

# Zralostní modely pro nasazení sémantických technologií ve firemním prostředí

Marek Nekvasil

Doktorand oboru informatika

**Abstrakt.** Sémantické technologie jsou v posledních letech zahrnovány stále do širších oblastí aplikačního nasazení a jejich působnost se neustále rozšiřuje. Rozdíly mezi nimi jsou však mnohdy závratné a úspěch příslušných investic přitom nejistý. Tato práce přináší možný náhled na kategorizaci sémantických aplikací a využívá je pro formulování kritických faktorů úspěchu nasazení těchto technologií ve firemním prostředí. Závěrem pak nastiňuje způsob, jakým je možné formulovat modely zralosti firmy pro předběžné posouzení vhodnosti tohoto zavedení.

**Klíčová slova:** sémantické technologie, kritické faktory úspěchu, zralost procesů

## 1 Úvod

Před rokem 2001 byl web širší veřejností považován za pouhý shluk statických webových stránek, poté však Tim Berners-Lee, někdejší ředitel konsorcia W3<sup>1</sup>, zavedl ve svém nejznámějším článku [1] pojem *Sémantický web*. Od té doby urazili sémantické technologie dlouhou cestu a dnes se setkáváme s jejich aplikacemi v mnohem širších oblastech počínaje nejrůznějšími aplikacemi pro integraci dat z různých zdrojů, pro podporu vyhledávání v různorodých datech, pro odvozování nových vztahů napříč heterogenními datovými základnami, přes aplikace pro podporu sociálních sítí, pro podporu rozhodování managementu, pro anotování a indexování libovolného obsahu až po tak rozdílné činnosti jako extrakce informací z nestrukturovaných zdrojů či dokonce tzv. Business Intelligence 2.0 [3].

Takto široký záběr sémantických (tj. znalostních) technologií je převážně způsoben tím, že obecné chápání toho, co vše lze považovat za sémantickou aplikaci je poněkud volné a neexistuje tudíž žádná obecně přijímaná definice. Například podle

---

<sup>1</sup> <http://www.w3.org/>

[6] lze za sémantickou technologii považovat libovolnou aplikaci, která uchovává význam odděleně od dat a souborů obsahu, a zároveň tento význam není zakomponován v aplikačním kódu. Toto pojetí tak zahrnuje zejména použití ontologických jazyků (jako RDF<sup>2</sup>, RDFS<sup>3</sup>, OWL<sup>4</sup> aj.) a pravidlových systémů.

Zájmová skupina při W3C pro vzdělávání a šíření sémantického webu shromažďuje a publikuje případové studie existujících aplikací využívajících sémantické technologie v praxi a také možné případy jejich užití [5]. Díky těmto studiím lze tak získat celkový náhled na současný stav a použitelnost znalostních technologií v produkčním prostředí.

Při snaze použít znalostní aplikace v komerční sféře je ovšem třeba obhájit příslušné investice. Existuje však hodně pohledů na to, co takové investice podniku přináší. Některé jsou zřejmé a dobře uchopitelné, jako například pohled finančních ukazatelů, které porovnávají výnosy a náklady. Jiné nejsou tak zřejmé ale jsou neméně důležité při hodnocení investic, obzvláště pak v tomto případě, kde přínosy příslušné aplikace lze kvantitativně vyjádřit jen velmi těžko. Takovými pohledy může být například hodnota pro zákazníky nebo zvýšení produktivity zaměstnanců. Mnohem zásadnějším problémem při zvažování podobné investice je ale odhadování (či obhajování) její uskutečnitelnosti a stanovování nutných podmínek, při kterých projekt nebude ztrátový.

Protože ale, jak jsem již zmínil, je povaha sémantických aplikací velmi různorodá, nelze podobné úvahy o proveditelnosti a návratnosti investic provádět souhrnně, nýbrž je třeba identifikovat jisté kategorie znalostních aplikací. Pro různé kategorie pak zcela jistě platí různé požadavky na proveditelnost projektu.

