

Про вивчення у ВНЗ системи аналізу даних на основі IDC`s Business Analytics Software Taxonomy

Актуальність вивчення у ВНЗ і ВВНЗ сучасних технологій аналітичної обробки інформації визначається тим, що використання сучасних інформаційно-аналітичних технологій подібне за ефектом до використання сучасних засобів виробництва чи озброєння і забезпечує перемогу у конкурентних змаганнях у різноманітних ділових сферах. Українські організації йдуть шляхом взаємодії або інтеграції зі світовими організаціями у різних сферах (економічна; політична; гуманітарна; інформаційна; охорони здоров'я населення; правоохоронна; оборона, ін.). Це вимагає впровадження або реінжинірингу, а також підтримки ділових інформаційно-аналітичних процесів в організаціях на основі загальноприйнятих у світі класифікацій (таксономій), що повинні знати студенти українських ВНЗ та ВВНЗ [1, 2].

Аналіз джерел знань про передові технології у сфері інформаційно-аналітичної обробки ділової інформації показує, що основні передові технології у цій сфері започатковані та реалізовані на практиці, як правило, у США.

Провідними приватними консалтинговими компаніями в ІТ-сфері у США є **International Data Corporation (IDC), Gartner, Forrester** [3,4,5]. Ці консалтингові компанії у своїх звітах визначають структуру, у тому числі порядок взаємодії інформаційно-аналітичних систем для менеджменту, економіки, фінансів, банківської справи, обліку, маркетингу, ін. Представництво консалтингової компанії IDC є в Україні [6].

Провідною державною консалтинговою організацією у США є **President's Council of Advisors on Science and Technology, PCAST** (Рада консультантів президента США з науки і технологій), яка входить до складу **Office of Science and Technology Policy, OSTP** (Офіс президента США з наукової і технологічної політики) [1]. Аналітичні консалтингові документи PCAST для Конгресу і Президента США готують провідні спеціалісти США з різних державних підсистем та з різних сфер консалтингу, науки і бізнесу США.

Аналіз звітів вказаних консалтингових груп та інших показує, що звіти IDC, Gartner, Forrester, PCAST доцільно використовувати як важливі джерела для визначення інформаційних технологій, які доцільно вивчати для уточнення змісту навчального процесу студентів українських ВНЗ та БВНЗ [3-7].

Також необхідно використовувати документи **AAAI (Association for the Advancement of Artificial Intelligence, formerly the American Association for Artificial Intelligence)** – Асоціації просування штучного інтелекту, спеціалізованих груп **ACM (Association for Computing Machinery)**, та інших [8-12].

Аналіз джерел знань по технологіям інформаційно-аналітичної обробки інформації на радянському і пострадянському просторі показує, що деякі важливі напрацювання не втратили свого значення і дозволяють краще зрозуміти переваги і недоліки сучасних технологій автоматизації інформаційно-аналітичної роботи, а також удосконалити їх. Див. для прикладу, [концепцію побудови ОГАС \(ЗДАС\) \(1960-ті роки\)](#), [концепцію ергатичного організму \(1970-ті роки\)](#) та розроблену на її основі [концепцію мережевого ергатичного організму \(2008 рік\)](#), яка підтримана багатьма вченими і практиками. Незалежно від назви, концепція мережевого ергатичного організму по суті є важливою концепцією реалізації сучасних великих мережевих ІТ систем (Google, Bing, інші). Всі відомі сучасні пострадянські технології аналітичної обробки інформації базуються чи узгоджені з Business Intelligence та Business Analytics та їх частковими складовими

(Data Mining, Text Mining, ін.), які засновані в США, у тому числі вихідцями з СРСР (П`ятецький-Шапіро – **Knowledge Discovery & Data Mining**, Сергій Брін – пошукові інформаційно-аналітичні алгоритми Google, тощо). Західні технології часто неточно позначають, для прикладу: **Data Mining** як “интеллектуальный анализ данных”, **Business Intelligence** як “бізнесова аналітика”.

Необхідно підкреслити, що “консервація” “енциклопедичної” термінології радянського часу (“АСУ”; ін.), яка у свій час була узгоджена з англійською термінологією і часто неточно перекладена (System Engineering як Системотехніка, Artificial Intelligence як “штучний інтелект”, Business Intelligence як бізнес-аналітика чи бізнесова аналітика, чи бізнес-розвідка; ін.), у деяких випадках гальмує вивчення і реалізацію важливих інформаційно-аналітичних hi-tech в Україні. Адже світ став глобальним, перейшов на рівень **Globalization 3.0** (“Глобалізація 3.0”), що потребує використання відповідної глобальної англійської термінології з урахуванням її бурхливого і суперечливого розвитку [2-13]. Відокремлення від цього процесу або замовчування і неявна протидія змінам приводить до закріплення відставання інформаційно-аналітичних технологій у деяких підсистемах.

Необхідно також зазначити, що всі провідні приватні компанії в Україні, насамперед компанії з іноземним капіталом – мають реалізовані передові інформаційно-аналітичні системи високого світового рівня, використовують сучасну англійську термінологію. У цих компаніях інформаційно-аналітичні процеси потребують постійної підтримки та подальшого удосконалення. Спеціалісти цих компаній, як правило, відвідують провідні консалтингові заходи IDC та ін.

