МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
УЧРЕЖДЕНИЕ ОБРАЗОВАНИЯ
«Гомельский государственный технический университет имени П.О. Сухого»

КАФЕДРА «ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ»

РЕФЕРАТ

на тему

ПРОГРАММНЫЙ КОМПЛЕКС ДЛЯ ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ РОЖДЕНИЯ Z' - БОЗОНОВ В ПРОТОН-ПРОТОННЫХ СТОЛКНОВЕНИЯХ С УЧЕТОМ ЭФФЕКТОВ Z - Z' СМЕШИВАНИЯ

подготовленный для прохождения итоговой аттестации по общеобразовательной дисциплине «Основы информационных технологи»

Выполнил:
магистрант гр. МАГ 40-12 специальности 1–40 80 04 «Математическое моделирование, численные методы и комплексы программ»
Бурим Илья Павлович

Проверил:
доцент кафедры «Информационные технологии»
Цитринов А.В.

Гомель 2017

СОДЕРЖАНИЕ

ВВЕДЕНИЕ **3**

1 BUSINESS INTELLIGENCE AND ANALYTICS: FROM BIG DATA TO BIG IMPACT **4**

1.1 BI&A 1.0 4
1.2 BI&A 2.0 5
1.3 BI&A 3.0 6
1.4 BI&A Applications: From Big Data to Big Impact 7
1.5 E-Commerce and Market Intelligence 8
1.6 E-Government and Politics 2.0 9
1.7 Science and Technology 10

2 Рождения Z' - бозонов в протон-протонных столкновениях с учетом эффектов Z - Z' смешивания **12**

ЗАКЛЮЧЕНИЕ **13**

Список использованных источников **14**

ПРИЛОЖЕНИЕ **16**

ВВЕДЕНИЕ

фывфы выфвыф вфывфы

ГЛАВА 1

BUSINESS INTELLIGENCE AND ANALYTICS: FROM BIG DATA TO BIG IMPACT

1.1 BI&A 1.0

The term intelligencehas been used by researchers in artificial intelligence since the 1950s. Business intelligence became a popular term in the business and IT communities only in the 1990s. In the late 2000s, business analyticswas introduced to represent the key analytical component in BI [1]. More recently big dataand big data analyticshave been used to describe the data sets and analytical techniques in applications that are so large (from terabytes to exabytes) and complex (from sensor to social media data) that they require advanced and unique data

As a data-centric approach, BI&A has its roots in the longstanding database management field. It relies heavily on various data collection, extraction, and analysis technologies (Chaudhuri et al. 2011; Turban et al. 2008; Watson and Wixom 2007). The BI&A technologies and applications currently adopted in industry can be considered as BI&A 1.0, where data are mostly structured, collected by companies through various legacy systems, and often stored in commercial relational database management systems (RDBMS). The analytical techniques commonly used in these systems, popularized in the 1990s, are grounded mainly in statistical methods developed in the 1970s and data mining techniques developed in the 1980s.

Data management and warehousing is considered the foundation of BI&A 1.0. Design of data marts and tools for extraction, transformation, and load (ETL) are essential for converting and integrating enterprise-specific data. Database query, online analytical processing (OLAP), and reporting tools based on intuitive, but simple, graphics are used to explore important data characteristics. Business performance management (BPM) using scorecards and dashboards help analyze and visualize a variety of performance metrics. In addition to these well-established business reporting functions, statistical analysis and data mining techniques are adopted for association analysis, data segmentation and clustering, classification and regression analysis, anomaly detection, and predictive modeling in various business applications. Most of these data processing and analytical technologies have already been incorporated into the leading commercial BI platforms offered by major IT

vendors including Microsoft, IBM, Oracle, and SAP (Sallam et al. 2011).

Among the 13 capabilities considered essential for BI platforms, according to the Gartner report by Sallam et al. (2011), the following eight are considered BI&A 1.0: reporting, dashboards, ad hocquery, search-based BI, OLAP, interactive visualization, scorecards, predictive modeling, and data mining. A few BI&A 1.0 areas are still under active development based on the Gartner BI Hype Cycle analysis for emerging BI technologies, which include data mining workbenchs, column-based DBMS, in-memory DBMS, and realtime decision tools (Bitterer 2011). Academic curricula in Information Systems (IS) and Computer Science (CS) often include well-structured courses such as database management systems, data mining, and multivariate statistics.

