

# Portál sociální sítě informatiků a jeho strukturované profily znalostí

Kamil Matoušek<sup>1</sup>, Jiří Kubalík<sup>1</sup>, Martin Nečaský<sup>2</sup>

<sup>1</sup> České vysoké učení technické v Praze  
Fakulta elektrotechnická  
Technická 2, 166 27 Praha 6  
{matousek,kubalik}@fel.cvut.cz

<sup>2</sup> Univerzita Karlova v Praze  
Matematicko-fyzikální fakulta  
Malostranské nám. 25, 118 00 Praha 1  
[necasky@ksi.mff.cuni.cz](mailto:necasky@ksi.mff.cuni.cz)

## 1 Úvod

V tomto článku popisujeme aktuální stav vývoje webového portálu [sosit.cz](http://www.sosit.cz), který je vyvíjen v rámci projektu Sociální síť informatiků v regionech ČR (SOSIREČR, <http://www.sosirecr.cz>). Portál je založen na dnes již běžném konceptu sociální sítě. Od ostatních portálů se liší svým zaměřením na komunitu informatiků v regionech ČR.

Impulzem pro práci na portálu byl fakt, že úspěšnost řešení každého projektu je založena na sestavení vhodného týmu. V oblasti aplikovaného výzkumu, kdy je zásadní schopnost navazování spolupráce mezi firemní a akademickou sférou, je však budování týmů v současnosti problematické. Je to způsobeno především nedostatkem výměny informací o existující nabídce a poptávce informatického výzkumu a výzkumných zdrojů (lidé, skupiny, výstupy z existujících projektů atp.). Prvním konkrétním cílem portálu je proto nabídnout prostředí pro výměnu a sdílení informací o probíhajícím informatickém výzkumu v regionech ČR a pro střet nabídky a poptávky v oblasti lidských zdrojů pro informatický výzkum.

I v případě výzkumného týmu jsou potřební pracovníci vykonávající standardní informatické profese, jako např. programátoři či systémoví správci. Druhým cílem portálu je proto také nabídnout prostředí pro střet nabídky a poptávky lidských zdrojů pro běžné informatické profese.

Je samozřejmě také nutné mít přehled o tom, jaké vzdělání nabízejí vysoké školy v ČR. Je dobré mít přehled o kvalitě jednotlivých studijních oborů, jak jsou jejich absolventi s oborem spokojeni a kde nacházejí uplatnění. Třetím cílem portálu je proto dále rozšířit nabídnuté prostředí o sdílení informací o studijních oborech vysokých škol, jejich hodnocení odbornou veřejností a sledování uplatnění jejich absolventů.

Článek je organizován následovně. V kapitole 2 formou uživatelských scénářů popisujeme základní funkční požadavky na portál. V kapitole 3 se potom věnujeme implementaci portálu. V kapitole 4 ukazujeme, jak současná implementace portálu naplňuje uvedené uživatelské scénáře. V poslední kapitole 5 se pak věnujeme zamýšleným rozšířením portálu.

## 2 Uživatelské scénáře

Než přistoupíme k popisu funkcionalit nabízených portálem, uvedeme několik scénářů, které demonstrují, pro jaké účely mohou uživatelé portál využít. V pozdějších kapitolách potom popíšeme, jak konkrétně může uvedené scénáře uživatel v portálu realizovat.