Tato práce si klade několik cílů a identifikace běžných kategorií sémantických aplikací je prvním z nich. Jakmile budou sémantické aplikace kategorizovány, bude možné identifikovat některé jejich podstatné vlastnosti izolovaně podle jednotlivých typů, a protože posuzování jejich přínosů na této úrovni by bylo až příliš obecné, načrtnu v závislosti na jednotlivých typech pak možné kritické faktory úspěchu (*CSF - Critical Success Factors* [4]). Dalším cílem práce je tedy určení nejpodstatnějších CSF při nasazování sémantické aplikace. V závěru práce se pak pokusím nastínit, jak by bylo možné na základě těchto kritických faktorů sestavit pro některé typy znalostních aplikací znalostní modely pro určení připravenosti

---

<sup>2</sup> <http://www.w3.org/RDF/>

<sup>3</sup> <http://www.w3.org/TR/rdf-schema/>

<sup>4</sup> <http://www.w3.org/TR/owl-features/>

firmy (z hlediska zralosti podnikových procesů ve smyslu původní práce W. Humphreyho [2]).

## 2 Kategorizace znalostních aplikací

Jak již bylo zmíněno výše, nelze znalostí aplikace posuzovat jako jednu celistvou oblast, neboť se vskutku jedná o velice různorodá použití příslušných technologií. Jednotlivé aplikace se mohou jedna od druhé radikálně lišit v mnoha aspektech, ať už je to rozsah datové základny, počet zúčastněných stran, typ vstupů i výstupů ba dokonce i samotná podstata činnosti. Z tohoto důvodu je tedy kategorizace sémantických aplikací mnohadimenzionální problém.

Na základě studia zmíněných případových studií publikovaných zájmovou skupinou pod W3C [5] (v době publikování tohoto příspěvku jich bylo vzato v úvahu 20) se objevily následující aspekty rozlišující sémantické aplikace mezi sebou (ne všechny ale mají vliv na formování faktorů úspěchu, tomu se budu věnovat v následující části):

- **zdroje informací** - z popisu povahy sémantických aplikací plyne, že každá musí používat minimálně jeden znalostní model (*ontologii* či alespoň *taxonomii*). Krom toho ale některé aplikace pracují s dalšími strukturovanými znalostmi a počítají s proměnnou znalostní základnou. Mimo to také samozřejmě mohou aplikace pracovat s daty jiné povahy. Znalostní aplikace lze tak dělit podle toho, zda zpracovávají strukturované znalosti, strukturovaná data či nestrukturovaná data.
- **povaha zdrojových informací** – sémantické aplikace lze rozlišit podle toho, zda informace, se kterými pracují, vznikají v jiných systémech (popřípadě jsou již ve strukturované podobě k dispozici) nebo zda vznikají speciálně pro tento systém. Pokud vznikají speciálně pro danou aplikaci lze ještě rozlišit případy, kdy se tak děje ručně, automaticky z jiných zdrojů či jako vedlejší efekt jiné činnosti (například běžné práce uživatele).
- **přesnost vstupů a výstupů aplikace** – u sémantických aplikací lze rozlišit rozdílný přístup ke zpracování vstupů na výstupy. Zde lze aplikace rozdělit na ty, které s daty pracují pevně a spoléhají na jejich úplnou přesnost, aplikace, které sice očekávají, že data mohou být neúplná, ale neočekávají v nich nekonzistence a nepracují s neurčitostí, a nakonec aplikace, které již obsahují zpracování neurčitosti.
- **doménová specifičnost a znovupoužitelnost aplikace** – z důvodu oddělení významu dat od jejich zpracování by sémantické aplikace měli být doménově závislé jen minimálně, i zde však existují výjimky, které zahrnují například specifické rozhraní přípůsobené konkrétní doméně.