In conclusion the opportunities associated with data and analysis in different organizations have helped generate significant interest in BI&A, which is often referred to as the techniques, technologies, systems, practices, methodologies, and applications that analyze critical business data to help an enterprise better understand its business and market and make timely business decisions.

1.2 BI&A 2.0

The Internet and the Web began to offer unique data collection and analytical research and development opportunities. In 2000s the HTTP-based Web 1.0 systems, characterized by Web search engines such as Google and Yahoo and e-commerce businesses such as Amazon and eBay, allow organizations to present their businesses online and interact with their customers directly. In addition to porting their traditional RDBMS-based product information and business contents online, detailed and IP-specific user search and interaction logs that are collected seamlessly through cookies and server logs have become a new gold mine for understanding customers' needs and identifying new business opportunities. Web intelligence, web analytics, and the user-generated content collected through Web 2.0-based social and crowd-sourcing systems (Doan et al. 2011; O'Reilly 2005) have ushered in a new and exciting era of BI&A 2.0 research in the 2000s, centered on text and web analytics for unstructured web contents.

An immense amount of company, industry, product, and customer information can be gathered from the web and organized and visualized through various text and web mining techniques. By analyzing customer clickstream data logs, web analytics tools such as Google Analytics can provide a trail of the user's online activities and reveal the user's browsing and purchasing patterns. Web site design, product placement optimization, customer transaction analysis, market structure analysis, and product

recommendations can be accomplished through web analytics. The many Web 2.0 applications developed after 2004 have also created an abundance of user-generated content from various online social media such as forums, online groups, web blogs, social networking sites, social multimedia sites (for photos and videos), and even virtual worlds and social games (O'Reilly 2005). In addition to capturing celebrity chatter, references to everyday events, and socio-political sentiments expressed in these media, Web 2.0 applications can efficiently gather a large volume of timely feedback and opinions from a diverse customer population for different types of businesses.

Many marketing researchers believe that social media analytics presents a unique opportunity for businesses to treat the market as a "conversation" between businesses and customers instead of the traditional business-to-customer, one-way "marketing" (Lusch et al. 2010). Unlike BI&A 1.0 technologies that are already integrated into commercial enterprise IT systems, future BI&A 2.0 systems will require the integration of mature and scalable techniques in text mining (e.g., information extraction, topic identification, opinion mining, question-answering), web mining, social network analysis, and spatial-temporal analysis with existing DBMS-based BI&A 1.0 systems.

Except for basic query and search capabilities, no advanced text analytics for unstructured content are currently considered in the 13 capabilities of the Gartner BI platforms. Several, however, are listed in the Gartner BI Hype Cycle, including information semantic services, natural language question answering, and content/text analytics (Bitterer 2011). New IS and CS courses in text mining and web mining have emerged to address needed technical training.

In conclusion it can be notified that in 2000s, BI&A created big jump in infrastructure natural language, information extraction, topic identification, opinion mining, question-answering. And created new directions in IT.

1.3 BI&A 3.0

This chapter describes most of the academic research on mobile BI, opening up exciting new steams of innovative applications and describes business intelligence and analytics in Web 3.0 area.

Whereas web-based BI&A 2.0 has attracted active research from academia and industry, a new research opportunity in BI&A 3.0 is emerging. As reported prominently in an October 2011 article in The Economist (2011), the number of mobile phones and tablets (about 480 million units) surpassed the number of laptops and PCs (about 380 million units) for the first time in 2011. Although the number of PCs in use surpassed 1 billion in 2008, the same article projected that the number of mobile connected devices would reach

10 billion in 2020. Mobile devices such as the iPad, iPhone, and other smart phones and their complete ecosystems of downloadable applicationss, from travel advisories to multi-player games, are transforming different facets of society, from education to healthcare and from entertainment to governments. Other sensor-based Internet-enabled devices equipped with RFID, barcodes, and radio tags (the "Internet of Things") are opening up exciting new steams of innovative applications. The ability of such mobile and Internet-enabled devices to support highly mobile, location-aware, person-centered, and context-relevant operations and transactions will continue to offer unique research challenges and opportunities throughout the 2010s. Mobile interface, visualization, and HCI (human–computer interaction) design are also promising research areas. Although the coming of the Web 3.0 (mobile and sensor-based) era seems certain, the underlying mobile analytics and location and context-aware techniques for collecting, processing, analyzing and visualizing such largescale and fluid mobile and sensor data are still unknown.