### 2.1 Vyhledávání výzkumných partnerů

Regionální firma *ContractsOnline* při realizaci informačního systému pro správu veřejných zakázek pro města ve svém regionu zjistila, že na Internetu existuje celá řada dostupných datových zdrojů nabízených

státní správou (obchodní rejstřík, insolvenční rejstřík, centralniadresa.cz, ...), které by bylo vhodné do připravovaného systému integrovat. Také se dozvěděla o existující iniciativě *OpenGov.eu*, jejímž cílem je zpřístupnění dat státní správy v otevřené a strojově zpracovatelné podobě. Firma se rozhodla, že potřebná data vytěží z existujících datových zdrojů pomocí technik zmiňovaných iniciativou *OpenGov.eu*. Studium webových stránek iniciativy firma zjistila, že hlavní technologickou myšlenkou je zpřístupnění dat státní správy v podobě formátu RDF v prostředí Linked Data a sémantického webu a že důležitým aspektem je také strojové zpracování existujících textů na webu do podoby RDF. *ContractsOnline* však nemá dostatečné know-how v této oblasti. Nezná formát RDF ani pojmy Linked Data a sémantický web. Neumí pracovat s metodami strojového zpracování textů. Ráda by proto využila prostředí, které by umožnilo vyhledání následujících informací:

- Jaké skupiny či osoby v ČR mají know-how v oblasti strojového zpracování textů, webových technologií a sémantického webu?
- Jaké skupiny či osoby v ČR spolupracují nebo se účastní iniciativy *OpenGov.eu*?
- Jsou nebo byly v ČR řešeny projekty v uvedených oblastech?

Na Internetu však nenalezneme portál, který by umožňoval požadované informace jednoduše nalézt. Firma *ContractsOnline* je odkázána buď na síť svých osobních kontaktů nebo na fulltextové vyhledávání v obecném prostředí Webu pomocí vyhledávacích robotů typu Google. Síť osobních kontaktů je však velmi úzká, často nepřesahuje hranice regionu a nezasahuje do akademického prostředí, ve kterém se právě hledané skupiny, osoby i projekty vyskytují. Fulltextové vyhledávání vrací velké množství irelevantních odkazů a jen málo relevantních. Typicky lze nalézt jen několik málo pracovišť, přičemž detailnější informace o realizovaných projektech a spolupráci jsou prakticky nedohledatelné. Celkově je tedy takové vyhledávání v prostředí Internetu dnes velmi časově náročné a neefektivní.

## 2.2 Hledání lidských zdrojů

Katedra softwarového inženýrství na Vysoké škole informatiky byla v jeden okamžik úspěšná v podání několika grantových projektů. Bohužel, její současné kapacity na pokrytí všech projektů nestačí. Jeden z projektů vyžaduje skupinu 5ti J2EE programátorů v oblasti mobilních aplikací, kterými katedra nedisponuje. Jiný z projektů vyžaduje experta na výzkum v oblasti databázového zpracování RDF dat. Jediný expert na katedře však odešel v mezidobí mezi podáním projektů a potvrzením jejich financování a není již dále možné jej v projektu využít. Je proto nutné velmi rychle doplnit požadované kapacity. Katedra by proto využila prostředí, které by umožnilo vyhledání následujících informací:

- Kdo se zkušenostmi z práce na grantových projektech a/nebo v oblasti programování webových aplikací nabízí své služby jako J2EE programátor?
- Jací jsou v ČR výzkumníci v oblasti databázového zpracování RDF dat s publikacemi v oblasti na významných konferencích, případně navíc se zkušenostmi z EU či národního výzkumného projektu v oblasti zpracování RDF dat?

Katedra má, podobně jako firma *ContractsOnline* v případě předchozího scénáře, možnost využít svoji vlastní síť osobních kontaktů nebo fulltextového vyhledávání. Žádná z možností však nemůže poskytnout dostatečně aktuální a úplné informace o osobách. Získání úplné informace vyžaduje značné vyhledávací úsilí.

## 2.3 Propagace výzkumného zaměření a výsledků

Výzkumná skupina XRG na MFF UK vytvořila nástroj pro navrhování a údržbu sady XML schémat. Nástroj je podložen teoretickým výzkumem publikovaným na odborných mezinárodních konferencích a ve vědeckých časopisech. Skupina vypracovala několik případových studií, ve kterých demonstrovala použitelnost svého nástroje. Chce nyní nástroj představit širší odborné veřejnosti a získat zpětnou vazbu.

