

Поява и развитие на складовете от знания

Латинка Тодоранова

The creation and development of knowledge warehouses

Latinka Todoranova

Abstract

The role of information technology in today's organizations is increasing. They are applied in the processing of large volumes of data in order to provide information to support effective decision making. However, only part of the needed information exists on computers. Important information exists in the heads of the employees. It is needed infrastructure to capture, storage and distribution this knowledge. The primary goal of a knowledge warehouse is to provide those opportunities.

Keywords: knowledge, knowledge management, knowledge warehouse.

Основна цел на развитието на информационните системи е осигуряването на възможност за подобряване качеството на решенията – много важен елемент във функционирането на всяка съвременна организация. Поставената цел може да се постигне чрез съхраняването, ефективното управление и използването на опита и на знанията в организацията. Много компании прилагат технологично-базирани инструменти за организиране на вътрешни иновативни дейности, основаващи се на знание. Те осъзнават, че знанията са стратегически ресурс и източник на възможности за подобряване на тяхната работа. Въпреки това обаче се отделя много малко внимание на това, как знанието може да бъде ефективно използвано за обогатяване на компетенциите на организациите. В същото време организациите са богати на информация и бедни на знания – независимо от нарастващата роля на съвременните информационни технологии. Една от причините за това е липсата на системни методи за идентифициране, представяне, съхраняване, преобразуване и трансфер на двата типа знание – явно (explicit) и неявно (tacit)¹. А всяка организация в днешно време трябва да е по-съсредоточена върху огромните ресурси от знания. За целите на настоящето изследване е необходимо първо да се разгледа общоприетото разграничаване на понятията данни, информация и знание.

Данните са форма на представяне на факти. В зависимост от различните дисциплини има различни дефиниции за данни. От гледна точка на форматите данните включват текст, мултимедия, временни редове, мрежов трафик и др. От гледна точка на структурата – структурирани, полуструктурирани, неструктурирани. Данните, на които не е приложена никаква физическа или аналитична обработка, могат да се нарекат „сурови”. Следователно характеристиките на „суровите” данни са специфичност, изолираност, хаотичност. След осъществяване на достъпа до тези данни са необходими подходящи обработки за превръщането им в абстрактна и универсална информация.

Информация са всички данни, които са преработени от гледна точка на определени цели, т.е. това са данни, интерпретирани от човек чрез използване на определени средства.

Знанията по своята същност са информация за свойствата на обектите и явленията от дадена област и връзките между тях, като знанието може да се разглежда в две насоки. От една страна, като резултат от обработка, интерпретиране, селектиране и трансформиране на данни, а от друга, като убеждения, намерения, гледни точки и др., отразяващи определена позиция, т.е. знанието винаги съществува за „определени цели”.

В изследването разграничаваме две категории знание – „сурово знание” и „експертно знание”. Като за „сурово знание” приемаме знанието, което е получено в резултат на процеса

¹ Nonaka, I., The Knowledge-Creating Company, <<http://midwestkm.com/wp-content/uploads/2013/09/The-Knowledge-Creating-Company-Nonaka.pdf>>, (20.06.2014).

на извличане на знания (*Data mining, DM*), следователно това е информация, която е анализирана чрез използване на алгоритми, техники и средства за извличане на знания. Прилагането на алгоритмите за извличане на знания представлява аналитичен процес, при който се използват изкуствен интелект, статистика, оптимизация и други математически алгоритми за извършване на по-задълбочен анализ. Преобразуването на данните в „сурово знание“ обхваща следните етапи: 1) предварителна обработка – преобразуване на данни в информация; 2) анализ – преобразуване на информацията в сурово знание.

Характеристиките на „суровото знание“ са:

- Необработеност – без по-нататъшно пречистване „суровото знание“ съдържа много излишества, едностранни или дори погрешни знания.
- Разнообразие – знанията трябва да се представят чрез определен модел за отнасяне към вземането на решения. Има много форми за представяне на „сурово знание“, – асоциативни правила, дърво на решенията, невронни мрежи, вероятностна карта, клъстери, формули и др. Някои представяния са лесни за разбиране (например дърво на решенията), докато други са трудни за интерпретиране (като невронните мрежи).
- Наличие на определен източник – това са резултати от анализи.
- Време на получаване – „суровото знание“ се извлича в определен момент от време. Следователно може да има конфликти между знанието, генерирано в различни периоди.
- Частична лекота на използване – съществува възможност за подпомагане дейността на организацията.