No integrated, commercial BI&A 3.0 systems are foreseen for the near future. Most of the academic research on mobile BI is still in an embryonic stage. Although not included in the current BI platform core capabilities, mobile BI has been included in the Gartner BI Hype Cycle analysis as one of the new technologies that has the potential to disrupt the BI market significantly (Bitterer 2011). The uncertainty associated with BI&A 3.0 presents another unique research direction for the IS community. Table 1 summarizes the key characteristics of BI&A 1.0, 2.0, and 3.0 in relation to the Gartner BI platforms core capabilities and hype cycle.

In conclusion it can be notified that the decade of the 2010s was an exciting one for high-impact BI&A research and development for both industry and academia. IS research and education programs need to carefully evaluate future directions, curricula, and action plans, from BI&A 1.0 to 3.0. The business community and industry have already taken important steps to adopt BI&A for their needs. The IS community faces unique challenges and opportunities in making scientific and societal impacts that are relevant and long-lasting (Chen 2011a).

1.4 BI&A Applications: From Big Data to Big Impact

This chapter describes new streams where business intelligence and analytics big data will be used. Streams like international travel, high-speed network connections, global supply-chain, and outsourcing have created a tremendous opportunity for IT advancement. It predicted by Thomas Freeman in his seminal book, The World is Flat (2005).

Several global business and IT trends have helped shape past and present BI&A research directions. In addition to ultra-fast global IT connections, the development and deployment of business-related data standards, electronic data interchange (EDI) formats, and business databases and information systems have greatly facilitated business data creation and utilization. The development of the Internet in the 1970s and the subsequent large-scale adoption of the World Wide Web since the 1990s have increased business data generation and collection speeds exponentially. Recently, the Big Data era has quietly descended on many communities, from governments and e-commerce to health organizations. With an overwhelming amount of web-based, mobile, and sensorgenerated data arriving at a terabyte and even exabyte scale (The Economist 2010a, 2010b), new science, discovery, and insights can be obtained from the highly detailed, contextualized, and rich contents of relevance to any business or organization.

In addition to being data driven, BI&A is highly applied and can leverage opportunities presented by the abundant data and domain-specific analytics needed in many critical and highimpact application areas. Several of these promising and high-impact BI&A applications are presented below, with a discussion of the data and analytics characteristics, potential impacts, and selected illustrative examples or studies: (1) ecommerce and market intelligence, (2) e-government and politics 2.0, (3) science and technology, (4) smart health and well-being, and (5) security and public safety. By carefully analyzing the application and data characteristics, researchers and practitioners can then adopt or develop the appropriate analytical techniques to derive the intended impact. IS departments thus face unique opportunities and challenges in developing integrated BI&A research and education programs for the new generation of data/analyticssavvy and business-relevant students and professionals (Chen 2011a).

In conclusion to technical system implementation, significant business or domain knowledge as well as effective communication skills are needed for the successful completion of such BI&A projects.

1.5 E-Commerce and Market Intelligence

This chapter describes new impacts and development directions in business intelligence and analytics big data for e-commerce organizations.

The excitement surrounding BI&A and Big Data has arguably been generated primarily from the web and e-commerce communities. Significant market transformation has been accomplished by leading e-commerce vendors such Amazon and eBay through their innovative and highly scalable ecommerce

platforms and product recommender systems. Major Internet firms such as Google, Amazon, and Facebook continue to lead the development of web analytics, cloud computing, and social media platforms. The emergence of customer-generated Web 2.0 content on various forums, newsgroups, social media platforms, and crowd-sourcing systems offers another opportunity for researchers and practitioners to "listen" to the voice of the market from a vast number of business constituents that includes customers, employees, investors, and the media (Doan et al. 2011; O'Rielly 2005). Unlike traditional transaction records collected from various legacy systems of the 1980s, the data that e-commerce systems collect from the web are less structured and often contain rich customer opinion and behavioral information. For social media analytics of customer opinions, text analysis and sentiment analysis techniques are frequently adopted (Pang and Lee 2008). Various analytical techniques have also been developed for product recommender systems, such as association rule mining, database segmentation and clustering, anomaly detection, and graph mining (Adomavicius and Tuzhilin 2005). Long-tail marketing accomplished by reaching the millions of niche markets at the shallow end of the product bitstream has become possible via highly targeted searches and personalized recommendations (Anderson 2004). The Netfix Prize competition for the best collaborative filtering algorithm to predict user movie ratings helped generate significant academic and industry interest in recommender systems development and resulted in awarding the grand prize of $1 million to the Bellkor's Pragmatic Chaos team, which surpassed Netflix's own algorithm for predicting ratings by 10.06 percent. However, the publicity associated with the competition also raised major unintended customer privacy concerns.