Метою статті є визначення таксономії (класифікації) прикладних інформаційно-аналітичних технологій та відповідних програмних додатків для вивчення студентами у ВНЗ та ВВНЗ – на основі аналізу звітних документів провідних консалтингових організацій IDC, Gartner, Forrester, PCAST та ін.

Необхідно зазначити, що замість терміну «Інформатика» у англомовному просторі використовують, як правило, термін **Computer Science** – це потрібно враховувати при викладанні навчальної дисципліни «Інформатика» та зв'язаних навчальних дисциплін.

З аналізу документів PCAST та зв'язаних публікацій можна зробити наступні важливі висновки (скорочено) [1].

1. У зв'язку із активним поєднанням інформаційних технологій з мережевими технологіями, у останнє десятиліття замість терміну **IT (Information Technology)** часто використовують більш точний термін **NIT- Networking and Information Technology** (Мережеві та Інформаційні Технології), а дослідження і розвиток позначають **R&D – Research and Development** (Дослідження і Розвиток). У США на державному рівні фінансується програми у сфері **NIT R&D**, вартість яких обчислюється мільярдами доларів [14].

2. Орієнтовно з 2010 року визначено початок **Big Data Age** (Ери Великих Даних) або **Zettabyte Age** (Ери Зеттабайт, Zetta (ZB) -1021). Вказано, що за останні десятиліття NIT підвищили продуктивність праці у державі (США) більше ніж будь-який інший набір сил, адже у 90-х роках 20-го століття інформація стала надвеликим і глобальним ресурсом. Інформаційний ресурс став прирівнюватися до глобальних матеріальних ресурсів і став таким же значним джерелом переробки та отримання прибутків [1, 2, 14].

3. У грудні 2010 року, у звіті PCAST, «Report to the President And Congress «**Designing a Digital Future: Federally Funded Research and Development in Networking and Information Technology**» («Доповідь Конгресу і Президенту США «Конструювання Цифрового майбутнього: що фінансуються з федерального бюджету досліджень і розробок в області мережних

та інформаційних технологій») вказано, що **Data Mining** («Розкопка Даних») і **Machine Learning** («Машинне Навчання») полегшують «перетворення даних у знання, а знань у дії» і зазначено, що кожне федеральне агентство повинне мати свою **Big Data Strategy** (Стратегію Великих Даних) [2, 14].

При цьому необхідно враховувати, що **Data Mining** і **Machine Learning** є специфічним спрощеним позначенням більш широкого комплексу інформаційних hi-tech, частковими складовими **Business Intelligence** і використовуються у комплексі з іншими засобами **Business Intelligence (Data Warehouse, Data Marts, OLAP, Dashboards, Scorecard, Text Mining, Visual Mining, ін.)**.

У авторських публікаціях на початку 2009 року запропоновано реалізувати такі стратегії (концепції), а до початку 2010 року мав бути розробленим початковий варіант такої концепції, проте з причини недоліків науково-організаційної роботи ця робота була не реалізована. Якщо мовчати про подібні недоліки, то вони будуть закріплювати відставання у важливих сферах **MIT R&D**. Прикладом практичних недоліків є відсутність допустимого і необхідного обміну інформацією між різними сховищами даних у підсистемах РНБОУ, що приводить до резонансних порушень права в Україні.

Консалтингові компанії Gartner, Forrester, IDC, як правило, щороку видає аналітичні звіти про стан і розвиток ІТ-технологій у сфері бізнесу, визначає провідних вендорів (постачальників) відповідних програмних засобів та визначає структуру ІТ у сфері бізнесу. Важливою частиною цих звітів є специфічні графіки та схеми (таксономії), які наочно і узагальнено відображають зміст рекомендацій. Для прикладу, у компанії **Gartner** – це діаграми, подібні до [3-5, 2]:

- у компанії **Gartner** – це **Gartner Magic Quadrant for Business Intelligence and Analytics Platforms...**,
- у компанії **Forrester** – це **The Forrester Wave™: Agile Business Intelligence Platforms ...**,

