System (TDRSS). Unfortunately, this type of system is not available for all satellites.

7.2.2 Data Access and Processing

Once ground stations download data, the main challenge is to distribute it among users. The users are scientists, researchers, or the general public. The challenge of a data dissemination system is to provide data to various users and bridge the gap between the providers and the potential users.

Even then, current technology, including state-of-the-art artificial intelligence and machine learning struggles to understand the image data that represents much of space data. Current artificial intelligence and machine learning technologies are effective text-based data processors, which has led to the fast development of internet big data utilization. As a result, data processing and data mining is still limited by humans' ability to analyze it.

7.3 Standardization

Data analysts currently spend 80-90% of their time preparing, locating, reformatting, and filtering data rather than analyzing it (Brown, 2015). Until technology to automatically process data is developed, humans must be able to effectively process it. The lack of standardization despite the efforts of the International Organization of Standardization (ISU) has immensely limited our usage of space big data.

This lack of standardization is also due to a lack of political support for change and adaptation to the increasingly connected network of data. While many governments and organizations agree that standardization is a concern, it is difficult to change practices, and it would not solve the problem of the immense volume of data already created.

7.4 Privacy vs. Openness

Openness and privacy are related challenges in big data. The need to promote liberal data use must be tempered with enabling entities to keep data private. This issue has become more apparent due to the transition from government centric SBD to a balance of publicly and privately generated SBD. The commercial world has a big interest now in SBD for profitability, and may not welcome openness as easily as a government agency.

7.4.1 Openness

Openness is an essential element of big data uses and interface. The easier it is to access information, the more likely it gets to use it and connect it with other sources to produce solutions. Data availability has been a concern for government, scientists, and businesses alike. That said, it also presents a challenge in relation to humanities because different cultures and countries have different values and ethics. Navigating this difficult challenge absolutely requires input from the different disciplines.

7.4.2 Privacy

Privacy has two elements with regard to SBD. First, the obligations on entities to ensure they comply with the appropriate privacy regulations to enable end users and the public's privacy rights. For instance, where there is a need to log into a service, the entity must ensure that the login details comply with privacy law. Likewise, companies must ensure that when they are collecting the space data, such as via remote sensing, they comply with privacy laws. Second is the right of the entity to keep its own data private. For instance, when companies gather or process this data, entities have a right to keep it confidential. The overarching problem, though, is that different countries and regions have different laws and rights to privacy. While space data includes data related to multiple countries, any regulations and laws must abide by each country's laws.

8. Roadmap

The reality is that all of the challenges are interdependent. Without business, technology, and policy progressing, effective use of space big data will continue to be hindered. However, the ways in which each should progress differ.

8.1 Business

Investment in the space industry has risen exponentially (Omohundro, 2015). However, in order to cut costs of business and lower barriers and risks to market entry, government and large industry must further normalize the use of space big data in decision-making. In reducing the risk and creating more opportunities, this will encourage technological innovation and further reductions to costs. By having more non-space industry government programs relying on space data, it will encourage more progressive policies towards data privacy and sharing.

8.2 Technology

Data downloading and data mining concerns discussed in section 4.2.2 above can both be remedied through improved technologies. We must improve data downloading capacity. New technologies such as laser communication offers a promising opportunity for increased download speeds up to 1 Gbps. Increased use and development of these technologies will enable more effective downloading of currently limited data (Ricklin et al, 2006).

Additionally, onboard pre-processing techniques can reduce the sampling rate and filter noises, while various compression algorithms can maximize the download link efficiency. This will require increased investment in promising technologies such as AI and machine learning as much of the data requires image processing that is currently unavailable. This technology must be able to determine what information is worth downloading to reduce the final load.

With appropriate political changes, a global data dissemination system should be established. This would not only aid in the standardization of data and storage, but also create a network where data downloading would no longer rely on a limited time interval where it can download the data. While this global data collection and distribution network would be optimal from an international perspective, the political hurdles of this level of data sharing will likely prove impossible to overcome.

That said, the development and application of existing cloud computing frameworks such as Hadoop would be a good middle ground. The distributed computing, storage, and caching would produce some of the same results as a large network once the data is collected, but would not address the data downloading challenge. Utilizing the cloud also has the added capability of storing big data with effective scalability potential (Ma et al., 2015).