In conclusion much BI&A-related e-commerce research and development information is appearing in academic IS and CS papers as well as in popular IT magazines.

1.6 E-Government and Politics 2.0

This chapter describes cahnges e-commerce area with coming Web 2.0 and with coming new technologies in BI&A research. The advent of Web 2.0 has generated much excitement for reinventing governments.

The 2008 U.S. House, Senate, and presidential elections provided the first signs of success for online campaigning and political participation. Dubbed «politics 2.0», politicians use the highly participatory and multimedia web platforms for successful policy discussions, campaign advertising, voter mobilization, event announcements, and online donations. As government and political processes become more transparent, participatory, online, and

multimedia-rich, there is a great opportunity for adopting BI&A research in e-government and politics 2.0 applications. Selected opinion mining, social network analysis, and social media analytics techniques can be used to support online political participation, e-democracy, political blogs and forums analysis, e-government service delivery, and process transparency and accountability (Chen 2009; Chen et al. 2007). For e-government applications, semantic information directory and ontological development (as exemplified below) can also be developed to better serve their target citizens.

Despite the significant transformational potential for BI&A in e-government research, there has been less academic research than, for example, e-commerce-related BI&A research. Egovernment research often involves researchers from political science and public policy. For example, Karpf (2009) analyzed the growth of the political blogosphere in the United States and found significant innovation of existing political institutions in adopting blogging platforms into their Web offerings. In his research, 2D blogspace mapping with composite rankings helped reveal the partisan makeup of the American political blogsphere. Yang and Callan (2009) demonstrated the value for ontology development for government services through their development of the OntoCop system, which works interactively with a user to organize and summarize online public comments from citizens.

In conclusion it can be notified that e-commerce area will be grow with BI&A science. Also it can be notified that more important for e-commerce will be the BI&A in e-government research and e-commerce-related BI&A research directions.

1.7 Science and Technology

This chapter describes using business intelligence and analytics big data for science. Describes tools and algorithms for successful opening of new achievements and researchs.

Many areas of science and technology (S&T) are reaping the benefits of high-throughput sensors and instruments, from astrophysics and oceanography, to genomics and environmental research. To facilitate information sharing and data analytics, the National Science Foundation (NSF) recently mandated that every project is required to provide a data management plan. Cyber-infrastructure, in particular, has become critical for supporting such data-sharing initiatives.

The 2012 NSF BIGDATA program solicitation is an obvious example of the U.S. government funding agency's concerted efforts to promote big data analytics. The program aims to advance the core scientific and technological

means of managing, analyzing, visualizing, and extracting useful information from large, diverse, distributed and heterogeneous data sets so as to accelerate the progress of scientific discovery and innovation; lead to new fields of inquiry that would not otherwise be possible; encourage the development of new data analytic tools and algorithms; facilitate scalable, accessible, and sustainable data infrastructure; increase understanding of human and social processes and interactions; and promote economic growth and improved health and quality of life.

Several S&T disciplines have already begun their journey toward big data analytics. For example, in biology, the NSF funded iPlant Collaborative is using cyberinfrastructure to support a community of researchers, educators, and students working in plant sciences. iPlant is intended to foster a new generation of biologists equipped to harness rapidly expanding computational techniques and growing data sets to address the grand challenges of plant biology. The iPlant data set is diverse and includes canonical or reference data, experimental data, simulation and model data, observational data, and other derived data. It also offers various open source data processing and analytics tools.

In astronomy, the Sloan Digital Sky Survey (SDSS) shows how computational methods and big data can support and facilitate sense making and decision making at both the macroscopic and the microscopic level in a rapidly growing and globalized research field. The SDSS is one of the most ambitious and influential surveys in the history of astronomy. Over its eight years of operation, it has obtained deep, multicolor images covering more than a quarter of the sky and created three-dimensional maps containing more than 930,000 galaxies and over 120,000 quasars. Continuing to gather data at a rate of 200 gigabytes per night, SDSS has amassed more than 140 terabytes of data. The international Large Hadron Collider (LHC) effort for high-energy physics is another example of big data, producing about 13 petabytes of data in a year (Brumfiel 2011).