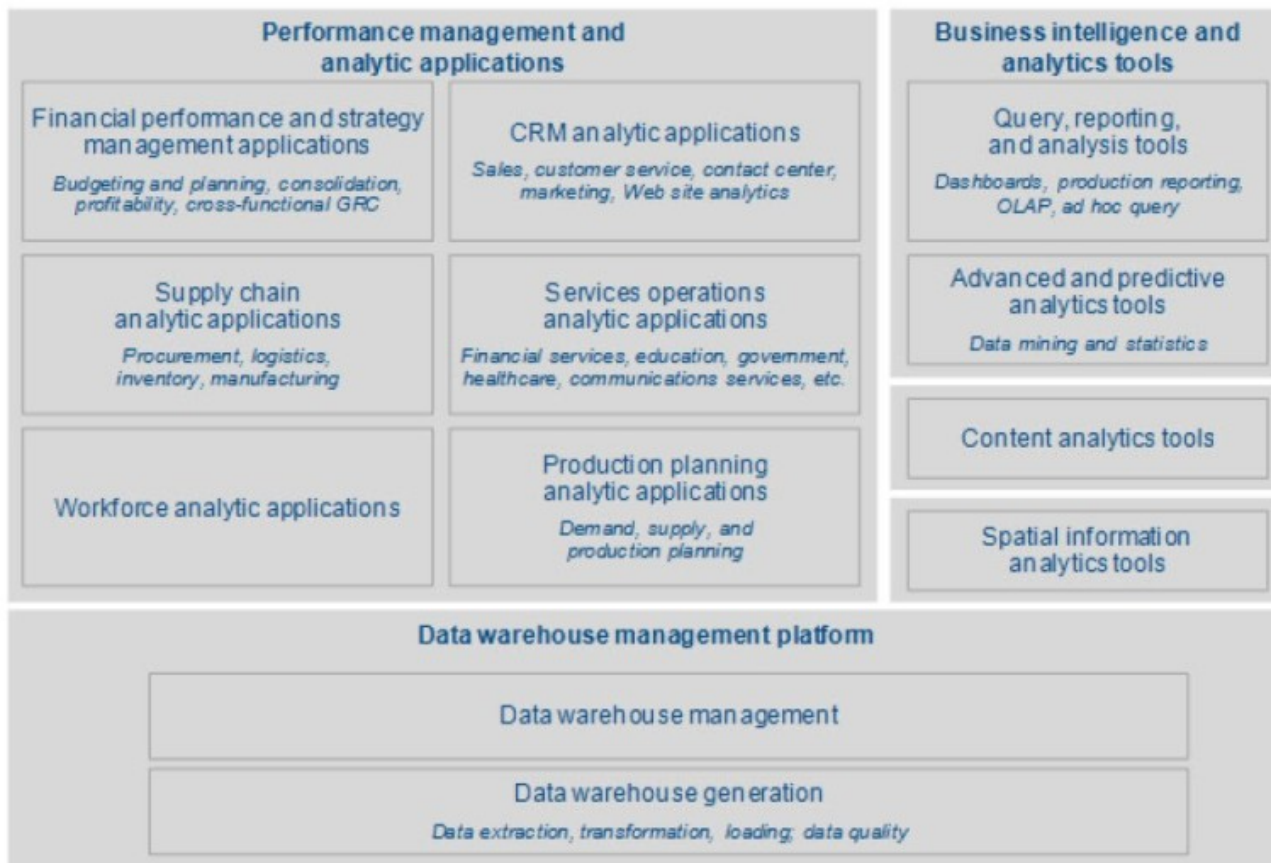
- у компанії IDC – це **IDC's Business Analytics Software Taxonomy ...** .

Аналіз доступних звітів провідних світових консалтингових компаній показує, що найкраще зрозуміти структуру системи сучасних програмних засобів для ділової аналітики (статистики) дозволяє **IDC's Business Analytics Taxonomy (у останній час позначають Business Analytics Software Market Taxonomy – таксономія (класифікація) програмного забезпечення для бізнес-аналітики)**, які оновлюється раз на кілька років чи щорічно.

Таксономію програмних засобів для ділової аналітики (**Business Analytics Taxonomy**) доцільно у різній мірі вивчати чи посылатися на неї під час вивчення особливостей інформаційних технологій і програмних засобів у навчальних дисциплінах, подібних до інформаційних систем і технологій, інформатики, статистики, систем підтримки прийняття рішень, веб-технологій, АСУ (SCADA), захист інформації, економічний аналіз, бухгалтерський облік, географічні інформаційні системи, ін. Доцільність такого вивчення підтверджена результатами анкетування і співбесід зі студентами очної і заочної форм навчання Житомирського військового інституту, а також Міжрегіональної академії управління персоналом.

Повні звіти IDC платні, тому нижче подано доступний в інтернеті варіант IDC Business Analytics Software Market Taxonomy 2014 року, а також варіант IDC Business Analytics Taxonomy, який перекладений на українську мову, з додатковою нумерацією.

IDC's Business Analytics Software Market Taxonomy, | 2014



Source: IDC, 2014

IDC`s Business Analytics Software Market Taxonomy (forbes.com)

На рисунку нижче показаний український переклад Круковського І.А. IDC`s Business Analytics Software Market Taxonomy з додатковою нумерацією її основних компонентів, що дозволяє більш зручно пояснювати зміст її підсистем та елементів.