Také hledá společnost, která by pomohla nástroj transferovat do praxe. Konkrétně by skupina využila následující služby:

- zveřejnění nabídky svého know-how a nástroje,
- šíření nabídky potenciálním zájemcům z řad akademických pracovníků i firem.

Uvedený scénář může nastat i v případě jednotlivce, který nabízí svoji odbornost jiným skupinám či projektům. Podobně jako v předchozích scénářích není síť osobních kontaktů dostatečná. Inzerce prostřednictvím pracovních portálů není vhodná, protože neumožňují popsat know-how a zkušenosti výzkumných týmů i jednotlivců. Zveřejnění nabídky je tak fakticky možné (a také se tak děje), prostřednictvím oficiálních webových stránek výzkumné skupiny, případně mateřské instituce. Pro cílené šíření k potenciálním partnerům však není tento způsob dostatečný.

## 3 Implementace portálu

### 3.1 Základní entity portálu

Popišme si nyní základní entity, se kterými se na portálu Síť IT můžeme setkat. Typické pro ně je, že mohou mít vlastní stránku, kde se mohou prezentovat, v některých případech jsou tyto stránky dále obohaceny o prostředky pro komunikaci a diskusi na související témata. Vedle textových informací mohou být entity také strukturovaně charakterizovány prostřednictvím několika znalostních resp. vědeckých profilů (viz níže).

- **Uživatelé** jsou základní aktivní entity portálu, které jsou automaticky vytvářeny pro všechny registrované uživatele. Uživatelé mají k dispozici možnost zaslání zpráv ostatním uživatelům portálu a samozřejmě také jejich přijímání. Dále portál zobrazuje entity **neaktivních osob**, které figurují pouze ve vztahu k jiným entitám portálu, jako jsou například členové vedení univerzit. Jejich stránky jsou napojeny na příslušné instituce a jsou zobrazovány obdobným způsobem, jako v případě stránek aktivních uživatelů, pouze v omezeném rozsahu.
- **Skupiny uživatelů** jsou vytvářeny aktivními uživateli za účelem spolupráce či pouhé výměny názorů na daná témata. Může jít o skupiny příznivců moderních technologií, stejně jako o pracovní skupiny řešící konkrétní výzkumné úkoly. Zájemce o účast ve skupině musí systém požádat o přidání příslušné vazby a tato vazba musí být potvrzena správcem skupiny. Tím je zamezeno nežádoucím uživatelům, aby se účastnili privátní diskuse členů skupiny.
- Na stránkách **firem** lze prezentovat informace a zveřejnit kontakty zajímavých společností v oblasti IT. Při spuštění portálu jsou již založeny entity vybraných společností a další „**aktivní**“ **firmy** mohou v rámci portálu zakládat a spravovat sami registrovaní uživatelé. Diskusní prostor příslušný těmto entitám je veřejně přístupný všem uživatelům portálu, například jako možnost pro vyjádření názoru na spolupráci s danou společností.
- **Instituce, vysoké školy** a další organizace jsou obsahem údajů obdobné firmám, ale většinou se liší svým zaměřením; u vysokých škol jsou doplněny vazby na jejich organizační složky a informace o příslušných studijních oborech. Seznamy těchto organizací jsou předem dány při spuštění portálu a uživatelé nemohou zakládat nové entity. Toto mohou případně na požádání zařídit správci portálu.
- **Projekty** obsahují jednak entity získané sběrem informací o projektech podporovaných z veřejných prostředků ČR. Vedle toho uživatelé mohou zakládat další vlastní projekty, na kterých pracují a prezentovat tak informace o nich ostatním uživatelům.
- **Inzeráty** představují vhodný prostor pro vyjádření nabídky a poptávky v oblasti IT, například v případě spolupráce na projektu. Poptávat i nabízet jejich prostřednictvím lze jednotlivce nebo skupiny osob, např. odborné týmy. Na příslušných stránkách pak mohou uživatelé vyjádřit zájem, o

němž systém uvědomí správce inzerátu. Obsah portálu je navíc obohacován o aktuální poptávku po relevantních profesích v oblasti IT na úřadech práce České republiky.