In conclusion it can be notified that business intelligence and analytics big data contributed to the big jump in astronomy, physics high-energy. The relationship between BI&A and science contributed to delevopment of new data processing and analytics tools, algorithms.

ЗАКЛЮЧЕНИЕ

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Davenport, T. H. 2006. «Competing on Analytics,» Harvard Business Review(84:1), p. 98-107.
2. Основные понятия ОС [Электронный ресурс]. — Режим доступа: http://technomag.bmstu.ru/doc/48639.html — Дата доступа: 11.12.2017.
3. Операционные системы Linux [Электронный ресурс]. — Режим доступа: http://help.ubuntu.ru/wiki/linux — Дата доступа: 11.12.2017.
4. Linux-2017: самые перспективные дистрибутивы [Электронный ресурс]. — Режим доступа: https://habrahabr.ru/company/ruvds/blog/320002/ — Дата доступа: 11.12.2017.
5. За пределами Стандартной модели [Электронный ресурс]. — Режим доступа: https://elementy.ru/LHC/HEP/SM/beyondSM — Дата доступа: 11.12.2017.
6. Н. В. Красников, В. А. Матвеев. Поиск новой физики на LHC [Электронный ресурс]. — Режим доступа: http://nuclphys.sinp.msu.ru/ATLAS_exp/at03.htm — Дата доступа: 11.12.2017.
7. Official documentation [Электронный ресурс]. — Режим доступа: http://home.thep.lu.se/ torbjorn/Pythia.html — Дата доступа: 11.12.2017.
8. Бобовников, И.Д. Эффекты $Z - Z'$-смешивания в процессах рождения пары $W^{\pm}$-бозонов на адронных и лептонных коллайдерах высоких энергий / И.Д. Бобовников, А.А. Панков. — Письма в ЭЧАЯ, 2016. Т. 13, №1(199). С.8-35
9. Слабое взаимодействие [Электронныйресурс]. — Режим доступа: http://nuclphys.sinp.msu.ru/enc/e149.htm — Дата доступа: 11.12.2017.
10. Osland, P. Probing $Z - Z'$ mixing with ATLAS and CMS resonant diboson production data at the LHC at $\sqrt{s} = 13$ TeV / P. Osland, A.A. Pankov, A.V. Tsytrinov // Physical Review D. – 2012. – Vol. 86. – P. 12.

11. Andreev, V. V. Constraints on the $Z - Z'$ mmixing angle from data measured for the process $e^+e^- \to W^+W^-$ at the LEP2 collider / V.V. Andreev, A.A. Pankov // Phys. At. Nucl. – 2012. – Vol. 75. – P. 76.

12. ALEPH and DELPHI and L3 and OPAL and SLD Collaborations and LEP Electroweak Working Group and SLD Electroweak Group and SLD Heavy Flavour Group / Schael, S. [et al.] // Precision electroweak measurements on the Zresonance, Phys. Rep. 2006. – P. 427.

13. Search for massive resonances decaying into WW, WZ or ZZ bosons in proton-proton collisions at $\sqrt{s} = 13$ TeV / Sirunyan, A. M. [et al.] // J. High Energy Phys. – 2017. – Vol. 162. – P. 56.

14. Measurement of W^+W^-−production in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector and limits on anomalous WWZ and WW_y couplings / Ada, G. [et al.] // Phys. Rev. D. – 2013. – Vol. 88. – P. 29.

15. Search for new phenomena in the $WW \to lvl'v'$ final state in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector / Ada, G. [et al.] // Physics Letters B. – 2013. – Vol. 3. – P. 878.

ПРИЛОЖЕНИЕ

BI&A 1.0

As a data-centric approach, BI&A has its roots in the longstanding database management field. It relies heavily on various data collection, extraction, and analysis technologies (Chaudhuri et al. 2011; Turban et al. 2008; Watson and Wixom 2007). The BI&A technologies and applications currently adopted in industry can be considered as BI&A 1.0, where data are mostly structured, collected by companies through various legacy systems, and often stored in commercial relational database management systems (RDBMS). The analytical techniques commonly used in these systems, popularized in the 1990s, are grounded mainly in statistical methods developed in the 1970s and data mining techniques developed in the 1980s.

