A Glimpse into the State and Future of (Big) Data Analytics in Austria
Results from an Online Survey

Ralf Bierig[1], Allan Hanbury[1], Martina Haas[2], Florina Piroi[1],
Helmut Berger[2], Mihai Lupu[1] and Michael Dittenbach[2]
[1]*Institute of Software Technology and Interactive Systems, Vienna University of Technology, Favoritenstr. 9-11/188-1, A-1040 Vienna, Austria*
[2]*max.recall information systems GmbH, Künstlergasse 11/1, A-1150 Vienna, Austria*

Keywords: Data Analysis, Data Analytics, Big Data, Questionnaire, Survey, Austria.

Abstract: We present results from questionnaire data that were collected from leading data analytics experts in Austria. The online survey addresses very current and pressing questions in the area of (big) data analysis. Our findings provide valuable insights about what top Austrian data scientists think about data analytics, what they consider as important application areas that can benefit from big data and data processing, the challenges of the future and how soon these challenges will become important, and the potential research topics of tomorrow. We visualize results, summarize our findings and suggest a possible roadmap for future decision making.

1 INTRODUCTION

The time has come for continuous and large-scale data analytics as our digital lifes now generate the impressive amount of 200 exabytes of data each year. This is equivalent to the volume of 20 million Libraries of Congress (Dandawate, 2013). In 2012 each Internet minute has witnessed 100,000 tweets, 277,000 Facebook logins, 204 million email exchanges, and more than 2 million search queries fired to satisfy our increasing hunger for information (Temple, 2013).

This trend is accelerated technologically by devices that primarily generate digital data without the need for any intermediary step to first digitize analog data (e.g. digital cameras vs. film photography combined with scanning). Additional information is often automatically attached to the content (e.g. the exchangeable image file format 'Extif') that generates contextual metadata on a very fine-grained level. This means, when exchanging pictures, one also exchanges his or her favorite travel destination, time (zone), specific camera configuration and the light conditions of the place with more to come as devices evolve. Such sensors lead to a flood of machine-generated information that create a much higher spatial and temporal resolution than possible before. This 'Internet of Things' turns previously data-silent devices into autonomous hubs that collect, emit and process data at a scale that make it necessary to have automated information processing and analysis (Dandawate, 2013) to extract more value from data than possible with manual procedures. Today's enterprises are also increasing their data volumes. For example, energy providers now receive energy consumption readings from Smart Meters on a quarter-hour basis instead of once or twice per year. In hospitals it is becoming common to store multidimensional medical imaging instead of flat high-resolution images. Surveillance cameras and satellites are increasing in numbers and generate output with increasingly higher resolution and quality. Therefore, the focus today is on discovery, integration, consolidation, exploitation and analysis of this overwhelming data (Dandawate, 2013). Paramount is the question of how all this (big) data should be analyzed and put to work. Collecting data is not an end but a means for doing something sensible and beneficial for the data owner, the business and the society at large. Technologies to harvest, store and process data efficiently have transformed our society and interesting questions and challenges have emerged of how society should handle these opportunities. While people are generally comfortable with storing large quantities of personal data remotely in a cloud there is also rising concern about data ownership, privacy and the dangers of data being intercepted and potentially misused (Boyd and Crawford, 2012).

In this paper, we present results of a study (Berger

Bierig R., Hanbury A., Haas M., Piroi F., Berger H., Lupu M. and Dittenbach M..
A Glimpse into the State and Future of (Big) Data Analytics in Austria - Results from an Online Survey.
DOI: 10.5220/0004999101780188
In *Proceedings of 3rd International Conference on Data Management Technologies and Applications* (DATA-2014), pages 178-188
ISBN: 978-989-758-035-2

et al., 2014) that was conducted between June 2013 and January 2014 on the topic of (big) data analysis in Austria. Specifically, we present and highlight results obtained from an online survey that involved leading data scientists from Austrian companies and the public sector. The questionnaire was targeted to identify the status quo of (big) data analytics in Austria and the future challenges and developments in this area. We surveyed opinion from 56 experts and asked them about their understanding of data analytics and their projections on future developments and future research.

The paper first discusses related work in the next section before describing the method that was used for creating the questionnaire and for collecting and analyzing the feedback in section 3. Results are presented and discussed in section 4. In section 5 we conclude and summarize our findings and suggest actions that are based on our findings.

2 RELATED WORK

Many countries and regions are currently developing strategies for dealing with Big Data. Prominent examples are the consultation process to create a Public-Private Partnership in Big Data currently underway in Europe [1], work by the National Institute of Standards and Technology (NIST) Big Data Public Working Group[2] as well as other groups (Agrawal et al., 2012) in the USA, and the creation of the Smart Data Innovation Lab[3] in Germany.

The recent and representative McKinsey report (Manyika et al., 2013) estimates the potential global economic value of Big Data analytics between $3.2 trillion to $5.4 trillion every year. This value arises by intersecting open data with commercial data and thus providing more insights for customised products and services and enabling better decision making. The report identified the seven areas of education, transportation, consumer products, electricity, oil and gas, healthcare and consumer finance. We expanded this selection by specifically focusing on the Austrian market and its conditions before prompting participants with a comprehensive selection of application areas as described later in section 4.3.