- Prezentace významných vědeckých **časopisů** na portálu obsahují především užitečné informace o jejich impakt faktoru, počtech článků apod. a příslušné odkazy na tato periodika.
- V prezentacích **studijních oborů** souvisejících s oblastí informačních technologií jsou spolu s uvedením studijního programu, uvedeny také informace o charakteru studia (zda jde např. o magisterské studium) s vazbou na vysokou školu, kde je vyučován, včetně identifikátoru příslušné akreditace a informací, dokdy je akreditace platná.

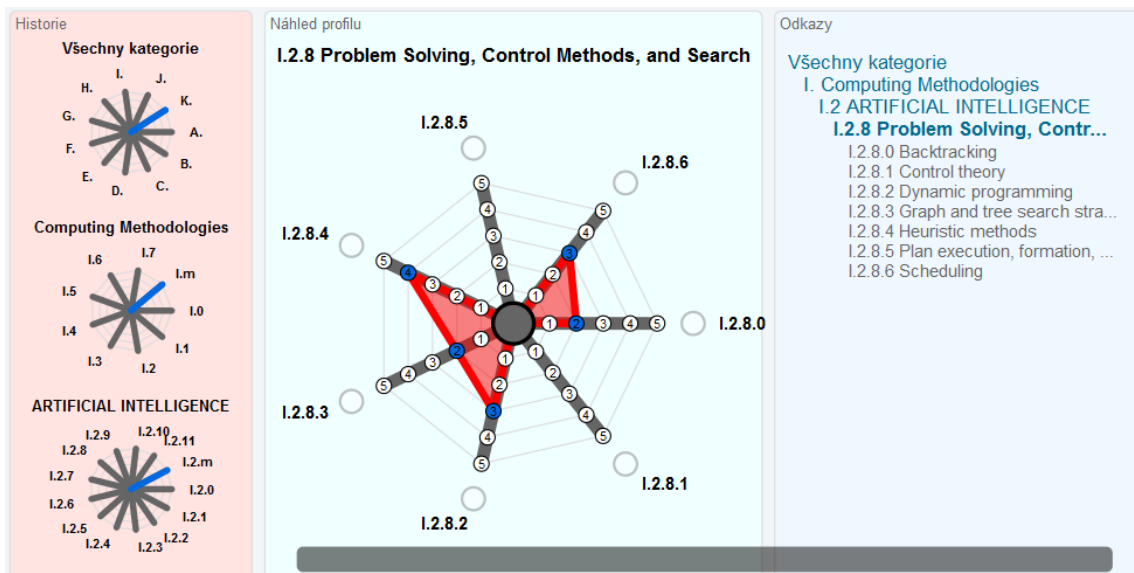
Mezi svými entitami portál Síť IT eviduje určité typy vztahů, které si lze zobrazit z příslušných stránek entit. Většina z nich byla popsána v předcházejícím výčtu entit. Obecně pro ně platí, že vznik či změnu vazby musí potvrdit druhá strana, tj. například zájemce o členství v konkrétní skupině portál ke skupině přiřadí až po potvrzení příslušným správcem této skupiny.

Uživatel - *správce entity* také disponuje dalšími právy: může upravovat obsah jejich stránek, vystupuje jako moderátor příslušných diskusí, může definovat další správce spravovaných entit a může publikovat novinky s případnými přílohami v podobě souborů. Ke sledování (tj. odběru) novinek vybraných entit se mohou přihlásit všichni ostatní uživatelé portálu. Založené entity (například své neaktuální inzeráty) pak může také odstranit.