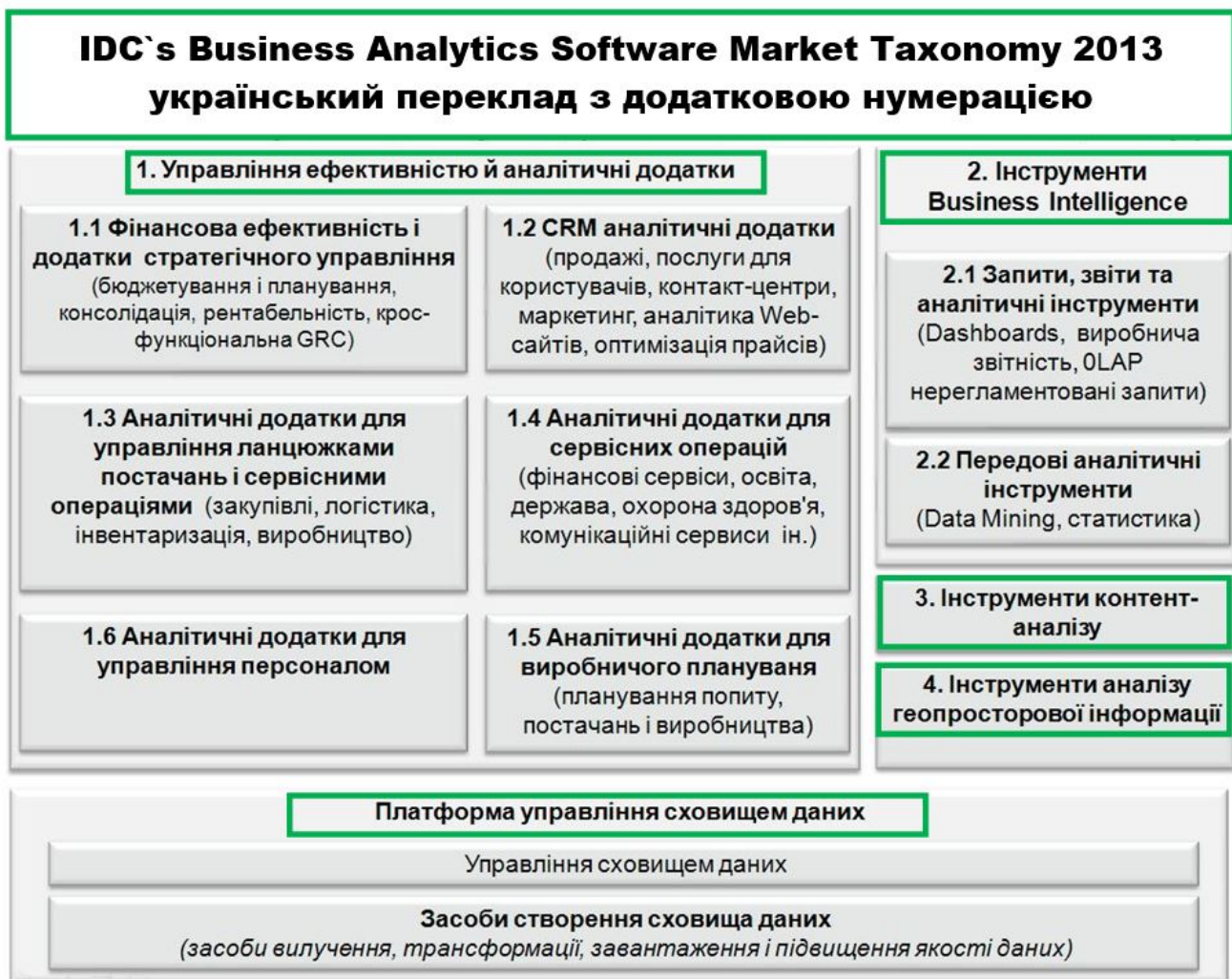
Основні підсистеми програмних засобів для ділової аналітики згідно IDC`s Business Analytics Software Market Taxonomy є такі:

- програмні додатки для аналітики й управління ефективністю організацію – **Performance Management and**

Analytic Applications;

- аналітичні інструменти для аналізу числової інформації, як правило у табличній електронній формі – **Business Intelligence;**
- інструменти контент-аналізу на основі Text Mining – **Content analytics tools;**
- аналітичні інструменти для просторової інформації на основі Geographic Information System (GIS) – **Spatial information analytics tools;**
- платформа управління сховищем даних організації – **Data warehouse management.**

Більш детальне викладання змісту компонентів IDC`s Business Analytics Software Market Taxonomy виходить за рамки цієї статті.



український переклад Круковського І.А.

Для часткових сфер інформаційно-аналітичної роботи (прикладної статистики) у державних сферах доцільно удосконалити існуючу таксономію узгоджено з Business Analytics у підсистемі Performance Management and Analytic Applications та інтегрувати її із засобами BI – Business Intelligence, GIS, Text Mining, ін. Як вказано вище, прикладом і результатом практичних недоліків у цій сфері є відсутність допустимого і необхідного обміну інформацією між різними сховищами даних у підсистемах РНБОУ, що приводить до резонансних порушень права в Україні.