Many other surveys have been conducted on the topic of big data and big data analytics by consulting companies, but these surveys usually concentrate on large enterprises[4]. A summary of the 2013 surveys is available (Press, 2013). A survey among people defining themselves to be Data Scientists has also been done to better define the role of Data Scientists (Harris et al., 2013). In this paper, a survey that takes the views of mostly academic scientists working in multiple areas related to data analytics is presented, and hence provides an unusual "academic" view of the area.

3 METHOD

Surveys are powerful tools when collecting opinion from the masses. Our main objective was to further specify our understanding of data analytics in Austria and to identify future challenges in this emerging field.

We followed the strategy of active sampling. The identification of Austrian stakeholders in data analytics formed the starting point: We first scanned and reviewed Austrian industry and research institutions based on their activities and research areas. We then identified key people from these institutions and asked about their opinions, attitudes, feedback and participation during a roadmapping process.

Our final contact list comprised 258 experts, all of them senior and visible data scientists, that we contacted twice and invited them to complete our questionnaire. This means our contact list has consensus-quality and represents the current situation and strength of senior data scientists in Austria. The survey was online between the beginning of September 2013 until the middle of October 2013. A total of 105 people followed the link to the survey resulting in a general response rate of 39%. However, several of them turned down the questionnaire or cancelled their efforts after only one or two questions. We took a strict measure and removed those incomplete cases from the list of responses to increase the quality of the data. This reduced the original 105 responses (39%) further down to 56 responses (21.7%).

The general advantages of online surveys, such as truthfulness, increased statistical variation and improved possibilities for data analysis (e.g. (Batinic, 2003; Döring, 2003)), unfortunately suffer from the problems of limited control, a higher demand on participants in terms of time and patience and the poten-

[1] http://europa.eu/rapid/press-release_SPEECH-13-893_en.htm

[2] http://bigdatawg.nist.gov

[3] http://www.sdil.de/de/

[4] Some examples: http://www-935.ibm.com/services/us/gbs/thoughtleadership/ibv-big-data-at-work.html, http://www.sas.com/resources/whitepaper/wp_58466.pdf, the Computing Research Association (CRA) http://www.cra.org/ccc/files/docs/init/bigdatawhitepaper.pdf and SAS http://www.sas.com/resources/whitepaper/wp_55536.pdf

tial that people may be engaged in other, distracting activities that alter the results and increase the dropout rate (e.g. (Birnbaum, 2004)). While our response rate of nearly 40% is normal for online surveys (Batinic, 2003), the high dropout rate in our specific case can be attributed to the complex nature of the subject.

The data of the survey was collected anonymously with LimeSurvey[5] and was analyzed with R, a statistical software package[6].

4 SURVEY RESULTS AND DISCUSSION

This section highlights the results we obtained from the data and focuses on four areas. First, the demographic information about the participants (e.g. their age, area of work, and their work experience) that helps us to get a better understanding of the characteristics of a typical data scientist in Austria. Second, we look at the application areas of data analytics and how participants projected the future relevance of these areas. Third, we investigated the future challenges of data analytics. Forth, we analysed free text submissions from questions about research priorities and the need for immediate funding to get a better understanding of possible future directions, interests and desires. We omitted replies for the questions 9, 13 and 16 that inquired about other application areas and additional general comments (see appendix 5. These questions only had a very limited text response that would not be meaningful to analyze statistically.

4.1 Participants

The data presented in this paper is based on the opinions of 56 people who completed the questionnaire — four female (7.1%) and 52 male (92.9%). This gender distribution is similar in the original contact list — 26 female (10.1%) and 232 male (89.9%) — and therefore represents the current gender situation in the data science profession in Austria. Participants were mostly Austrians (96%) and the majority of them were working in the research and academic sector. About a fifth (21.4%) of all responses came from the industry. The larger part worked for academic (55.4%) or non-industry (33.9%) research organisations[7]. The majority of participants (80.3%) had an extended experience of nine or more years. This defines our sample as a group of mostly academic, male, and Austrian data scientists.

4.2 What is Data Analytics?

We asked participants to describe the term 'Data Analytics' in their own words as an open question to get an idea about the dimensions of the concept and the individual views on the subject. Figure 1 depicts a summary word cloud from the collected free-text responses for all those terms that repeatedly appeared in the response[8]. It further depicts a small set of representative extracts from the comments and definitions that participants submitted. Overall, the comments where very much focused on the issue of large data volumes, the process of knowledge extraction with specific methods and algorithms and the aggregation and combination of data in order to get new insights. Often it was related to machine learning and data mining but as a wider and more integrative approach. Only very few respondents labeled Data Analytics to be simply a modern and fashionable word for data mining or pattern recognition.

4.3 Important Application Areas

Based on the literature review that preceded this survey, we identified the main application areas of data analytics in Austria as healthcare, commerce, manufacturing and logistics, transportation, energy and utilities, the public sector and the government, education, tourism, telecommunication, e-science, law enforcement, and finance and insurance. Figure 2 shows the relative importance of these areas as attributed by participants. Selections were made in binary form with multiple selections possible. The figure shows that the area of healthcare is perceived as a strong sector for data analytics (66.1%) followed by energy (53.6%), manufacturing and e-science (both 50.0%). As a sector that is perceived to benefit only little from (big) data analytics are tourism and commerce (both 23.2%). This is despite the fact that these areas are large in Austria based on demographic data as provided by the Austrian department of Statistics[9].

[5]http://www.limesurvey.org/de/

[6]http://www.r-project.org/

[7]Multiple selections were possible which means that these numbers do not add up to 100%.