- **počet a druh uživatelů** – uživatele sémantické aplikace mohou tvořit jednotliví neodborní uživatelé, odborní specialisti (doménoví experti), znalostní experti či management. Rozlišit lze aplikace také podle toho, zda jsou určeny pro jednotlivce, pracovní skupiny či tisíce uživatelů v podobě sociální sítě.
- **vztah poskytovatel x uživatel** – zde se podařilo identifikovat několik možností provozu aplikace: uživatel je jednotlivec a provozuje aplikaci pro vlastní potřebu, uživatele tvoří více rovnoprávných uživatelů či sociální síť a provoz je zajištěn komerčně, komunitou či jinak neziskově, uživatele tvoří zákazníci provozovatele a nebo uživatele tvoří zaměstnanci firmy (provozovatele). U posledních dvou možností lze ještě rozlišit případy, kdy provoz aplikace tvoří hlavní činnost firmy a kdy tvoří jen podpůrný proces a lze tudíž uvažovat o jeho outsourcingu. V případě, že je provoz aplikace zajištěn komunitou, může být dělen podle toho, je-li provozován centralizovaně či decentralizovaně.
- **frekvence přístupu k aplikaci a její dostupnost** – k aplikacím může být přístupováno neustále (24/7), náhodně, pravidelně či jednorázově. Krom toho je třeba rozlišit také požadavky na dostupnost aplikace: buď musí být aplikace dostupná neustále, v definovaných intervalech nebo na vyžádání (daném reaktivně např. ručním spuštěním aplikace).
- **druh činnosti aplikace** – z analýzy případových studií se podařilo identifikovat několik hlavních druhů činností sémantických aplikací. Jsou jimi indexace dat, integrace dat a odvozování. Tyto činnosti jsou však ve většině případů spíše prostředkem než účelem činnosti aplikace (až na výjimku integrace heterogenních dat). Z nich pak lze odvodit několik dalších činností, pro které je již aplikace povětšinou tvořena, jsou jimi např. podpora lepšího vyhledávání (indexace + integrace), procházení heterogenní datové základny a navigace v doméně (integrace + indexace), doporučování nových vztahů entit (odvozování) či adaptabilita systému na změnu datové struktury (integrace dat).

Pochopitelně si lze představit sémantickou aplikaci zařazenou v nastíněné klasifikaci víceméně libovolně, ale v předkládaných případových studiích lze vypozařovat určité shluky a koincidence mezi jednotlivými aspekty. Toto setřídění tak umožňuje identifikovat a pojmenovat jisté základní archetypy sémantických aplikací:

- **„vylepšený vyhledávač“** – tyto aplikace se soustředí na indexaci dat, často spojenou s integrací dat z různých jiných systémů, kde tato data vznikají automaticky. Tato automatická zdrojová data bývají také často doplněna ručními anotacemi. Aplikace tohoto typu pracuje jak s daty strukturovanými, tak s daty nestrukturovanými (pomocí automatických filtrů a wrapperů). Hlav-

ním přínosem těchto aplikací je jednak umožnění prohledávání heterogenní datové základny a jednak tvorba komplikovanějších dotazů bez nutnosti apriorní znalosti datové struktury. V ostatních aspektech se však již mohou velmi lišit, a tak sem lze zařadit tak rozdílné aplikace jako např. podporu pro anotování a vyhledávání v souborech na osobním počítači<sup>5</sup>, veřejný portál pro vyhledávání v poznacích čínské medicíny<sup>6</sup> či správa zvukových nahrávek archivu norské rozhlasové stanice<sup>7</sup>.

- **„rozhraní k prohlížení dat“** – tyto aplikace navazují na schopnosti předchozího typu, avšak rozvíjejí je nejen o možnosti zobrazování různorodého obsahu (videonahrávky, odborné články, chemické vzorce atd.), ale také o možnost vizuálního procházení daty nezávisle na jejich struktuře. Tyto aplikace bývají většinou zaměřeny na použití odbornými specialisty a provozány jsou buď komerčně pro vnitropodnikovou potřebu nebo neziskově pro podporu odborné veřejnosti (a zároveň jako propagace vlastní technologie). Příklady tohoto archetypu mohou být např. systémy pro sdružování medicínských informací ať už za účelem usnadnění léčení pacientů<sup>8</sup> či dosažení úspor při vývoji nových léčiv<sup>9</sup> nebo portál pro sdružení programátorských znalostí společnosti Oracle<sup>10</sup>.
- **„doporučovací nástroj“** – podstatou těchto aplikací je odvozování nových vztahů mezi jednotlivými entitami. Krom všech ostatních typů zdrojových dat mívají tyto aplikace často za vstup i data generovaná automaticky jako vedlejší efekt běžné činnosti uživatele, díky čemuž mohou krom jiného navrhovat nové vztahy i na základě aktuálního kontextu. Uživatel je pak nežádka zákazníkem poskytovatele a to ať už se jedná o placenou službu, veřejnou službu (např. navrhování individuální prohlídky města Zaragoza<sup>11</sup>) či komerční způsob pro personalizované cílení reklamy spolu s poskytováním služby (jako například systém pro doporučení služeb uživatelům mobil-