Систему програмних засобів для часткових сфер (економіка, менеджмент, фінанси, банківська справа та страхування, облік і оподаткування, медицина, психологія, соціологія, маркетинг, політологія, правознавство, національна безпека і оборона, ін.) можна будувати як модифіковану систему систем підтримки прийняття рішень на основі Business Intelligence (DSS BI 2.0), яка передбачає реалізацію концепції Business Intelligence+KMS (**Knowledge Management System**) або DSS BI 3.0 (Cloud Computing) [2]. Збереження інформації та її аналіз з використанням “хмарних технологій” (**Cloud Computing**) на практиці у багатьох випадків дуже небезпечно чи неприпустимо з причин небезпеки для конфіденційної інформації. Неконфіденційна інформація організації також є об’єктом **Competition Intelligence** (Конкурентної розвідки) і з неї можуть бути сформовані конфіденційні знання про діяльність організації за допомогою засобів Business Intelligence (**OLAP, Data Mining, Visual Mining, Text Mining, Dashboard, Scorecard, і т.п.**) Тому під питанням також є можливість зберігання неконфіденційної інформації про організацію у чужій чи публічній IT-“хмарі”

Особливості побудови DSS BI 2.0 описані у статті “Узагальнена архітектура системи підтримки прийняття рішень на основі Business Intelligence у розширеному тлумаченні” (адреса на сайті НБУВ кілька разів змінювалася).

Особливості концепції Business Intelligence+KMS (Knowledge Management System) та її часткових складових і прикладних застосувань описані на головній сторінці сайту Business Intelligence+KMS та у системі зв'язаних наукових публікацій [2, 17-23].

Зазначимо, що Business Intelligence+KMS відображає підходи, які реалізовані на практиці у всіх сучасних аналітичних системах, проте під різними маркетинговими назвами. Для прикладу: подібно побудована основа веб-системи Google Analytics; у необхідному форматі для застосування засобів Business Intelligence без операцій ETL (Extract, Transform, Load) почали видавати лог файл у панелі управління хостингом C-panel; ін.

Висновки і перспективи подальших досліджень

1. За результатами проведеного навчально-наукового і практичного дослідження встановлено, що найбільш доцільно вивчати технології аналітичної обробки інформації на основі IDC`s Business Analytics Software Taxonomy з урахуванням "Gartner Magic Quadrant for Business Intelligence and Analytics Platforms...", "The Forrester Wave™: Agile Business Intelligence Platforms...". Їх назви час від часу дещо змінюються, остання доступна таксономія IDC`s Business Analytics Software Taxonomy 2015.

2. У процес навчання студентів у ВНЗ та студентів і курсантів у ВВНЗ доцільно включити завдання перекладу і публікації на спеціальному сайті щорічних консалтингових документів, подібних до публікацій IDC, Gartner, Forrester, PCAST, AAAI, ACM, ін. Це підвищує рівень їх підготовки, що підтверджено у підготовці студентів навчально-наукового підрозділу ЖВІ, у МАУП та відображено у публікаціях більше ніж 90-та студентів на цьому сайті.

3. Знання і використання звітних документів провідних світових консалтингових груп (IDC, Gartner, Forrester, PCAST, AAAI, ін.) забезпечує підтримання рівня компетентності спеціалістів протягом всього життя.

4. Кожен звітний документ провідних світових консалтингових груп має вартість більше ніж 1000 \$, тому у більшості випадків можна ознайомитися лише з їх безкоштовними відкритими версіями, у яких не гарантована повна достовірність і повнота консалтингової інформації. Вивчення і дослідження таксономій приводить до їх доповнення й уточнення.

5. Кожна велика і мала організація у державній сфері та у сфері бізнесу потребує своєї Концепції Big Data (обробки табличної і текстової інформації), яка повинна бути узгоджена з IDC Business Analytics Таксоному, відповідати своїй частковій проблемній області роботи та бути узгодженою з концепціями, технологіями і програмними засобами у взаємодіючих організаціях.

Перспективою подальших досліджень є опис архітектури удосконалених систем підтримки прийняття рішень на основі Business Intelligence (DSS BI), які уточнюють і удосконалюють існуючу таксономію систем підтримки прийняття рішень (DSS – Decision Support System). Система узагальнених архітектур DSS/BI представлена на початку 2016 року у [22].

ЛІТЕРАТУРА:

1. President's Council of Advisors on Science and Technology [Електронний ресурс]. – Режим доступу : <https://www.whitehouse.gov/administration/eop/ostp/pcast> .
2. Круковський І.А. Business Intelligence+KMS. DSS-BI.com.ua. [Електронний ресурс]. – Режим доступу : <http://dss-bi.com.ua/>
3. IDC. Analyze the Future. [Електронний ресурс]. – Режим доступу : <http://www.idc.com> .
4. Gartner [Електронний ресурс]. – Режим доступу

:<http://www.gartner.com> .

5. Forrester [Электронный ресурс]. – Режим доступа :<http://www.forrester.com>.

6. IDC Ukraine. [Электронный ресурс]. – Режим доступа :: <http://idcukraine.com/ru/> .

7. PCAST (President's Council of Advisors on Science and Technology). Report to the President And Congress «Designing a Digital Future: Federally Funded Research and Development in Networking and Information Technology». December 16, 2010. [Электронный ресурс]. – Режим доступа :<http://www.whitehouse.gov/administration/eop/ostp/pcast>.

8. AAAI (Association for the Advancement of Artificial Intelligence, formerly the American Association for Artificial Intelligence), AAAI DIGITAL LIBRARY. [Электронный ресурс]. – Режим доступа :<http://www.aaai.org/Library/library.php> .