[8]We only included terms that appeared a least three times and we filtered with an english and a topical stop word list (e.g. terms like 'and' or 'etc' and terms like 'data' or 'analytics').

[9]In 2010, 19.3% of the employed worked in commerce and 9.1% in the gastronomical and leisure sector (source: 'Ergebnisse im Ueberblick: Statistik zur Unternehmensdemografie 2004 bis 2010', available at http://www.statistik.at/web_de/statistiken/unternehmen_arbeitsstaetten/arbeitgeberunternehmensdemografie/index.html, extracted July 15, 2014.

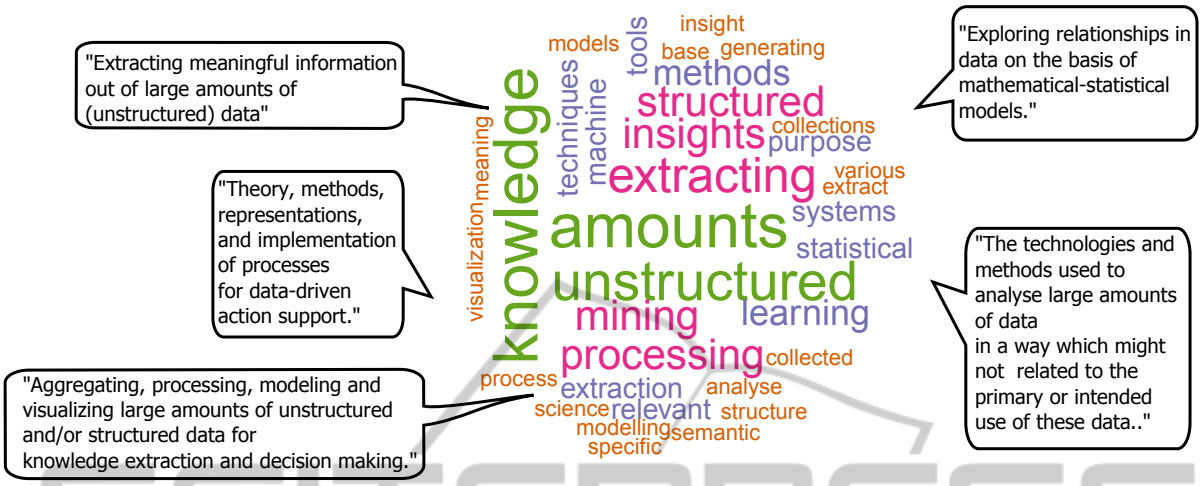

Figure 1: What participants understood as data analytics.

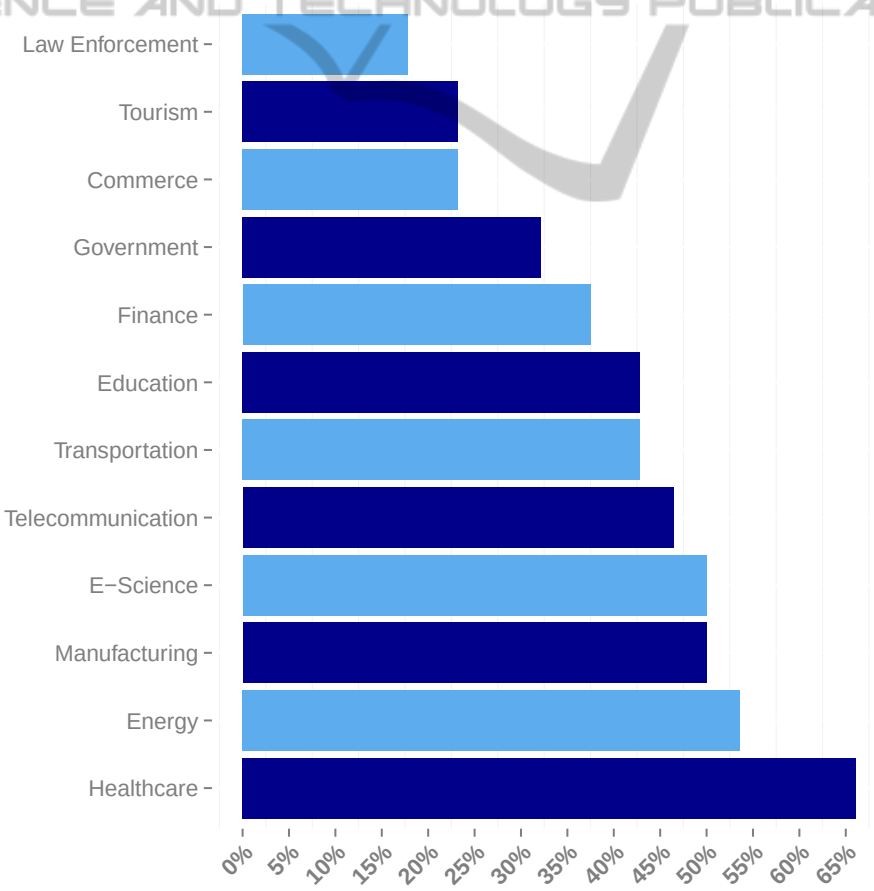

Figure 2: Important application areas for data analytics.