Data management and warehousing is considered the foundation of BI&A 1.0. Design of data marts and tools for extraction, transformation, and load (ETL) are essential for converting and integrating enterprise-specific data. Database query, online analytical processing (OLAP), and reporting tools based on intuitive, but simple, graphics are used to explore important data characteristics. Business performance management (BPM) using scorecards and dashboards help analyze and visualize a variety of performance metrics. In addition to these well-established business reporting functions, statistical analysis and data mining techniques are adopted for association analysis, data segmentation and clustering, classification and regression analysis, anomaly detection, and predictive modeling in various business applications. Most of these data processing and analytical technologies have already been incorporated into the leading commercial BI platforms offered by major IT vendors including Microsoft, IBM, Oracle, and SAP (Sallam et al. 2011).

Among the 13 capabilities considered essential for BI platforms, according to the Gartner report by Sallam et al. (2011), the following eight are considered BI&A 1.0: reporting, dashboards, ad hocquery, search-based BI, OLAP, interactive visualization, scorecards, predictive modeling, and data mining. A few BI&A 1.0 areas are still under active development based on the Gartner BI Hype Cycle analysis for emerging BI technologies, which include data mining workbenchs, column-based DBMS, in-memory DBMS, and realtime decision tools (Bitterer 2011). Academic curricula in Information Systems (IS) and Computer Science (CS) often include well-structured courses such as database management systems, data mining, and multivariate statistics.

BI&A 2.0

Since the early 2000s, the Internet and the Web began to offer unique data collection and analytical research and development opportunities. The HTTP-based Web 1.0 systems, characterized by Web search engines such as Google and Yahoo and e-commerce businesses such as Amazon and eBay, allow organizations to present their businesses online and interact with their customers directly. In addition to porting their traditional RDBMS-based product information and business contents online, detailed and IP-specific user search and interaction logs that are collected seamlessly through cookies and server logs have become a new gold mine for understanding customers' needs and identifying new business opportunities. Web intelligence, web analytics, and the user-generated content collected through Web 2.0-based social and crowd-sourcing systems (Doan et al. 2011; O'Reilly 2005) have ushered in a new and exciting era of BI&A 2.0 research in the 2000s, centered on text and web analytics for unstructured web contents.

An immense amount of company, industry, product, and customer information can be gathered from the web and organized and visualized through various text and web mining techniques. By analyzing customer clickstream data logs, web analytics tools such as Google Analytics can provide a trail of the user's online activities and reveal the user's browsing and purchasing patterns. Web site design, product placement optimization, customer transaction analysis, market structure analysis, and product recommendations can be accomplished through web analytics. The many Web 2.0 applications developed after 2004 have also created an abundance of user-generated content from various online social media such as forums, online groups, web blogs, social networking sites, social multimedia sites (for photos and videos), and even virtual worlds and social games (O'Reilly 2005). In addition to capturing celebrity chatter, references to everyday events, and socio-political sentiments expressed in these media, Web 2.0 applications can efficiently gather a large volume of timely feedback and opinions from a diverse customer population for different types of businesses.

Many marketing researchers believe that social media analytics presents a unique opportunity for businesses to treat the market as a "conversation" between businesses and customers instead of the traditional business-to-customer, one-way "marketing" (Lusch et al. 2010). Unlike BI&A 1.0 technologies that are already integrated into commercial enterprise IT systems, future BI&A 2.0 systems will require the integration of mature and scalable techniques in text mining (e.g., information extraction, topic identification, opinion mining, question-answering), web mining, social network analysis, and

spatial-temporal analysis with existing DBMS-based BI&A 1.0 systems.

Except for basic query and search capabilities, no advanced text analytics for unstructured content are currently considered in the 13 capabilities of the Gartner BI platforms. Several, however, are listed in the Gartner BI Hype Cycle, including information semantic services, natural language question answering, and content/text analytics (Bitterer 2011). New IS and CS courses in text mining and web mining have emerged to address needed technical training.

BI&A 3.0

Whereas web-based BI&A 2.0 has attracted active research from academia and industry, a new research opportunity in BI&A 3.0 is emerging. As reported prominently in an October 2011 article in The Economist (2011), the number of mobile phones and tablets (about 480 million units) surpassed the number of laptops and PCs (about 380 million units) for the first time in 2011. Although the number of PCs in use surpassed 1 billion in 2008, the same article projected that the number of mobile connected devices would reach 10 billion in 2020. Mobile devices such as the iPad, iPhone, and other smart phones and their complete ecosystems of downloadable applicationss, from travel advisories to multi-player games, are transforming different facets of society, from education to healthcare and from entertainment to governments. Other sensor-based Internet-enabled devices equipped with RFID, barcodes, and radio tags (the "Internet of Things") are opening up exciting new steams of innovative applications. The ability of such mobile and Internet-enabled devices to support highly mobile, location-aware, person-centered, and context-relevant operations and transactions will continue to offer unique research challenges and opportunities throughout the 2010s. Mobile interface, visualization, and HCI (human–computer interaction) design are also promising research areas. Although the coming of the Web 3.0 (mobile and sensor-based) era seems certain, the underlying mobile analytics and location and context-aware techniques for collecting, processing, analyzing and visualizing such largescale and fluid mobile and sensor data are still unknown.