9. ACM SIGKDD (Association for Computing Machinery, Special Interest Group on Knowledge Discovery in Data), ACM Digital Library, SIGKDD. [Электронный ресурс]. – Режим доступа :<http://portal.acm.org/sig.cfm?id=SP936> .

10. D.J. Power. Ask Dan! about DSS, Brief History of Decision Support Systems (version4.1). Editor, DSSResources.COM. – Режим доступа : <http://dssresources.com/history/dsshhistory.html> .

11. US Army ESCC (The U.S. Army Enterprise Solutions Competency Center, Army Business Intelligence Competency Center). Business Intelligence Reference Guide. – Режим доступа: <http://escs.army.mil> .

12. The World is Flat (ISBN 1-59397-668-2), Thomas L. Friedman, pg 421

13. The Networking and Information Technology Research and Development (NITRD) Program. – Режим доступа: <https://www.nitrd.gov/> .

14. PCAST (President's Council of Advisors on Science and Technology) // Report to the President And Congress «Designing a Digital Future: Federally Funded Research and Development in Networking and Information Technology» (December 16, 2010) [Электронный ресурс]. – Режим доступа : <http://www.whitehouse.gov/sites/default/files/microsites/ostp/>

pcast-nitrd-report-2010.pdf .

15. Круковський І.А. Нотатки про Business Intelligence+. DSS-BI.com.ua. – Режим доступу : <http://dss-bi.com.ua/WP/> .

16. Круковський І.А. Узагальнена архітектура системи підтримки прийняття рішень на основі Business Intelligence у розширеному тлумаченні / І.А. Круковський // Вісник ЖДТУ. – 2010. – Вип. 2 (53). – С. 103–111.

17. Круковський І.А. Удосконалені вимоги до реалізації OLAP у DSS для часткових проблемних областей інформаційно-аналітичної роботи : військ.-техн. зб. / І.А. Круковський // Академія сухопутних військ. – 2010. – Вип. 3. – С. 26–32.

18. Круковський І.А. Архітектура експертної системи з розширеним виведенням на трикомпонентній гібридній моделі подання знань / І.А. Круковський // зб. наук. праць ВІТІ НТУ України «КПІ». – 2009. – Вип. 3. – С. 20–24.

19. Валюх А.І. Експертна система, узгоджена з Business Intelligence 2.0 / А.І. Валюх, І.А. Круковський, В.Л. Сімаков // Вісник ЖДТУ. – Житомир, 2011. – Вип. 2 (57). – С. 53–62 [Електронний ресурс]. – Режим доступу : http://archive.nbuv.gov.ua/portal/natural/Vzhdtu/2011_2/8.pdf .

20. Круковський І.А. Проблемні питання розробки і реалізації Geospatial Business Intelligence / І.А. Круковський // Геоінформаційні системи у військових задачах : ІІ наук.-техн. семінар 21–22 січн. 2011 року. – Львів : Академія Сухопутних військ, 2011. – С. 117–125 [Електронний ресурс]. – Режим доступу : <http://www.asv.gov.ua/content/nauka/gis.pdf> .

21. Круковський І.А. Проблемні питання використання і розвитку засобів Social Media Analytics, їх інтеграції з Business Intelligence та з елементами ГІС – на прикладі платформи SemanticForce / І.А. Круковський, В.Л. Гаврилюк, Б.А. Хомів // “ІVСічневі ГІСи”: Інтелектуальна оборона” (науково-практичний форум) / Академія Сухопутних військ імені гетьмана Петра Сагайдачного: Львів, 22-24 січня 2013 р. – С. 42-45.

22. Google+, Sistem Plus. [Електронний ресурс]. – Режим доступу : <https://plus.google.com/+Sistemplk7> .

23. Система+. [Електронний ресурс]. – Режим доступу :

<http://dss-bi.com.ua/System/> .

Автор публікації: Ігор Круковський, system.k7@gmail.com.