Основните задачи на извличането на „сурово знание“ са охарактеризиране, разграничаване, релевантност, класификация, клъстеризация, откриване на необичайни зависимости, анализ за отклонения, намиране на сходства и др. Технологиите, използвани при този процес, включват статистически анализи, оптимизация, машинно обучение, визуализация, складиране на данни и др. Могат да се разграничат следните видове представяне на „сурово знание“: правила, класификационни маркери, тагове за клъстеризация и др.

„Експертното знание“ е знанието на специалистите в определена област. То зависи от индивидуалните умения, получени по време на обучения, по време на ежедневната работа, както и при преодоляването на критични ситуации. Това знание има строги граници при определяне на неговите значения. В рамките на даден обект то е необходимо и полезно знание, а извън тях вече не е приложимо в същата степен. Качеството на „експертното знание“ зависи до известна степен от професионалния опит на служителите. Към „експертното знание“ можем да отнесем факти и правила, които нямат нужда от доказване.

В областта на управлението на знания всяко знание, използвано за подпомагане дейностите на организацията, се разглежда като интелектуален капитал и по-точно това са: човешко знание; знание, вградено в организационните процеси и услуги; взаимоотношенията в организацията – връзки между служители, работещи в организацията и връзки между организацията и нейната среда. В съвременните условия обаче организациите работят в много бързо променяща се среда и са под постоянен натиск от страна на конкуренти, доставчици, партньори, клиенти, като в същото време са изправени пред опасността да загубят знания, когато служител напусне организацията. В този случай идентифицирането, съхраняването и използването на знания е най-важната дейност, която поддържа интелектуалния им капитал в актуално състояние. Създаването на интелектуален капитал, от своя страна, е свързано с два основни проблема:

- служителите не желаят да споделят своето знание – тези, които напускат организацията, вземат своите знания, опит и умения със себе си.
- новопостъпилите служители придобиват необходимите знания и умения за значителен период от време.

За преодоляване на тези проблеми е необходимо създаването на вътрешна култура за подпомагане споделянето на знания. Освен това трябва да се осигури съхраняването на

знанието на служителите. Тази цел може да се реализира по различни начини, като се започне от насърчаване на комуникацията между служителите и се завърши със създаването на бази от знания и склад от знания – това са части от процеса на преобразуване на неявното знание в явно. Същността на този подход е да се използват наличните средства и техники, за да се направи процесът на придобиване на нови знания и умения по-лесен за новоназначените служители. Следователно, за да развие интелектуалния си капитал всяка организация трябва да:

1. Идентифицира и възприема интелектуални ценности, които са в обкръжаващата я среда, както и вътре в самата организация.
2. Определи дали идентифицираните интелектуални ценности са достатъчни за постигане на предварително определени цели, за развиване на работни процеси и повишаване на нейната ефективност.
3. Създаде допълнителна стойност от своя интелектуален капитал чрез вземане на по-рационални решения и предприемане на съответни действия.

Именно поддържането на потока от знание е средството за генериране на нова допълнителна стойност от интелектуалния капитал. Това ще доведе до подобряване на функционирането на организациите в бъдеще. Следователно организациите, с използването на подходящите решения, могат много бързо да осигурят съвместна работа, сътрудничество и прозрачност. Поради тази причина сме изследвали появата и развитието на складовете от знания. Те се свързват с процеса на откриване и извличане на знания от неструктурирана информация. Тази информация представлява съвкупност от величини, които не могат да се представят в числов формат. Такива са голяма част от знанията на експертите, които са от съществено значение при постигане на стратегическите, тактическите и оперативните цели на организациите. А откриването и извличането на знания е изключително важен процес, защото от качеството на знанията се определя в голяма степен качеството на работа на организациите.

Съществува обаче фундаментална разлика между възможностите, които предоставят базите от знания (БЗ) и базите от данни (респективно складовете от данни и складовете от знания). При базите от данни може да се извлича само информация, представена в явен вид в базата, докато при базите от знания, в резултат на разсъждения се генерира нова информация, неприсъстваща в явен вид в базата². Наличието на база от знания и съответен механизъм за извод позволява реализирането на символни изчисления. От своя страна, символните системи се самообучават и адаптират към новопостъпилата информация. Символните системи, и по-точно експертните системи, предлагат и обяснения за полученото решение. Освен това, за разлика от данните в базите от данни, между отделните знания в БЗ съществуват връзки, които определят структурата на знанията. Следователно интеграцията между системите, използващи бази от данни и тези, използващи бази от знания води до много по-ефективни приложения.