---

<sup>5</sup> <http://nepomuk.semanticdesktop.org>

<sup>6</sup> <http://www.cintcm.com>

<sup>7</sup> <http://www.nrk.no/>

<sup>8</sup> <http://www.pharmasurveyor.com/>

<sup>9</sup> <http://www.lilly.com>

<sup>10</sup> <http://otnsemanticweb.oracle.com/>

<sup>11</sup> [http://www.zaragoza.es/turruta/Turruta/en/index\\_Ruta](http://www.zaragoza.es/turruta/Turruta/en/index_Ruta)

ních zařízení<sup>12</sup>). Aplikace v této kategorii často pracují s neurčitostí, a tak sem lze zařadit také nejrůznější expertní systémy.

- **„rámec pro výměnu dat“** – aplikace této kategorie jsou provozovány distribuovaně, a dovolují „jen“ sjednotit struktury vyměňovaných dat mezi účastníky, nezávisle na jejich obsahu. Tento obsah se tak může časem vyvíjet a být přizpůsobován potřebám konkrétního bilaterálního vztahu výměny a přesto být předáván ve standardizovaném formátu. Jako příklad lze zde uvést iniciativu za zavedení sémantické výměny dat v ropném a plynárenském průmyslu<sup>13</sup>.

Jistě by bylo možné objevit i další archetypy sémantických aplikací, tyto čtyři však považuji za nejčastější. Samozřejmě existují i aplikace, které žádnému z těchto archetypů přiřadit nelze, stejně tak jako jiné naopak leží na pomezí dvou nebo více.

### 3 Kritické faktory úspěchu

Jak již bylo naznačeno v minulé části, syntézou rizik zmiňovaných v jednotlivých případových studiích [5] lze nastínit nejčastější kritické faktory úspěchu při vývoji a nasazování sémantických aplikací do produkčního prostředí. Tyto faktory ale nejsou univerzální, nýbrž každý platí vždy jen pro určitou skupinu aplikací, danou aspekty jejich kategorizace, zmiňovanými výše. Dosud identifikované kritické faktory úspěchu sémantických aplikací tedy jsou:

- **správnost použité ontologie/taxonomie** – tento faktor je platí pro všechny znalostní aplikace, a čím je použitý znalostní model komplikovanější a fixovanější, tím je také jeho správnost zásadnější. Naplnění tohoto faktoru úspěchu s sebou nese nutnost najímání kvalitních analytiků a znalostních inženýrů v období vývoje a zavádění aplikace, což implikuje nemalé náklady.
- **dostatečná strmost učící křivky koncových uživatelů** – platí pro aplikace, u nichž je koncovým uživatelem jednatel. Sémantika u tohoto typu aplikací s sebou nese dosti atypický způsob ovládání oproti standardním aplikacím a proto učící křivka má velmi pozvolný náběh. Nutností tak je nejen sro-

---

<sup>12</sup> <http://www.w3.org/2001/sw/sweo/public/UseCases/SaltLux-KTF/KTF.pdf>

<sup>13</sup> <http://www.w3.org/2001/sw/sweo/public/UseCases/Chevron/>

zumitelné a pokud možno intuitivní ovládání systému, ale také srozumitelnost, přehlednost a přesnost poskytovaných výstupů.