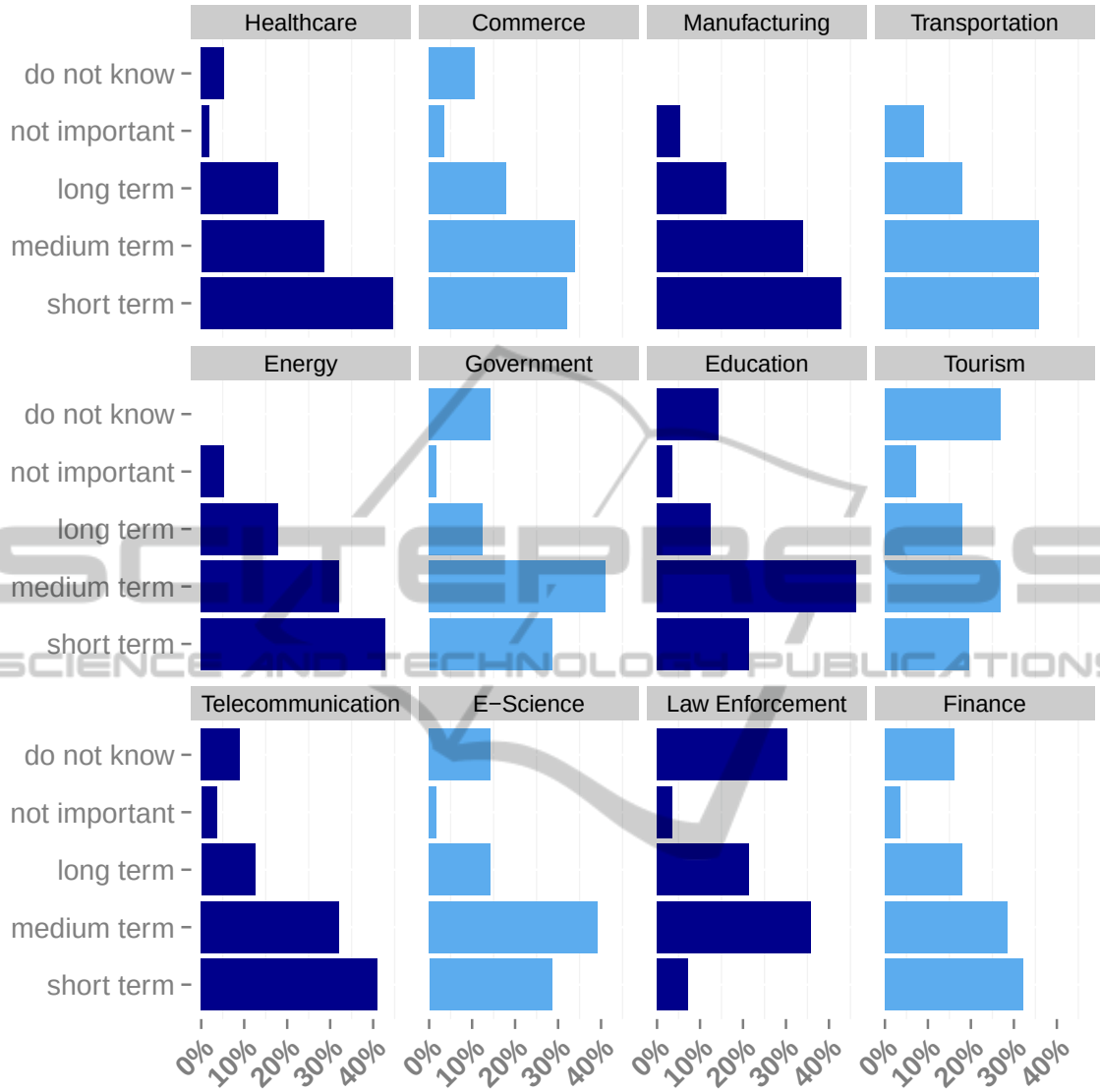

Figure 3: Application areas where data analytics will become important in the short, middle and long term.

We additionally asked participants to provide future projections for these application areas w.r.t. how they think these areas would become important for data analytics in the future (see figure 3). Here, participants rated the application areas based on their relevance for the *short*, *middle* and *long term* future. The diagram also visualizes the amount of uncertainty in these projections as participants could select if they were unsure or even declare an application area as unimportant. The figure shows that application areas that are perceived as strong candidates (e.g. healthcare, energy and telecommunication) are all marked as relevant for the short term future with decreasing ratings on the longer timeline. Less strongly perceived application areas, such as law enforcement and tourism have results that are less clearly expressed with a stronger emphasis on a longer time frame. The amount of uncertainty about these areas is also much higher. Law enforcement is perceived as both less important and not benefiting from data analytics. This is conceivable as law enforcement may not be perceived as an independent sector, as this is the case in the United States (Norton, 2013) where data analytics already assists the crime prediction process with data mining, e.g. with the use of clustering, classification, deviation detection, social network analysis, entity extraction, association rule mining, string comparison and sequential pattern mining. It comes a bit as a surprise that tourism in Austria was both perceived as rather unimportant and also as an area that would only benefit from data analytics in the mid- and long-term future. The large proportion of uncertainty