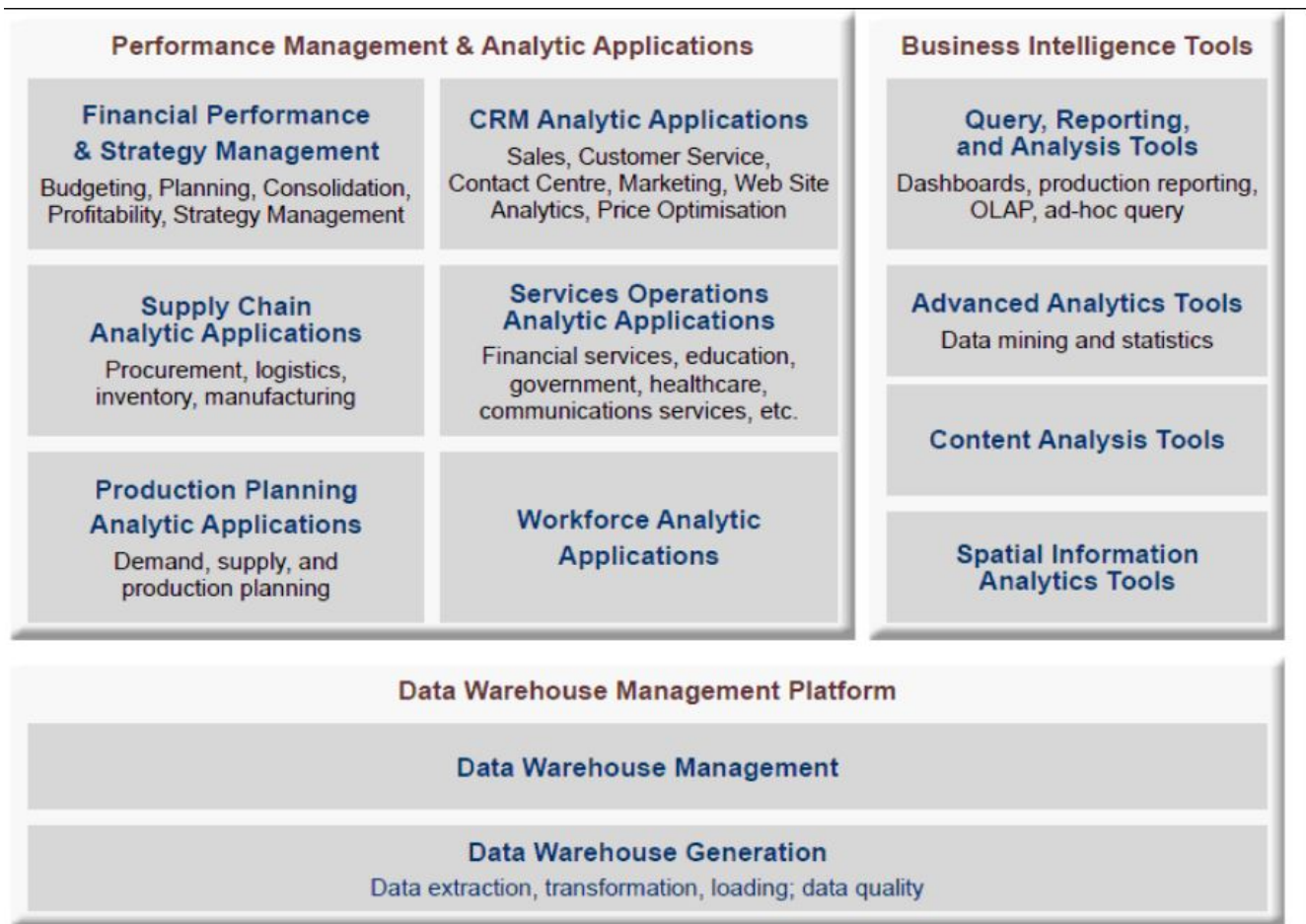
Приклади таксономій для ретроспективного аналізу:

IDC's Business Analytics Software Taxonomy, 2008



Source: IDC, 2008

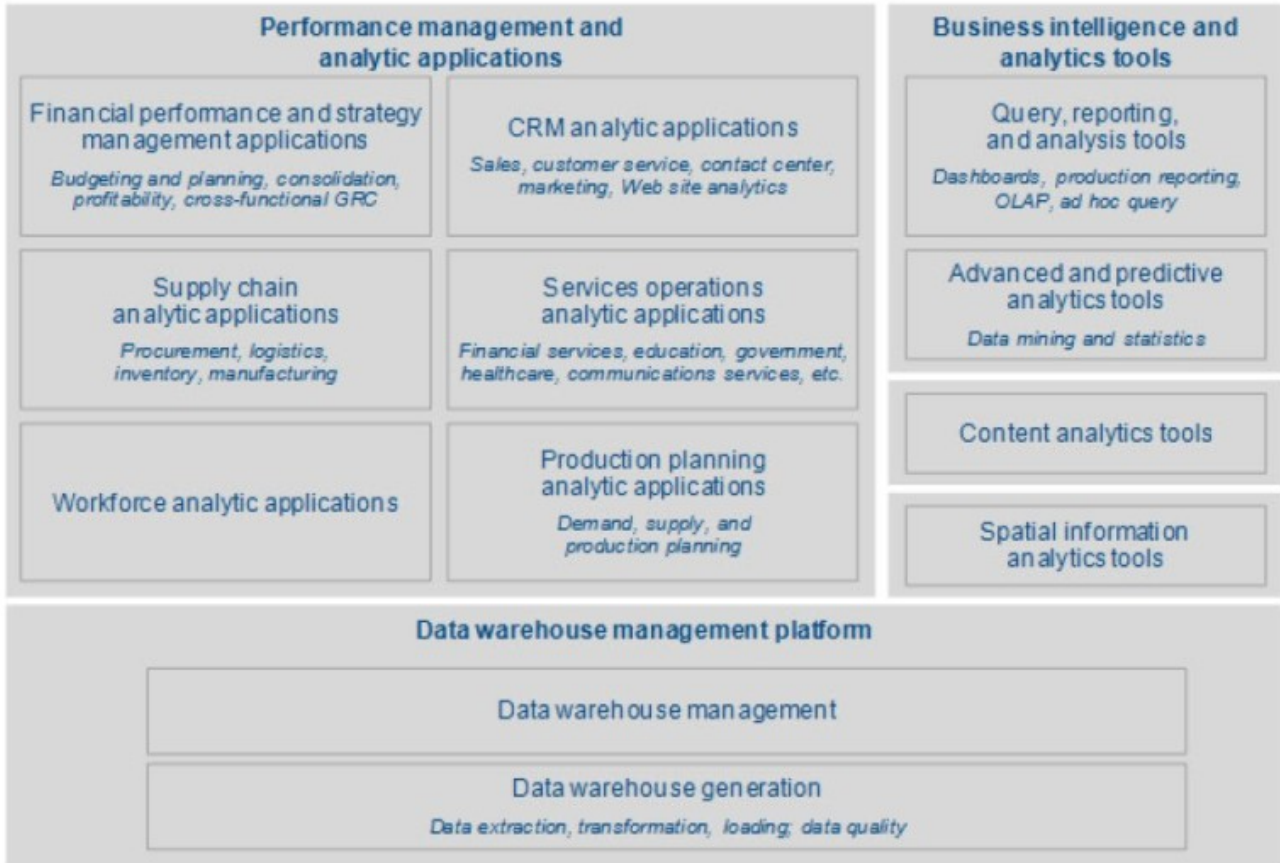
IDC`s Business Analytics Software Taxonomy 2008