- **potenciál možných přínosů aplikace kompenzovat dočasné snížení produktivity v období zavádění a učení** (a také provozní náklady) – přínosy aplikací bývají velmi různorodé a často velmi vágní (oproti klasickým řešením) a tak se jen velmi nelehko odhaduje jejich výše a význam v době zavádění aplikace. Provozní náklady jsou povětšinou srovnatelné se standardními aplikacemi, ale při zavádění je nutné počítat s dočasným snížením produktivity uživatelů (viz předchozí bod). Aby aplikace byla úspěšná, nemělo by toto dočasné zhoršení být natolik závažné, aby zastínilo její možné přínosy.
- **vůle a kázeň všech stran používat stejný znalostní model** – v případě, že je znalostní aplikace provozována distribuovaně, je nutné, aby všechny zúčastněné strany používaly společný znalostní model. Objevuje se zde tedy možné riziko v podobě nutnosti vyjednávání o jeho podobě a obsahu.
- **synchronizovaná distribuce centrální ontologie** – postupně může docházet k úpravám centrálního znalostního modelu a v případě, že je aplikace provozována distribuovaně, je nutné, aby byly tyto změny patřičně šířeny mezi zúčastněné strany, nebo by mohlo docházet k nekonzistencím. Při těchto změnách samozřejmě platí také předchozí bod.
- **dostatečný počet uživatelů** - pokud se příslušná sémantická aplikace zakládá na sociálních datech, je její úspěch podmíněn dostatkem uživatelů, kteří mají tato data produkovat. Riziko se v tomto případě vyskytuje v podobně nutných nákladů na propagaci nově vznikajícího systému.
- **motivace uživatelů** – tento kritický faktor se vyskytuje na dvou úrovních. První je při období zavádění aplikace, kdy na uživatele působí negativní stimulace v podobě pomalého náběhu učící křivky, uživatel tak v první chvíli postrádá motivaci se se systémem učit zacházet. Druhá pak při samotném provozu, častým zdrojem dat pro sémantické systémy libovolného rozsahu jsou totiž ruční anotace, jejichž tvorba je pro uživatele pracná. Částečnou motivací může být potenciální přínos lepších výsledků či usnadnění práce v budoucnu. Krom toho se při učení může uživatel motivovat možností využít nabyté zkušenosti jinde, neboť napříč různými aplikacemi se používají obdobné technologie (např. SPARQL dotazování).
- **dostatek zdrojových dat** - u aplikací provádějících odvozování je dostatek zdrojových přímo zásadní pro poskytování přínosných výsledků (tj. přidaná hodnota sémantiky). Dokonce i u aplikací zakládajících na indexování dat je dostatek dat kritický pro úspěch aplikace, protože při malém objemu dat sice sémantické indexování dává obdobné výsledky jako tradiční metody, ale s vyššími počátečními náklady.

- **různorodost zdrojů i formy dat** – čím je komplikovanost použitého jazyka z hlediska použitých znalostí vyšší, tím přínosnější výsledky mohou dávat aplikace zakládající na odvozování vztahů. Obdobně pak čím je různorodost obsahu dat vyšší, tím přínosnější budou výsledky budou dávat aplikace provádějící jejich sémantickou integraci. Na triviálních systémech se tedy použití těchto technologií pravděpodobně nevyplatí.
- **udržení minimálně stejné přesnosti výsledků jako u dílčích systémů** - u aplikací provádějích integraci dat vzniká riziko nekonzistence dat ze zdrojových systémů. Samotná funkčnost sémantických aplikací není podmíněna konzistentními daty, ale je třeba nekonzistenci při návrhu očekávat. Dobře odhadnutá spolehlivost dat je tak v tomto bodě rozhodující.
- **spolehlivost parserů a wrapperů** - pokud aplikace zpracovává nestrukturovaná data, je závislá na výstupech parserů a wrapperů nejrůznějšího obsahu, popřípadě na systémech zpracování přirozeného jazyka. Zde platí opět totéž co v předchozím bodě, a sice že je nutné předem správně odhadnout spolehlivost takto získávaných informací.