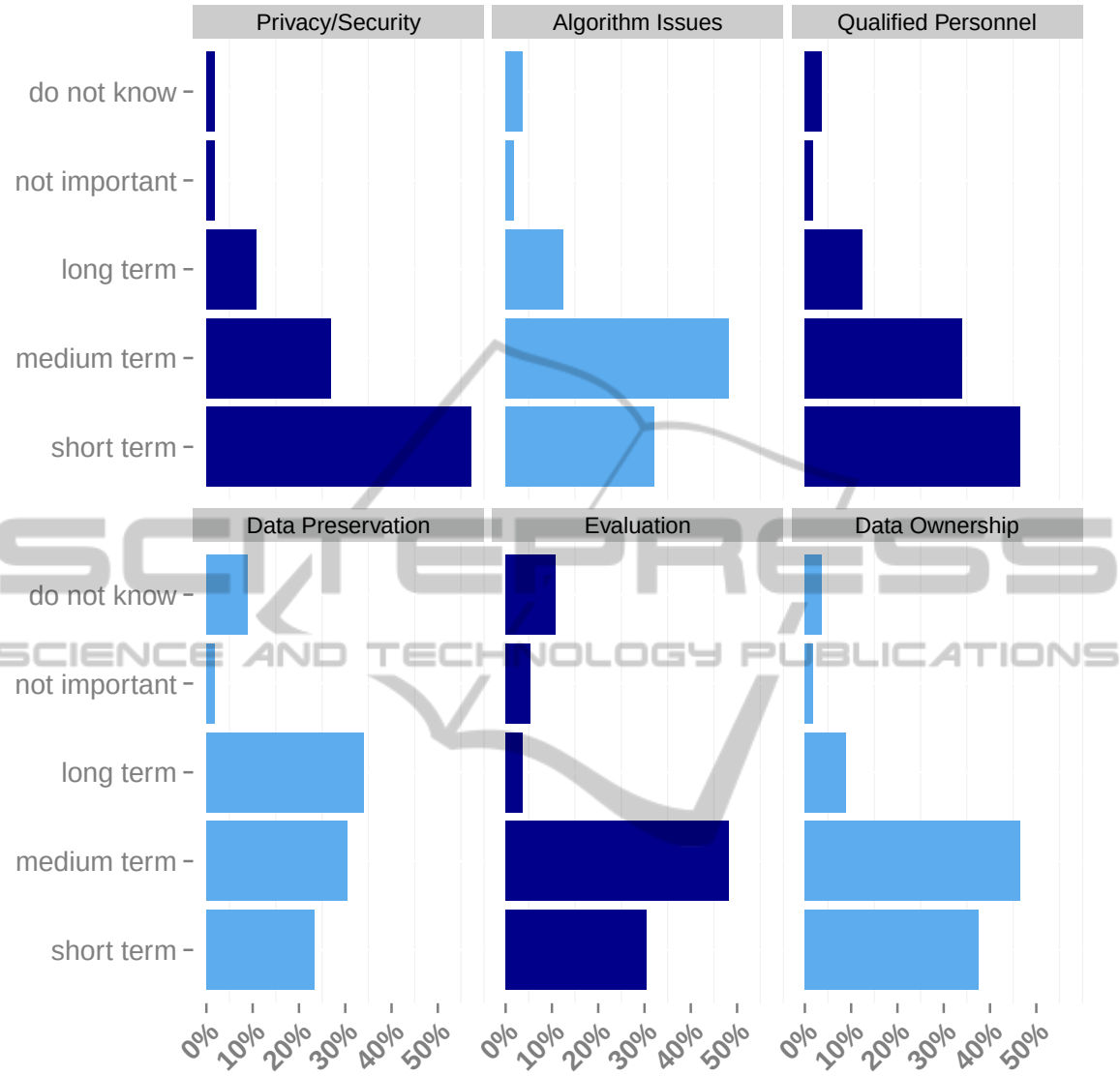

Figure 4: Future challenges in data analytics for the short, middle and long term.

shows that experts seem to be rather unsure about the future of these two sectors.

4.4 Current and Future Challenges

Based on the literature review, we identified the main challenges of data analytics in the areas of privacy and security, algorithmic and scalability, getting qualified personnel, the preservation and curation process, the evaluation and benchmarking, and data ownership and open data. We now asked participants to categorize these challenges into three groups: *Short term* if they see it as an issue of the very near future, *medium term* if there is still time and *long term* if this might become an issue some time in the far future. Our intent was to obtain a priority that can help us to identify possible actions and recommendations for decision making. Figure 4 depicts all responses for all categories and also includes the amount of uncertainty (*do not know*) and how unimportant people thought it to be (*not important*).

All challenges share that they are perceived as all being relevant in the short and middle-term future, with high certainty throughout. Upon closer investigation, the response can then be further divided into two groups.

The first group of challenges consists of privacy and security and the issue of qualified personnel. These issues are perceived as considerably more important and pressing in the short term future than the middle and the long term future. It was especially striking how important *privacy and security* was em-

Table 1: Comparison of Big Data Challenges as identified in three different studies.

CRA Study	SAS Study	Our Study
Lack of Skills / Experience	-	Personnel
Accessing data / Sharing	Human Collaboration	Data Ownership / Data Preservation
Effective Use	-	-
Analysis and Understanding	Data Heterogenity	Algorithm Issues
Scalibility	Scalibility/Timeliness	Evaluation
-	Privacy	Privacy and Security

phasised throughout the entire study — including this online questionnaire. This strong impact might be partially attributed to the very recent NSA scandal. However, it might also be that our target group quite naturally possesses a heightened sensitivity to the potential dangers of (big) data analytics and the often unprotected flow of personal data on the open web. *Qualified personnel* is a problem of the near future and has been discussed in the literature throughout many studies. This is well confirmed in our own findings and an important issue to address in future decision making.