## 3.2 Odborné profily

Takzvané odborné profily (OP) představují specifický nástroj portálu pro strojově čitelný popis znalostí a odbornosti v ICT. Obecně se jedná o strukturované datové typy, reprezentující hierarchicky členěné kategorie. Rozlišujeme dva základní typy OP – jednoúrovňový znalostní profil a víceúrovňový vědecký profil. Znalostní profil, zavedený v [3], se skládá ze 16 os, z nichž každá reprezentuje určitou oblast inženýrských znalostí, např. modelování procesů či datové inženýrství. Vědecký profil charakterizuje vědecké zaměření dané entity v podobě hierarchicky strukturované podle všeobecně uznávané ACM klasifikace [4]. ACM klasifikace má hierarchickou strukturu – kategorie může být buď listová a nebo může být členěna na podkategorie. U obou typů OP se používá pro každou osu ohodnocení na stupnici 0-5. Obecně nic nebrání tomu, aby byl nadefinován a použit libovolný další hierarchický OP. Příklad vyplněného vědeckého profilu je na Obrázku 1.

OP půjde použít jak k popisu znalostního profilu či odbornosti celé entity, ale také k popisu jednotlivých dat prezentovaných na stránkách entit. Vědecký profil může např. charakterizovat znalosti a odborné zaměření osoby nebo celé skupiny osob (např. výzkumného týmu). Může také charakterizovat vědecké zaměření projektu či volného pracovního místa v rámci projektu. Znalostní profil studijního oboru může charakterizovat znalosti typického absolventa tohoto oboru. Znalostní profil také může charakterizovat požadované znalosti ideálního kandidáta na pracovní pozici.



Obrázek 1: Příklad vyplněného vědeckého profilu podle klasifikace ACM.

### 3.3 Profil uživatele

Základní rozvržení informací na stránkách si popíšeme na příkladu stránek uživatele, konkrétně vlastní stránky přihlášeného uživatele. Stránka se dělí na tři hlavní části, viz Obrázek 2:


- Blok hlavních informací o dané entitě obsahuje jméno, foto, odkaz na domácí webovou stránku, mapku zobrazující domácí region uživatele a náhled na strukturovaný vědecký nebo znalostní profil. Uživatel může pomocí vědeckého a znalostního profilu specifikovat například vystudovaný obor, téma své diplomové/disertační práce apod.
- Nově vytvořenou stránku si uživatel (správce) může dočasně ponechat v rozpracovaném stavu, přičemž takováto rozpracovaná stránka (ani žádné na ní vyplněné informace) není přístupná žádným jiným uživatelům portálu. Rozpracovanou stránku lze nevratně převést do stavu Publikováno. Od tohoto okamžiku už není možné tuto stránku ani data na ní vložená před uživateli portálu opětovně skrýt. Samozřejmě, uživatel má kdykoliv možnost svoji stránku na portálu zrušit.
- Podrobné informace o entitě jsou umístěny ve spodním bloku oranžových záložek. Na záložce Detail profilu si uživatel může vyplnit kontaktní informace jako adresa bydliště, sídla instituce a pracoviště, kontaktní email a telefonní číslo. Dále informace o svém pracovním zařazení s odkazem na instituci/firmu, odkazy na zdroje relevantní k jeho odbornému zaměření, informace o studiu a pracovních zkušenostech, zajímavé webové zdroje, klíčová slova. Kromě výše zmíněných kategorií uživatel může zadat i libovolnou jinou textovou informaci, ke které může přiřadit jeden nebo více nahraných souborů.

Na záložce Novinky má uživatel prostor pro vkládání zpráv, upoutávek a poznámek, které chce zprostředkovat ostatním uživatelům. Každá novinka může být opět doplněna nahrávanými soubory.

Na záložce Zprávy má uživatel k dispozici jednoduchý interní zprávací systém, který využívá prostředků integrovaného fóra <http://www.sosirecr.cz/mybb/>. Systém umožňuje zasílání zpráv pouze mezi uživateli uvnitř portálu.