Tyto kritické faktory samozřejmě budou mít u různých aplikací různou váhu. Pokud bude například aplikace zdrojová data sbírat automaticky a výstupy podle kontextu sdělovat uživateli v téměř přirozeném jazyce, lze očekávat poměrně strmou učití křivku, takže období snížené produktivity bude vcelku minimální a naopak na jeho kompenzování budou stačit i téměř nepatrné přínosy. Tyto univerzální kritické faktory lze tedy brát pouze jako výchozí body při posuzování konkrétního případu.

#### 4 Nástin zralostních modelů

Zralostní modely [2] se rozvinuly v uplynulých dvou desetiletích, aby umožnili posoudit připravenost podniku na provedení nějakého typu strukturální investice. Nejčastěji se používají při nasazování nějakých aplikací IT, jako jsou CRM systémy, ERP či Business Intelligence.

Dle mého názoru by mělo být možné na základě výše uvedených aspektů kategorie a jim přidružených kritických faktorů úspěchu sestavit zralostní modely podniku vzhledem k zavádění sémantických technologií určitého typu. Ač by tyto modely byly bez faktického obsahu, bez cílového prohlášení, mohly by být formulovány konkrétními příklady a minimálně pro každý archetyp by tak bylo možné formulovat určité úrovně požadavků na nějakou stránku podniku, které by měly být splněny, aby bylo možné uvažovat o přínosnosti toho kterého řešení.

Příklad pro archetyp sémantického vyhledávače by mohl vypadat takto: Pokud podnik používá jeden zdroj dat a proprietární datovou strukturu, pak je pro zavádě-

ní podobného systému nepřipraven. Pokud používá více systémů s heterogenní datovou strukturou, může zavedení vyhledávače se sémantickým indexováním přinést zlepšení výsledků vyhledávání. Další úroveň podnik dosáhne, pokud použije více systémů se standardizovanou datovou strukturou, pak může začít přemýšlet o integraci se systémy výměny dat, atd...

## 5 Závěr

V této práci byl nastíněn způsob kategorizace sémantických aplikací a za pomoci analýzy katalogu případových studií skupiny konsorcia W3 pak identifikovány jejich čtyři základní archetypy. Vzhledem k navržené kategorizaci pak byly formulovány nejčastější kritické faktory úspěchu při nasazování sémantických aplikací ve firemním prostředí spolu s mezemi jejich platnosti.

V závěrečné části práce byl také nastíněn způsob, jakým lze nalezené kritické faktory a navržené členění využít při formulaci modelů zralosti podniku. Samotná formulace konkrétních zralostních modelů je pak předmětem mé budoucí činnosti.

## Literatura

- [1] Berners-Lee, T., Hendler, J., Lassila, O.: The Semantic Web, *Scientific American*, May 2001. <http://www.sciam.com/article.cfm?id=the-semantic-web>
- [2] Humphrey, W.: *Managing the Software Process*, Addison-Wesley Professional, Massachusetts, USA, 1989
- [3] Raden, N.: Business Intelligence 2.0: Simpler, More Accessible, Inevitable, *Intelligent Enterprise*, <http://www.intelligententerprise.com/showArticle.jhtml?articleID=197002610>
- [4] Rockart, J. F.: Critical Success Factors, *Harvard Business Review*, 1979, str. 81-91
- [5] Semantic Web Education and Outreach (SWEO) Interest Group: *Semantic Web Case Studies and Use Cases*, <http://www.w3.org/2001/sw/sweo/public/UseCases/>
- [6] Staab, S., Studer, R.: Handbook on Ontologies. *International Handbooks on Information Systems*, Springer Verlag, 2004

## **Summary**

### **Maturity Models for Deployment of Semantic Technologies in Enterprises**

In recent years semantic technologies are included in broader and broader areas of application, and their scope is constantly expanding. The differences between them, however, are often vast and the success of the investments and the uncertain. This work provides a possible approach to the categorization of the semantic applications and uses it to formulate the critical factor of success for deployment of these technologies in an enterprise environment. Finally, it outlines how it is possible to formulate a capability maturity model of a company for a preliminary assessment of the appropriateness of such deployment.