Още през 2002 г. Nemati, Steiger, Iyer, Herschel³ и Dymond⁴ разглеждат концепцията за складовете от знания като развитие на складовете от данни в две независими разработки. Няколко години по-късно – през 2009 г., Lee⁵ също анализира възможността за комбиниране на извличането и управлението на знания със склад от знания. Определението, което той дава, е следното: складът от знания е „информационно хранилище“, в което компонентите от знание са организирани и съхранени за повторно използване⁶. Според него целта на склада от знания е да осигури на вземащите решения платформа за интелигентни анализи, която да поддържа всички фази на процеса по управление на знания. Следователно потребителите на склада от

² Атанасова, Т., Интелигентни компютърни системи, Унив. изд. „Наука и икономика“, Варна, 2011.

³ Виж: Nemati, H., David Steiger, Lakshmi Iyer, Richard Hershel, Knowledge warehouse: an architectural integration of knowledge management, decision support, artificial intelligence and data warehousing, Elsevier Science, 2002.

⁴ Dymond, A., The Knowledge Warehouse: The Next Step Beyond the Data Warehouse, SUGI 27 Data Warehousing and Enterprise Solutions, Paper 144-27, 2002.

⁵ Lee, Ming-Chang, The Combination of Knowledge Management and Data Mining with Knowledge Warehouse, International Journal of Advancements in Computing Technology, Volume 1, 2009.

⁶ Пак там.

знания са вземащите решения. Чрез неговото добавяне и включването на изкуствен интелект се разширяват когнитивните им възможности при преобразуването на неявно (*tacit*) в явно (*explicit*) знание. Явното знание може лесно да се програмира, докато неявното знание трудно може да се формулира, тъй като то включва научна и технологична експертиза, личен опит, умения, интуиция. В практиката много често опитите за преобразуване на неявното знание в явно водят до сериозни проблеми⁷. Явното знание се интегрира в анализите с цел откриване на нови модели и връзки, разбиране на новото знание чрез аналогии и обяснения. Въз основа на изложените факти могат да се обобщят основните изисквания към склада от знания, отнасящи се до неговите възможности за:

- ефикасно генериране, съхраняване, извеждане и управление на знания в различни форми;
- съхраняване, изпълнение и управление на аналитични задачи и подпомагащите ги технологии;
- валидиране на моделите и разкриване на връзката на новите знания с поставените цели.

Най-общо складът от знания осигурява същите услуги за знанията, каквито осигурява складът от данни за данните. Например, ако знанията са съхранени под формата на модели и отделни примери към моделите, то складът от знания трябва да съхранява, извлича и манипулира тези модели. Освен това всеки пример трябва да е логически свързан със съответния модел, за да се улесни генерирането на допълнителни примери от един и същ модел, както и да се подобри ефикасността на съхраняването и извличането им. Важно е да се отбележи обаче, че знанията, организирани в склада, подлежат на анализи, които се различават от тези върху данните. При анализите на знанията се включват различни технологии като: невронни мрежи, продукционни правила, генетични алгоритми и др. Всяка от тези технологии има свои собствени изисквания по отношение на входните параметри; параметрите за работа на избраната технология (например броя на слоевете в невронната мрежа); изходния формат. Използването на склада от знания и средствата от изкуствения интелект за манипулирането му дават възможност за генериране на обяснения, извеждане на аргументи, които са част от фактите и които показват дали даден хипотетичен модел подпомага или не поставената цел.