No integrated, commercial BI&A 3.0 systems are foreseen for the near future. Most of the academic research on mobile BI is still in an embryonic stage. Although not included in the current BI platform core capabilities, mobile BI has been included in the Gartner BI Hype Cycle analysis as one of the new technologies that has the potential to disrupt the BI market significantly (Bitterer 2011). The uncertainty associated with BI&A 3.0 presents another unique research direction for the IS community. Table 1 summarizes the key characteristics of BI&A 1.0, 2.0, and 3.0 in relation to the Gartner BI platforms

core capabilities and hype cycle. The decade of the 2010s promises to be an exciting one for high-impact BI&A research and development for both industry and academia. The business community and industry have already taken important steps to adopt BI&A for their needs. The IS community faces unique challenges and opportunities in making scientific and societal impacts that are relevant and long-lasting (Chen 2011a). IS research and education programs need to carefully evaluate future directions, curricula, and action plans, from BI&A 1.0 to 3.0.

BI&A Applications: From Big Data to Big Impact

Several global business and IT trends have helped shape past and present BI&A research directions. International travel, high-speed network connections, global supply-chain, and outsourcing have created a tremendous opportunity for IT advancement, as predicted by Thomas Freeman in his seminal book, The World is Flat (2005). In addition to ultra-fast global IT connections, the development and deployment of business-related data standards, electronic data interchange (EDI) formats, and business databases and information systems have greatly facilitated business data creation and utilization. The development of the Internet in the 1970s and the subsequent large-scale adoption of the World Wide Web since the 1990s have increased business data generation and collection speeds exponentially. Recently, the Big Data era has quietly descended on many communities, from governments and e-commerce to health organizations. With an overwhelming amount of web-based, mobile, and sensorgenerated data arriving at a terabyte and even exabyte scale (The Economist 2010a, 2010b), new science, discovery, and insights can be obtained from the highly detailed, contextualized, and rich contents of relevance to any business or organization.

In addition to being data driven, BI&A is highly applied and can leverage opportunities presented by the abundant data and domain-specific analytics needed in many critical and highimpact application areas. Several of these promising and high-impact BI&A applications are presented below, with a discussion of the data and analytics characteristics, potential impacts, and selected illustrative examples or studies: (1) ecommerce and market intelligence, (2) e-government and politics 2.0, (3) science and technology, (4) smart health and well-being, and (5) security and public safety. By carefully analyzing the application and data characteristics, researchers and practitioners can then adopt or develop the appropriate analytical techniques to derive the intended impact. In addition to technical system implementation, significant business or domain knowledge as well as effective communication skills are needed for

the successful completion of such BI&A projects. IS departments thus face unique opportunities and challenges in developing integrated BI&A research and education programs for the new generation of data/analyticssavvy and business-relevant students and professionals (Chen 2011a).

E-Commerce and Market Intelligence

The excitement surrounding BI&A and Big Data has arguably been generated primarily from the web and e-commerce communities. Significant market transformation has been accomplished by leading e-commerce vendors such Amazon and eBay through their innovative and highly scalable ecommerce platforms and product recommender systems. Major Internet firms such as Google, Amazon, and Facebook continue to lead the development of web analytics, cloud computing, and social media platforms. The emergence of customer-generated Web 2.0 content on various forums, newsgroups, social media platforms, and crowd-sourcing systems offers another opportunity for researchers and practitioners to "listen" to the voice of the market from a vast number of business constituents that includes customers, employees, investors, and the media (Doan et al. 2011; O'Rielly 2005). Unlike traditional transaction records collected from various legacy systems of the 1980s, the data that e-commerce systems collect from the web are less structured and often contain rich customer opinion and behavioral information. For social media analytics of customer opinions, text analysis and sentiment analysis techniques are frequently adopted (Pang and Lee 2008). Various analytical techniques have also been developed for product recommender systems, such as association rule mining, database segmentation and clustering, anomaly detection, and graph mining (Adomavicius and Tuzhilin 2005). Long-tail marketing accomplished by reaching the millions of niche markets at the shallow end of the product bitstream has become possible via highly targeted searches and personalized recommendations (Anderson 2004). The Netfix Prize competition for the best collaborative filtering algorithm to predict user movie ratings helped generate significant academic and industry interest in recommender systems development and resulted in awarding the grand prize of $1 million to the Bellkor's Pragmatic Chaos team, which surpassed Netflix's own algorithm for predicting ratings by 10.06 percent. However, the publicity associated with the competition also raised major unintended customer privacy concerns.