8.3 Openness and Privacy

All data will likely never become open to everyone, nor will all privacy disputes be put to rest. However, the current lack of clear regulations on what may infringe on people's rights to privacy on both a domestic and international level must be addressed. Even as the debate continues, policymakers must begin to form a framework with what decisions have been made. This will alleviate some of the legal fears would-be businesses have upon market entry.

This initial framework should include clear policy relating to later political decisions about privacy. This policy should create a legal framework protecting companies from future decisions that may impact their viability due to added restrictions. This will encourage further business and technological development that would otherwise not occur out of fear.

By starting with a general framework and growing policies based on new technologies and industry needs, policy will be able to grow alongside them.

8.4 Concluding Thoughts

If business, technology, or policy fails to progress, it will immensely impact the overall development of space big data. In the current space industry, technology and policy relies directly on the inputs from private companies. Private companies rely directly on the technology available as well as the policies in place to dictate their viability. Technology development can only stem from the same demand that dictates business's decisions and has largely grown in recent years as the result of research.

The roadmap is not a straight line, but rather an interwoven network of developing aspects of the same space industry.

9. Conclusion

The timing of our space big data project could not be better. After the big data boom of 2014 and subsequent increase in commercial and governmental use of internet big data, space big data is on the verge of a similar development. Project Space Big Data's aim is to shed some much needed light on the landscape of space big data, provide insight into the key challenges for further growth and lay the foundation for a roadmap to overcome barriers (technical, legal, etc.) in the space big data industry.

As expected, there are both similarities and differences in the development and current state of space big data when compared to internet big data. As a whole, the space big data industry is still undeveloped, relying primarily on technologies and marketability from ancillary industries. That said, there is a clear trend towards increased reliance on data from space in general, and by extension the use of big data. After thorough research and analysis, the TP Big Data Team identified clear challenges and barriers to a paradigm shift that many experts say is soon to occur. These will eventually lead to the normalization of space big data in both government and commercial settings.

The space big data ecosystem is truly interconnected. As with much of the space sector at large, the recent shift has been from a government-centric industry towards a more commercial approach to its development and utilization. End users, customers, will become the starting block and motivation for future developments. These commercial actions both rely on, and propagate technological development, and rely on political and humanities-based discourse as perceptions of space big data and privacy change.

One of the main takeaways from our research is that space big data has the potential to completely change not only the space industry, but also industry as a whole if developed and utilized properly. To do so, there must be not only a realization that space big data is a highly applicable field, but there must be action taken to move away from traditional approaches to data dissemination and usage. Space data at the scale it is now at has created completely novel opportunities, risks, and approaches to business and security. That said, space big data is still largely misunderstood. Before industry and technology can develop, space big data should be brought to the forefront of scientific industry as a primary tool for development.

10. Acknowledgements

The ISU and Team Project Space Big Data wish to express their sincere appreciation to Technion, Israel Institute of Technology for its sponsorship of this project.

Team Project Space Big Data would also like to thank the ISU editing team and faculty, including Laura Rose, David Cohen, Jonathan Faull, and Irina Thaler for their support and contributions to this final report.