The second group of challenges covers algorithmic and scalability issues, data preservation and curation, evaluation, and data ownership and open data. All of these were attributed more frequently to be issues of the future. Ironically, data preservation and curation has been attributed with being more relevant in the long-term future than the mid- and short-term with the highest amount of uncertainty in the entire response. This should ideally be the opposite. We would have also expected that data ownership and open data issues would be categorized very similar to the privacy and security response and that the algorithmic issues are more relevant on a short term scale as data is mounting very fast. The responses nevertheless demonstrate the feeling that the privacy and security and qualified personnel challenges need to be solved before progress can be made in the field.

We additionally compared our list of challenges with those that were identified in two related, recent studies: One study hosted by the Computing Research Association (CRA) that focused on Challenges and Opportunities of Big Data in general[10] and one study on Big Data visualization by SAS[11]. In Table 1, we refer to them as 'CRA Study', 'SAS Study' and 'Our Study' and compare 6 challenge categories that were identified across these studies. The challenges are presented in no particular order, however, the reader can compare challenge categories horizontally in the table. A dash (-) means that a particular challenge was not identified by a study. We related the categories with each other to give the reader an overview about the similarities and differences from three perspectives. Naturally, the categories did not always represent a perfect match. For example, the challenge of data access and data sharing was addressed as the need for human collaboration in the SAS study and our own study identified the challenge of data ownership and the challenge of preserving data in this category. However, the issue of a lack of personnel was also identified as a lack of skills and experience in the CRA study. Whereas privacy was clearly addressed in the SAS study and more comprehensively combined with security in our study, the CRA study did not consider it a challenge at all. Overall, this comparison shows that there is considerable agreement between studies with respect to future challenges. It would be interesting to further extend this comparison to a much wider range of studies in future work.

4.5 Future Research Topics

We prompted participants with two questions about future research in (big) data analytics. The first question asked them to enter free text on topics of their preferred future research which they had to prioritise by three levels (top priority, 2nd priority, 3rd priority). The second question can be seen as a refinement of the top priority level of the previous question and asked them to describe which research topic they would like to see publicly funded with 10 Million . Again, this was submitted as free text allowing participants to contribute their ideas in a completely free and unrestricted form. Figure 5 shows the term frequencies of those texts for all three priorities and also the text for the 10 Million research topic. This allows for easy comparisons. A sum across all four columns of frequencies provides an overview about the entire topic space.

The most frequent themes are privacy/security (mentioned 22 times) and healthcare (mentioned 17 times) which coincides with the findings from the previous questions. The importance of privacy and security was found to be the most pressing future chal-