Záložka Vazby obsahuje seznam(y) vazeb mezi uživatelem a ostatními entitami na portálu. Typicky se zde zobrazují vazby uživatele na instituci, ve které uživatel pracuje (případně její konkrétní

organizační složky), vazby na projekty, kterých se uživatel účastní, členství ve skupinách uživatelů a vazby typu „kolega“ na jiné uživatele.

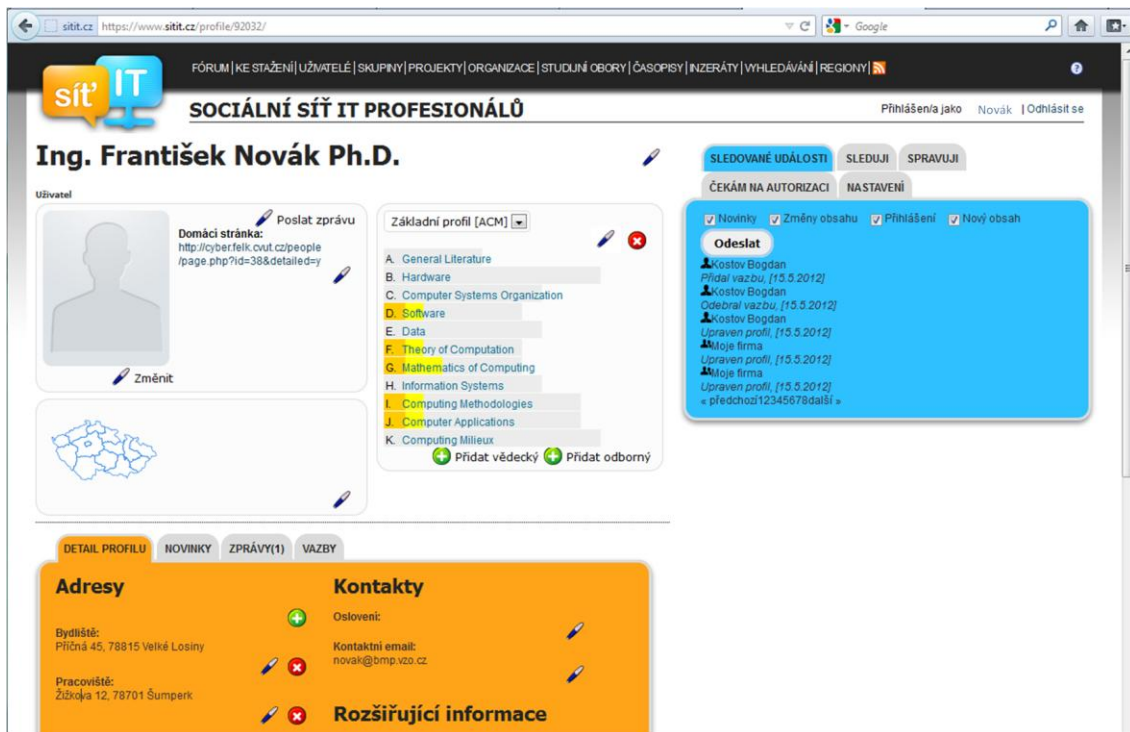
- V pravém bloku modrých záložek je záložka Sledované události se seznamem událostí ze sledovaných stránek portálu. Mezi sledované události patří novinky, přidání a úpravy obsahu a datum přihlášení. Seznam sledovaných stránek je na záložce Sleduji. Přidat stránku mezi sledované resp. odebrat stránku ze seznamu sledovaných stránek je možné kliknutím na ikonku  na příslušné stránce. Novinky ze sledovaných stránek lze odebírat také pomocí RSS kanálu.

Záložka Spravuji ukazuje aktuální seznam stránek, ke kterým má uživatel správcovská práva, tedy může editovat obsah těchto stránek.