1. AIA, 2012. Weather Observations from Space are Critical to the Nation's Health & Safety. [pdf] Available at: http://www.aia-aerospace.org/assets/AIA_FY14_Weather_Observations_Fact_Sheets.pdf [Accessed 22 August 2016].
2. Anhai, D., Raghu,R., Alon, H., 2011. Crowdsourcing Systems on the World-Wide-Web. Communications of the ACM, [e-journal] 54(4) pp.86-96. Available through: ACM.org http://cacm.acm.org/magazines/2011/4/106563-crowdsourcing-systems-on-the-world-wide-web/fulltext [Accessed 24 August 2016].
3. Armbruster, P., and Kearney, M., 2016. *CCDS Overview*. International Space University, unpublished.
4. Bargmeyer, B.E. and Gillman,D.W., n.d. Metadata Standards And Metadata Registries: An Overview.[pdf] Bureau of Labor Statistics. Available at: http://www.bls.gov/ore/pdf/st000010.pdf [Accessed 23 August 2016].
5. Bezanson, J., Edelman, A., Karpinski, S., and Shah, V.B., 2015. Julia: A fresh approach to numerical computing.Cornell University. [online] Available at: http://arxiv.org/abs/1411.1607 [Accessed 24 August 2016].
6. Brown, M.S., 2015. What Big Data Analytics Professionals Want From IT. [online] Available at: http://www.forbes.com/sites/metabrown/2015/06/27/what-big-data-analytics-professionals-want-from-it/#25cbf6bf3c78 [Accessed 24 August 2016].
7. Buczkowski, A., 2016. List of the top 100 geospatial start-ups and companies in the world. [online] Available at: http://geoawesomeness.com/list-top-100-geospatial-start-ups-companies-world/ [Accessed 22 August 2016].
8. Cecil, A., J., Pitt, L., P., Welch, S., J., Bryan, J., D., 2014. Enhanced International Space Station Ku-Band Telemetry Service. [pdf] Available at: http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20140010952.pdf
9. CHEOS, 2016. Homepage. [online] Available at: http://www.cheos.org.cn/n380375/n380413/c6656047/content.html [Accessed 26 August 2016].
10. Curry, E., 2016. The Big Data Value Chain: Definitions, Concepts, and Theoretical Approaches. In: Cavanillas, J.M., Curry, E. and Wahlster, W., 2016. New Horizons for a Data-Driven Economy, Ch.3. Springer Nature. Available through: Springer.com http://link.springer.com/chapter/10.1007/978-3-319-21569-3_3/fulltext.html [Accessed 23 August 2016].
11. Dale N., 2016. Doing Business in a Big Data World, Ch.10: Three ways to build a big data system. [online] Available at: http://searchdatamanagement.techtarget.com/feature/Three-ways-to-build-a-big-data-system [Accessed 23 August 2016].
12. ESA SSA NEO, 2016. Near-Earth Object Population Observation Program. [online] Available at: http://neo.ssa.esa.int/neo-population [Accessed 21 August 2016].
13. ESRI, 2016. Homepage. [online] Available at: http://www.esri.com/ [Accessed 22 August 2016].
14. Eumetsat, 2016a. Distributing Data. [online] Available at: http://www.eumetsat.int/website/home/AboutUs/WhatWeDo/DistributingData/index.html [Accessed 21 August 2016].
15. Euro-vo.org, 2016. Software. [online] Available at: http://www.euro-vo.org/?q=science/software [Accessed 21 August 2016].
16. French, C., 1996. Data Processing and Information Technology,10th ed. Thomson, pp79-80. [online]. Available at: https://books.google.co.il/books/about/Data_Processing_and_Information_Technolo.html?id=zVCdg7Tg6-AC&redir_esc=y [Accessed 23 Aug 2016].
17. Gançarski, P., 2014a. Big Data and Earth Observation : New Challenges in remote sensing images interpretation. University of Strasbourg. [pdf] Available at: http://www.cnrs.fr/mi/IMG/pdf/bigdata_remote_sensing_challenges_-_journee_imagerie_cnrs_-_pierre_gancarski.pdf [Accessed 24 August 2016]
18. GeoData, 2016. Space Science data, GeoData [e- journal]. Available at: GeoData.com http://spacescience.data.ac.cn/ [Accessed 26 August 2016].
19. Gilliland, B., 2015. An Evolving Guide To The (Unfinished) Universe (Op-Ed). [online] Available at: http://www.space.com/28464-the-ever-evolving-story-of-the-universe.html [Accessed 24 August 2016].
20. Grush, L., 2016. Breaking: SpaceX successfully lands its rocket on a floating drone ship for the first

time. [online] Available at: http://www.theverge.com/2016/4/8/11392138/spacex-landing-success-falcon-9-rocket-barge-at-sea [Accessed: 29 August 2016].