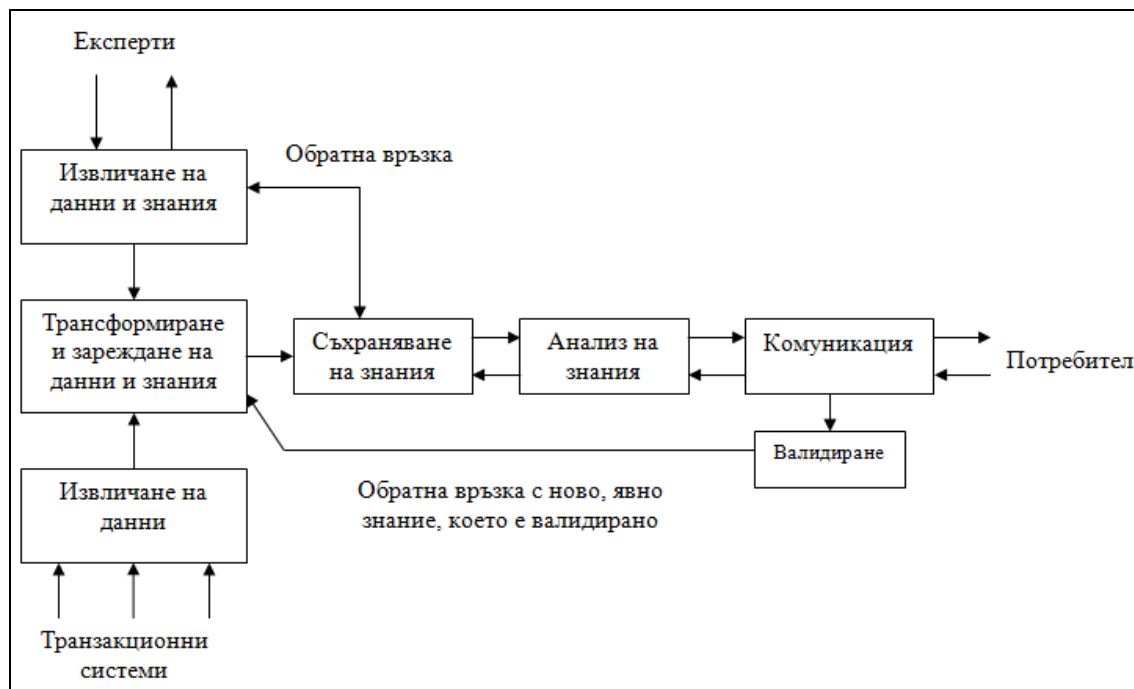
Nemati, Steiger, Iyer и Herschel предлагат архитектура на склад от знания, представена на фигура 1.

Предложената архитектура се състои от шест основни компонента: модул за извличане на данни и знания; два цикъла за обратна връзка; модул за трансформиране и зареждане на данни и знания; склад от знания (модул за съхраняване на знания); модул за анализ на знания и модул за комуникация. Първият от предложените компоненти е отговорен за преобразуването на неявното знание в явно. Той включва потребителски интерфейс, който подпомага добавянето и актуализирането на знания. Между модула за извличане на данни и знания и този за съхраняването им има двупосочна връзка, посредством модула за зареждане на знания. Тази връзка дава възможност за незабавно разпространяване на знанията от един потребител към друг, за представяне на актуални списъци с „какво-ако“ сценарии. Другата връзка – между модула за трансформиране и зареждане на знания и модула за комуникация, осигурява съхраняването на ново, валидирано знание, което е генерирано в системата.

Следващият компонент в разглежданата архитектура – модул за трансформиране и зареждане на данни и знания има функции, подобни на тези на средствата за извличане, трансформиране и зареждане (*Extract, transform, load, ETL*) на данни при складовете от данни. По отношение на модула за съхраняване на знания авторите предлагат използването на обектноориентирана система за управление на бази от знания (СУБЗ), интегрираща база от знания, база от модели и аналитични задачи. Тя включва голям набор от обекти от знания – числови данни, текст, проверени модели, метамоделите, както и средствата за манипулирането

⁷ Hansen, M., Nitin Nohria, Thomas Tierney, What's Your Strategy for Managing Knowledge?, <<http://consulting-ideas.com/wp-content/uploads/Whats-your-strat-art.pdf>>, (12.03.2014).

им. В следващия компонент – модул за анализ на знанията, е включен много важен елемент – контролер на задачите, който действа като „виртуален” потребител и по този начин не се налага реалният потребител да познава в детайли технологиите. Последният компонент – модулът за комуникация, управлява цялата връзка между СУБЗ и потребителския интерфейс. Той включва шест подмодула – когнитолог, „какво-ако” интерфейс, обработка на заявки, представяне на резултати, онлайн помощ, потребителски интерфейс.