Source: IDC, 2011

IDC's Business Analytics Software Taxonomy 2011

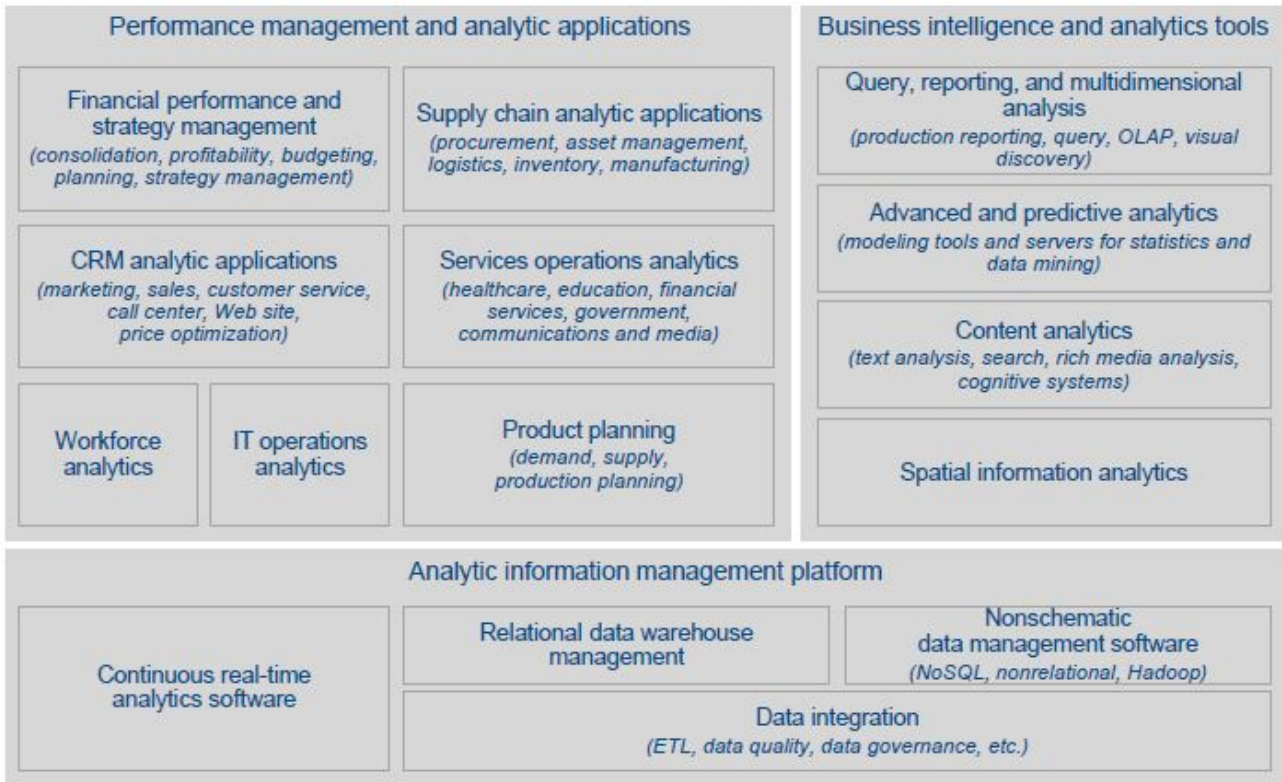
IDC's Business Analytics Software Market Taxonomy, | 2014



Source: IDC, 2014

IDC's Business Analytics Software Market Taxonomy 2014

IDC's Business Analytics Software Taxonomy, 2015



Source: IDC, 2015

IDC's Business Analytics Software Taxonomy 2015