[10] http://www.cra.org/ccc/files/docs/init/bigdatawhitepaper.pdf

[11] http://www.sas.com/resources/whitepaper/wp_55536.pdf

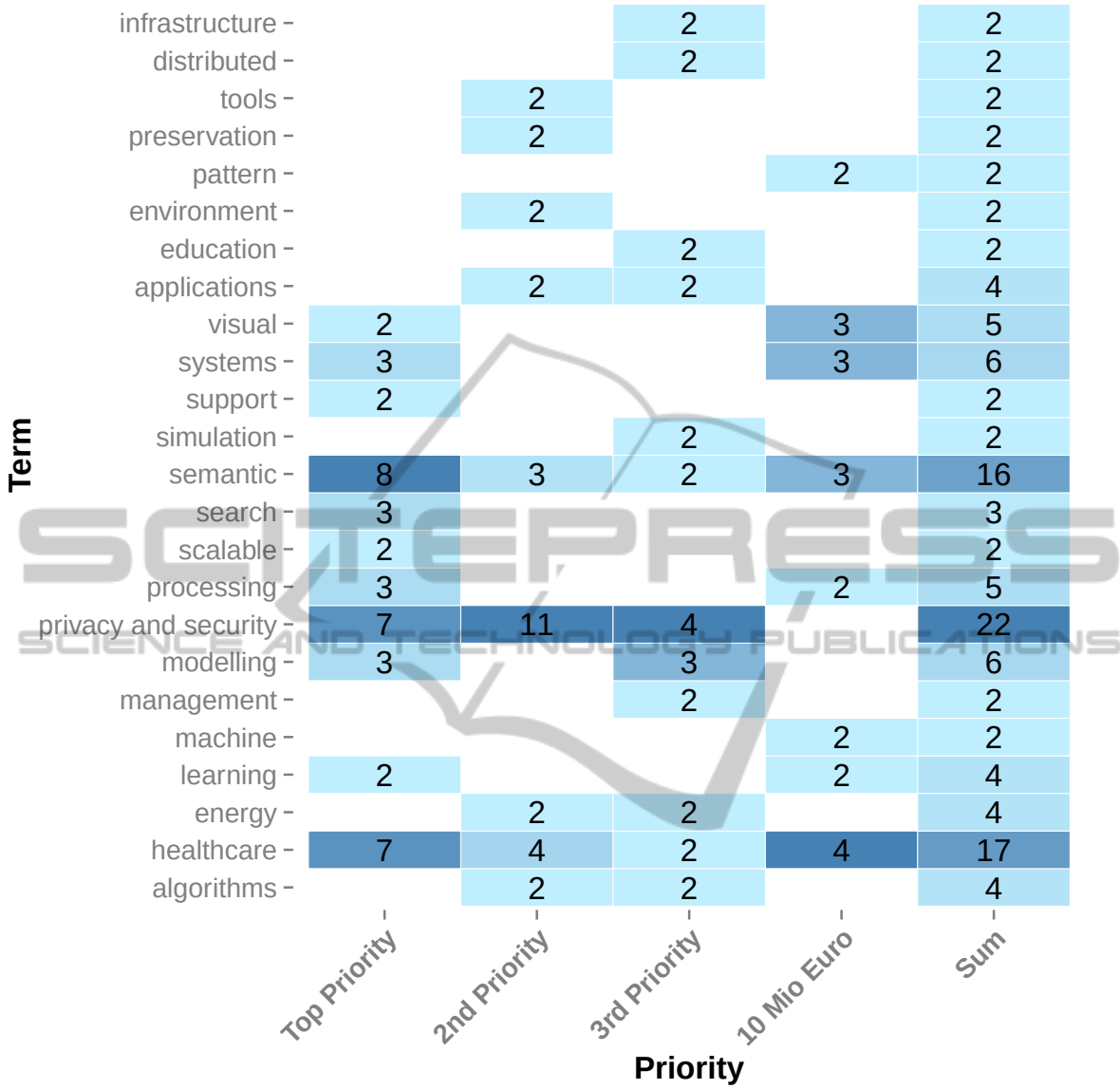

Figure 5: Priorities for future research as expressed by participants.

lenge (see Figure 4). Likewise, healthcare was also perceived as the most promising application domain (see Figure 2) with the most pressing time line that strongly leans toward the short-term future (see Figure 3). The third most frequent keyword were semantic issues (mentioned 16 times) that were more extensively investigated in a number of workshops and an expert interview that is documented in more detail in (Berger et al., 2014).

5 CONCLUSIONS AND TECHNOLOGY ROADMAP

This paper presented results from a study on (big) data analysis in Austria. We summarized the opinons of Austrian data scientists sampled from both industry and the academia and some of the most pressing and current issues in the field.

We found that data analytics is understood as dealing with large data volumes, where knowledge is extracted and aggregated to lead to new insights. It was interesting to see that it was often related to data mining but viewed more widely and more highly in-

tegrated. Healthcare was seen as the most important application area (66.1%), followed by energy (53.6%), manufacturing and logistics (50.0%) and e-science (50.0%) with big potential in the short term future. Other areas were judged less important, such as tourism (23.2%) and law enforcement (17.9%), with high uncertainty. We found that our literature-informed list of challenges were confirmed by our respondents, however only privacy/security and the challenge to get qualified personnel was strongly attributed to the very near future. Algorithm issues, data preservation, evaluation and data ownership were seen as challenges that become more relevant only in the longer run. Research priorities and funding requests where strongly targeted to privacy/security (mentioned 22 times), healthcare (mentioned 17 times) and semantic issues (mentioned 16 times). This result conforms largely to the findings in the other parts of the study.

Based on the results of the survey presented in this paper, along with the outcomes of three workshops and interviews, a technology roadmap consisting of a number of objectives was drawn up. The roadmap actions are described in much greater detail in (Berger et al., 2014) as part of the complete report that focuses on all parts of the study. This is outside the scope of this paper that focuses on the details of the online survey. In summary, the identified challenges, together with their careful evaluation, have led to three categories of actions that are manifested in this roadmap.

First, to meet the challenges of data sharing, evaluation and data preservation, an objective in the roadmap is to create a "Data-Services Ecosystem" in Austria. This is related to an objective to create a legal and regulatory framework that covers issues such as privacy, security and data ownership, as such a framework is necessary to have a functioning Ecosystem. In particular, it is suggested to fund a study project to develop the concept of such an Ecosystem, launch measures to educate and encourage data owners to make their data and problems available, and progress to lighthouse projects to implement and refine the Ecosystem and its corresponding infrastructure. Furthermore, it is recommended to develop a legal framework and create technological framework controls to address the pressing challenges of privacy and security in data analytics.

Second, technical objectives are to overcome challenges related to data integration and fusion and algorithmic efficiency, as well as to create actionable information and revolutionise the way that knowledge work is done. We suggest to fund research that focuses on future data preservation, to develop fusion approaches for very large amounts of data, to create methods that assure anonymity when combining data from many sources, to enable real time processing, and to launch algorithmic challenges. A full list of suggestions are described in more detail in (Berger et al., 2014).

The third and final objective in the roadmap is to increase the number of data scientists being trained. We suggest a comprehensive approach to create these human resources and competences through educational measures at all levels: from schools through universities and universities of applied sciences to companies. The issue of having more and highly skilled data scientists soon is an issue that requires immediate action to secure the future prosperity of the Austrian (big) data analytics landscape.

ACKNOWLEDGEMENTS

This study was commissioned and funded by the Austrian Research Promotion Agency (FFG) and the Austrian Federal Ministry for Transport, Innovation and Technology (BMVIT) as FFG ICT of the Future project number 840200. We thank Andreas Rauber for his valuable input. Information about the project and access to all deliverables are provided at http://www.conqueringdata.com/.

REFERENCES

Agrawal, D., Bernstein, P., Bertino, E., Davidson, S., Dayal, U., Franklin, M., Gehrke, J., Haas, L., Halevy, A., Han, J., Jagadish, H. V., Labrinidis, A., Madden, S., Papakonstantinou, Y., Patel, J. M., Ramakrishnan, R., Ross, K., Shahabi, C., Suciu, D., Vaithyanathan, S., and Widom, J. (2012). Challenges and Opportunities with Big Data. http://www.cra.org/ccc/files/docs/init/bigdatawhitepaper.pdf, last visited: August 2013.