21. Hertzsprung, E., 1908. Über die Sterne der Unterabteilungen c und ac nach der Spektralklassifikation von Antonia C. Maury. *Astronomische Nachrichten*, 179(24), pp.373-380.
22. Hertzsprung, E., 1911. Über die Verwendung photographischer effektiver Wellenlaengen zur Bestimmung von Farbenäquivalenten. *Publikationen des Astrophysikalischen Observatoriums zu Potsdam*, 63.
23. IVOA, 2016. Homepage. [online] Available at: http://www.ivoa.net/about/what-is-vo.html [Accessed 21 August 2016].
24. JPL, 2016. Managing the Deluge of Big Data From Space. [online] Available at: http://www.jpl.nasa.gov/news/news.php?release=2013-299 [Accessed 22 August 2016].
25. Lei,C., Dongwon,L., and Tova, M., 2015. Data-driven Crowdsourcing: Management, Mining, and Applications. 2015 IEEE 31st International Conference on Data Engineering. Seoul, South Korea 13-17 April 2015. IEEE [online] Available at: http://ieeexplore.ieee.org/document/7113418/?tp=&arnumber=7113418 [Accessed 25 August 2016].
26. Ma, Y., Wu, H., Wang, L., Huang, B., Ranjan, R., Zomaya, A. and Jie, W., 2015. Remote sensing big data computing: Challenges and opportunities. Future Generation Computer Systems, 51, pp. 47–60.[pdf] Available through: Sciencedirect http://www.sciencedirect.com/science/article/pii/S0167739X14002234 [Accessed 28 August 2016]
27. Marr, B., 2015. Why only one of the 5 Vs of big data really matters. [online]. Available at: http://www.ibmbigdatahub.com/blog/why-only-one-5-vs-big-data-really-matters [Accessed 24 August 2015].
28. med.ira.inaf.it, 2016. Receiver Design & Development. [online] Available at: http://www.med.ira.inaf.it/ricevitori.html [Accessed 28 August 2016]
29. Microsoft, 2016. Data Mining Algorithms (Analysis Services - Data Mining). Microsoft. [online] Available at: https://msdn.microsoft.com/en-us/library/ms175595.aspx [Accessed: 19 August 2016].
30. NASA Science, 2016b. Data Rights & Related Issues - NASA Science. [online] Available at: http://science.nasa.gov/earth-science/earth-science-data/data-information-policy/data-rights-related-issues/ [Accessed 22 August 2016].
31. NASA, 2015. Masters with Masters: What is the Big Deal about Big Data? [online] Available at: http://www.nasa.gov/offices/oce/home/MwM_Sep222015.html [Accessed 24 August 2016].
32. NASA, 2016a. NASA Exoplanet Archive. [online] Available at: http://exoplanetarchive.ipac.caltech.edu/ [Accessed 23 August 2016].
33. NASA, 2016b. NASA Exoplanet Archive: Exoplanet Transit Survey Service. [online] Available at: http://exoplanetarchive.ipac.caltech.edu/applications/ETSS/CoRoT_exo_index.html [Accessed 23 August 2016].
34. Nguyen, A.T., Hadjitheodosiou, M. and Baras, J.S., 2004. Broadband communication support for earth observation from the International Space Station. [pdf] University of Maryland: Computer Networks. Available at: https://www.isr.umd.edu/~baras/publications/journals/2005_Nguyen_Broadband_Communication.html [Accessed 22 August 2016]
35. NISO, 2004. Understanding Metadata. [pdf] NISO Press. Available at: http://www.niso.org/publications/press/UnderstandingMetadata.pdf [Accessed 23 August 2016].
36. OECD, 2011. The Space Economy at a Glance 2011. [online]. OECD Publishing pp.50-53. Available at: http://www.oecd-ilibrary.org/docserver/download/9211051ec006.pdf?expires=1472213133&id=id&accname=guest&checksum=0D15F4C028461900BF6EF5D7B3BA0C69 [Accessed 21 August 2016].
37. Omohundro, S., 2015. AI, Deep Learning, and the future of business. [pdf] Available at: https://omohundro.files.wordpress.com/2015/12/vlab-ai-deep-learning-and-the-future-of-business.pdf [Accessed 25 August 2016].
38. Orbital Insight, 2016b. Solutions. [online] Available at: https://orbitalinsight.com/solutions/ [Accessed 27 August 2016].