Фиг. 1 Архитектура на склад от знания

При проектирания от Dymond⁸ склад от знания структурата, в която се съхраняват знанията, е база от знания, конструирана като дърво с обекти във възлите. Обектите съхраняват данните в атрибутите, а програмният код – в методите. Използването на обектноориентиран подход е заложено и при реализирането на архитектурата, предложена от Nemat, Steiger, Iyer, Hershel, представена по-горе в изложението. Според предложението от Dymond модел за въвеждането на знанията от базите от знания в склада от знания е необходима допълнителна работа от страна на специалисти, които провеждат обектноориентирани анализи и проектиране. Освен това могат да се използват и автоматизирани техники, с помощта на които знанията да се преобразуват във форма, която дава възможност за тяхното зареждане в склада от знания. Базите от знания, обектите, атрибутите и методите сами по себе си се разглеждат като компоненти, които могат да се реорганизируют в нови структури. Самите атрибути и методи в базите от знания са достъпни чрез изпълнението на заявки.

През 2009 г. Lee⁹ анализира възможността за комбиниране на извличането и управлението на знания със склад от знания. Той разглежда склада от знания като вид база от данни, управлявана чрез СУБЗ и самата база от знания. Под СУБЗ се има предвид софтуерно приложение, което управлява БЗ, подобно на СУБД. За разлика от БД, които съхраняват записи, предложената БЗ съхранява мрежа от обекти и компоненти, капсулирани като данни и методи. Според него, за разлика от склада от данни, складът от знания е комбинация от постоянни и променливи обекти и компоненти и освен данни съхранява информация и знания,

⁸ Dymond, A., The Knowledge Warehouse: The Next Step Beyond the Data Warehouse, SUGI 27 Data Warehousing and Enterprise Solutions, Paper 144-27, 2002.

⁹ Lee, Ming-Chang, The Combination of Knowledge Management and Data Mining with Knowledge Warehouse, International Journal of Advancements in Computing Technology, Volume 1, 2009.

капсулирани в атрибути и методи. В този контекст Lee представя склада от знания като включващ компоненти, разположени в шест отделни слоя – въвеждане на знания; действия със знания, съхраняване на данни, сървър за приложения, приложна система, потребителски интерфейс.

Разгледаните подходи имат общи елементи, като авторите на първо място извеждат необходимостта от откриването и организирането на знанията в бази от знания, съобразно някои от най-често използваните модели за представяне на знания – продукционни правила, семантични мрежи, фрейми. Трябва да се има предвид, че сложните задачи изискват различни типове знания, както и различни модели за представяне на знанията с цел улесняване на разсъжденията. Но все още никой от разгледаните модели не е в състояние напълно да представи реалния свят, както и да симулира човешките разсъждения. Поради тази причина в много области се налага използването на повече от един модел за представяне на знанията. След като знанията са налични в информационните системи, те могат да бъдат извлечени, трансформирани и заредени в склад от знания.

В заключение можем да обобщим, че предстои по-широкото използване на складовете от знания като средство, което не само улеснява улавянето и кодирането на знания, но и подпомага извличането и споделянето на знания в организацията. А управлението на знанията осигурява възможност за по-ефективно използване на интелектуалния капитал в организацията и до подобряване качеството на вземаните решения.

Използвана литература:

1. Атанасова,Т., Интелигентни компютърни системи, Унив. изд. „Наука и икономика”, Варна, 2011
2. Dymond, A., The Knowledge Warehouse: The Next Step Beyond the Data Warehouse, SUGI 27 Data Warehousing and Enterprise Solutions, Paper 144-27, 2002
3. Hansen, M., Nitin Nohria, Thomas Tierney, What’s Your Strategy for Managing Knowledge?, <<http://consulting-ideas.com/wp-content/uploads/Whats-your-strat-art.pdf>>, (12.03.2014)
4. Lee, Ming-Chang, The Combination of Knowledge Management and Data Mining with Knowledge Warehouse, International Journal of Advancements in Computing Technology, Volume 1, 2009
5. Nemati, H., David Steiger, Lakshmi Iyer, Richard Hershel, Knowledge warehouse: an architectural integration of knowledge management, decision support, artificial intelligence and data warehousing, Elsevier Science, 2002
6. Nonaka, I., The Knowledge-Creating Company, <<http://midwestkm.com/wp-content/uploads/2013/09/The-Knowledge-Creating-Company-Nonaka.pdf>>, (20.06.2014)

За контакти:

Ас. д-р Латинка Тодоранова
Икономически университет – Варна
E-mail: todoranova@ue-varna.bg