Much BI&A-related e-commerce research and development information is appearing in academic IS and CS papers as well as in popular IT magazines.

E-Government and Politics 2.0

The advent of Web 2.0 has generated much excitement for reinventing governments. The 2008 U.S. House, Senate, and presidential elections provided the first signs of success for online campaigning and political participation. Dubbed «politics 2.0», politicians use the highly participatory and multimedia web platforms for successful policy discussions, campaign advertising, voter mobilization, event announcements, and online donations. As government and political processes become more transparent, participatory, online, and multimedia-rich, there is a great opportunity for adopting BI&A research in e-government and politics 2.0 applications. Selected opinion mining, social network analysis, and social media analytics techniques can be used to support online political participation, e-democracy, political blogs and forums analysis, e-government service delivery, and process transparency and accountability (Chen 2009; Chen et al. 2007). For e-government applications, semantic information directory and ontological development (as exemplified below) can also be developed to better serve their target citizens.

Despite the significant transformational potential for BI&A in e-government research, there has been less academic research than, for example, e-commerce-related BI&A research. Egovernment research often involves researchers from political science and public policy. For example, Karpf (2009) analyzed the growth of the political blogosphere in the United States and found significant innovation of existing political institutions in adopting blogging platforms into their Web offerings. In his research, 2D blogspace mapping with composite rankings helped reveal the partisan makeup of the American political blogsphere. Yang and Callan (2009) demonstrated the value for ontology development for government services through their development of the OntoCop system, which works interactively with a user to organize and summarize online public comments from citizens.

Science and Technology

Many areas of science and technology (S&T) are reaping the benefits of high-throughput sensors and instruments, from astrophysics and oceanography, to genomics and environmental research. To facilitate information sharing and data analytics, the National Science Foundation (NSF) recently mandated that every project is required to provide a data management plan. Cyber-infrastructure, in particular, has become critical for supporting such data-sharing initiatives.

The 2012 NSF BIGDATA program solicitation is an obvious example of

the U.S. government funding agency's concerted efforts to promote big data analytics. The program aims to advance the core scientific and technological means of managing, analyzing, visualizing, and extracting useful information from large, diverse, distributed and heterogeneous data sets so as to accelerate the progress of scientific discovery and innovation; lead to new fields of inquiry that would not otherwise be possible; encourage the development of new data analytic tools and algorithms; facilitate scalable, accessible, and sustainable data infrastructure; increase understanding of human and social processes and interactions; and promote economic growth and improved health and quality of life.

Several S&T disciplines have already begun their journey toward big data analytics. For example, in biology, the NSF funded iPlant Collaborative is using cyberinfrastructure to support a community of researchers, educators, and students working in plant sciences. iPlant is intended to foster a new generation of biologists equipped to harness rapidly expanding computational techniques and growing data sets to address the grand challenges of plant biology. The iPlant data set is diverse and includes canonical or reference data, experimental data, simulation and model data, observational data, and other derived data. It also offers various open source data processing and analytics tools.

In astronomy, the Sloan Digital Sky Survey (SDSS) shows how computational methods and big data can support and facilitate sense making and decision making at both the macroscopic and the microscopic level in a rapidly growing and globalized research field. The SDSS is one of the most ambitious and influential surveys in the history of astronomy. Over its eight years of operation, it has obtained deep, multicolor images covering more than a quarter of the sky and created three-dimensional maps containing more than 930,000 galaxies and over 120,000 quasars. Continuing to gather data at a rate of 200 gigabytes per night, SDSS has amassed more than 140 terabytes of data. The international Large Hadron Collider (LHC) effort for high-energy physics is another example of big data, producing about 13 petabytes of data in a year (Brumfiel 2011).