39. Porter, M.E.,1980. *Competitive strategy: Techniques for analyzing industries and competitors*. 16th ed. New York, NY: Free Press [u.a.].
40. Ricklin, J., Hammel, S., Eaton F., and Lachinova, S., 2006. Atmospheric channel effects on free-space laser communication. Journal of Optical and Fiber Communications Reports, 3(2). Springer pp.111–158. [pdf] Available at: http://download.springer.com/static/pdf/541/art%253A10.1007%252Fs10297-005-0056y.pdf?originUrl=http%3A%2F%2Flink.springer.com%2Farticle%2F10.1007%2Fs10297-005-0056-y&token2=exp=1472214509~acl=%2Fstatic%2Fpdf%2F541%2Fart%25253A10.1007%25252Fs10297-005-0056-y.pdf%3ForiginUrl%3Dhttp%253A%252F%252Flink.springer.com%252Farticle%252F10.1007%252Fs10297-005-0056-y*~hmac=de7df03ba7435defb757a11de1e8e2b6b1a6b0280599bfc6836e38787080fd9c [Accessed 25 August 2016].
41. Sala, S., 2016. Datafeed using big data to detect and predict natural hazards better and faster: lessons learned with hurricanes, earthquakes, floods. [blog] 10 February. Available at: http://datapopalliance.org/using-big-data-to-detect-and-predict-natural-hazards-better-and-faster-lessons-learned-with-hurricanes-earthquakes-floods/ [Accessed 22 August 2016].
42. SKA, 2016. Homepage. [online] Available at: https://www.skatelescope.org/ [Accessed 22 August 2016].
43. Technology Strategy Board, 2014. Downstream Space Applications. [pdf] Satellite Applications Catapult Ltd. Available at: http://ukseds.org/rispace/docs/Space%20Career%20Talk%20-%20UKSEDS.pdf [Accessed 22 August 2016].
44. TechTarget, 2009. Raw data (source data or atomic data). [online] Available at: http://searchdatamanagement.techtarget.com/definition/raw-data [Accessed 24 August 2016].
45. Tozer, T, 2016. Introduction to Link Budgets. [online]. Available at: http://www.satcom.co.uk/article.asp?article=21 [Accessed 22 August 2016].
46. Van Den Hof, P., Wahlberg, B., Weiland, S., 2004. System identification 2003. [e-book] Elsevier p.700. Available through: Google books https://books.google.co.il/books?id=Aa-cDAAAQBAJ&pg=PA700&lpg=PA700&dq=Onboard+processing+and+storage+capability+are+constrained+by+power+and+space+limits&source=bl&ots=c7senEZCsw&sig=rBI6QJO8RR3YU75OtgfGV737EtY&hl=en&sa=X&ved=0ahUKEwiv182d-dnOAhVBshQKHSxwAQsQ6AEIHDAA#v=onepage&q=Onboard%20processing%20and%20storage%20capability%20are%20constrained%20by%20power%20and%20space%20limits&f=false [Accessed 24 August 2016].
47. Vecchi, A., and Brennan, L., 2015. Innovating the business model: the case of space. adoption of innovation. [online] Springer. Available at: http://link.springer.com/chapter/10.1007/978-3-319-14523-5_10#page-1 [Accessed 24 August 2016].
48. Visa, A., 2013. Knowledge mining and big data. Course catalog 2013-2014 - SGN-43006. Tampere University of Technology, Unpublished. [online] Available at: http://www.tut.fi/wwwoppaat/opas2013-2014/kv/laitokset/Signaalinkasittely/SGN-43006.html [Accessed: 23 August 2016].
49. Wall, M., 2016. China Finishes Building World's Largest Radio Telescope. Space.com, [e-journal]. Available through: http://www.space.com/33357-china-largest-radio-telescope-alien-life.html [Accessed 27 August 2016].
50. Yang, H., 2016. Latest development progress and trends of foreign data relay satellite systems. Telecommunication Engineering, pp.109-116. [e-book] Available through: Doc88.com http://xueshu.baidu.com/s?wd=paperuri%3A%28cf57cee320085c25df0b2aae1b7e5eb8%29&filter=sc_long_sign&tn=SE_xueshusource_2kduw22v&sc_vurl=http%3A%2F%2Fwww.doc88.com%2Fp-1671589356841.html&ie=utf-8&sc_us=16582859970274079117 [Accessed 22 August 2016].
51. Zhang, D and Shuyan, X, 2009. High-speed ccd image data fiber transmission system. Optics and precision engineering. [e-book] pp. 669-675. Available through: Baidu.com http://xueshu.baidu.com/s?wd=paperuri:(dde3e8e2d958e517b99938644051ec33)&filter=sc_long_sign&sc_ks_para=q%3DHigh-speed+CCD+image+data+fiber+transmission+system&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_us=712658938479449505 [Accessed 22 August 2016].