Batinic, B. (2003). Internetbasierte Befragungsverfahren. *Österreichische Zeitschrift für Soziologie*, 28(4):6–18.

Berger, H., Dittenbach, M., Haas, M., Bierig, R., Hanbury, A., Lupu, M., and Piroi, F. (2014). *Conquering Data in Austria*. bmvit (Bundesministerium für Verkehr, Innovation and Technology, Vienna, Austria.

Birnbaum, M. H. (2004). Human research and data collection via the internet. *Annual review of Psychology*, 55:803–832.

Boyd, D. and Crawford, K. (2012). Critical Questions for Big Data. *Information, Communication & Society*, 15(5):662–679.

Dandawate, Y., editor (2013). *Big Data: Challenges and Opportunities*, volume 11 of *Infosys Labs Briefings*. Infosys Labs. http://www.infosys.com/infosys-labs/publications/Documents/bigdata-challenges-opportunities.pdf, last visited: August 2013.

Döring, N. (2003). *Sozialpsychologie des Internet. Die Bedeutung des Internet für Kommunikationsprozesse, Identitäten, soziale Beziehungen und Gruppen*. Hogrefe, Göttingen, 2nd edition.

Harris, H., Murphy, S., and Vaisman, M. (2013). *Analyzing the Analyzers: An Introspective Survey of Data Scientists and Their Work*. O'Reilly.

Manyika, J., Chui, M., Groves, P., Farrell, D., Kuiken, S. V., and Doshi, E. A. (2013). *Open data: Unlocking innovation and performance with liquid information*. McKinsey Global Institute.

Norton, A. (2013). Predictive Policing - The Future of Law Enforcement in the Trinidad and Tobago Police Service. *Int. J. of Computer Applications*, 62:32–36.

Press, G. (2013). The state of big data: What the surveys say. http://www.forbes.com/sites/gilpress/2013/11/30/the-state-of-big-data-what-the-surveys-say/, last visited: March 2014.

Temple, K. (2013). What Happens in an Internet Minute? http://scoop.intel.com/what-happens-in-an-internet-minute/, last visited: August 2013.

APPENDIX: QUESTIONS

The questions of the online survey are categorized in background information, research and development focus, challenges, and public funding.

Background information

Question 1: What is your current activity environment? (Provide comments if you wish to):

- Academia (University)
- Non-University Research
- Industry
- Public Office

Question 2: How many years of experience do you have in your activity area?

- 1-3
- 4-8
- 9 or more

Question 3: Would you consider yourself a... (Provide comments if you wish to)

- Researcher
- Service Provider
- Policy Maker
- User of Data Analytics technology
- Other:

Question 4: Gender

- Male
- Female

Question 5: In which country do you work?

- List of countries

Data Analytics Definition

Question 6: What is your understanding of Data Analytics?

- Free-test answer

Research and Development Focus

Question 7: Which of the following sub-fields do you focus on? (Provide specific details if you wish to)

- Search and Analysis
- Semantic Processing
- Cognitive Systems
- Visualisation and Presentation
- Other:

Question 8: Which of the following Application Domains do you find important today? (i.e. Application Domains you might already be working on. Provide specific details if you wish to)

- Healthcare
- Commerce
- Manufacturing and Logistics
- Transportation
- Energy and Utilities
- Public Sector / Government
- Education
- Tourism
- Telecommunications
- eScience (incl. Life Science)
- Law Enforcement
- Finance and Insurance

Question 9: Other Application Domains you find important

- Free-test answer

Challenges

Question 10: Which challenges do you see in Data Analytics?

- Free-test answer

Question 11: Following your previous answer, please judge if the following challenges will be important in the short, medium or long term.

- Privacy and Security – short term, medium term, long term, not important, don't know
- Algorithm Issues (e.g. Scalability) – short term, medium term, long term, not important, don't know
- Qualified Personnel – short term, medium term, long term, not important, don't know
- Data Preservation and Curation – short term, medium term, long term, not important, don't know
- Evaluation and Benchmarking – short term, medium term, long term, not important, don't know
- Data Ownership and Open Data – short term, medium term, long term, not important, don't know

Question 12: Which challenges do you see in Data Analytics?

- Healthcare – short term, medium term, long term, not important, don't know
- Commerce – short term, medium term, long term, not important, don't know
- Manufacturing and Logistics – short term, medium term, long term, not important, don't know
- Transportation – short term, medium term, long term, not important, don't know
- Energy and Utilities – short term, medium term, long term, not important, don't know
- Public Sector / Government – short term, medium term, long term, not important, don't know
- Education – short term, medium term, long term, not important, don't know
- Tourism – short term, medium term, long term, not important, don't know
- Telecommunications – short term, medium term, long term, not important, don't know
- eScience (incl. Life Science) – short term, medium term, long term, not important, don't know
- Law Enforcement – short term, medium term, long term, not important, don't know
- Finance and Insurance – short term, medium term, long term, not important, don't know

Question 13: Other Application Domains you find important (please indicate Short/Medium/Long Term)

- Free-test answer

Public Funding

Question 14: Which research areas or topics in the Data Analytics field are most important and should be prioritized by public funding (name 3 and rank)

- Top Priority: Free-test answer
- Second Priority: Free-test answer
- Third Priority: Free-test answer

Question 15: Please complete the following news headline: 10,000,000 Euro for...

- Free-test answer

Question 16: Other comments you might have about data analytics

- Free-test answer