Záložka Nastavení zatím nabízí pouze změnu hesla a zaslání zapomenutého hesla. V budoucnu mohou přibýt další možnosti uživatelského nastavení.

Struktura stránek ostatních entit na portálu je odvozena od tohoto základního modelu. Liší se pouze konkrétními údaji, které jsou relevantní pro danou entitu (např. jiné budou vazby u uživatele, jiné u skupiny).

Z hlediska orientace na stránkách portálu je důležité si uvědomit, že pravý blok modrých záložek obsahuje informace pro přihlášeného uživatele, které se nemění bez ohledu na aktuální navštívenou stránku portálu. Naproti tomu obsah záložek ve spodním bloku se vždy vztahuje k entitě prezentované na dané navštívené stránce. To znamená, že uživatel A, který si prohlíží svoji stránku, uvidí ve spodním bloku „svoje“ záložky Detail, Novinky atd. Zatímco, když bude uživatel A na stránce uživatele B, tak ve spodním bloku uvidí záložky vztahující se k uživateli B. V obou případech ale bude zobrazen stejný obsah záložek v pravém bloku, tedy seznam stránek sledovaných uživatelem A, seznam posledních událostí na stránkách sledovaných uživatelem A atd.



Obrázek 2: Náhled stránky přihlášeného uživatele.

### 3.4 Důvěryhodnost a vyhledávání

Jedním z hlavních přínosů portálu by měla být možnost vyhledání relevantních a důvěryhodných dat z oblasti ICT. U veřejných webových zdrojů je důvěryhodnost dat zajištěna a ověřena jejich poskytovateli. Věřitelnost dat vložených do portálu uživateli by měl pokud možno vhodně sledovat samotný portál. Pro tento účel byl v portálu implementován podpůrný mechanismus pro automatické odhadování důvěryhodnosti uživatelů a jimi vložených dat. Ten využívá tři faktory důvěryhodnosti - čestnost, odbornost a praxi. Čestnost je počítána zejména na základě sociální blízkosti osob (implicitní zdroj) a na základě explicitního vyjádření o čestnosti druhých. Pro výpočet sociální blízkosti mezi dvěma entitami byla zvolena variace energy spreading algoritmu [6]. Konkrétně je tato sociální blízkost na portálu získávána z vazeb mezi uživateli (jejich příslušnost ke stejné skupině/instituci, vazby typu “kolega” atd.). Dále má každý uživatel možnost explicitně vyjádřit svůj názor buď na celou stránku daného uživatele, kliknutím na ikonku ★ (můj oblíbený/důvěryhodný uživatel) nebo jen na určitou množinu informací, které na své stránce uživatel prezentuje. Konkrétně může ohodnotit jeho odborné profily a novinky. Kromě přednastavené hodnoty „nemám na informaci názor“ (-) může vyjádřit buď svůj souhlas (👍) nebo nesouhlas (👎) s prezentovaným odborným profilem resp. novinkou (pozor: v tomto případě nelze zaměňovat s obvyklou interpretací ve stylu „líbí/nelíbí“). Praxe a odbornost jsou kvantifikovány s využitím PageRank-based algoritmu na základě objevených publikací entit na webových portálech DBLP, ACM a IEEE.

V portálu jsou implementovány dva způsoby vyhledávání – fultextové vyhledávání a vyhledávání v odborných profilech, podrobně v [5]. V obou případech jsou výsledky dotazu seřazeny na základě jejich relevantnosti s přihlédnutím k portálem vypočítané míře důvěryhodnosti.

### 3.5 Služby integrovaného fóra SoSIREČR

Mezi klíčové funkčnosti portálů jako je tento patří nástroje pro komunikaci mezi uživateli. V našem portálu pro tyto účely nabízíme následující možnosti komunikace, které jsou realizovány prostředky integrovaného fóra SoSIREČR:

- Každé entitě kromě *uživatele* je přiřazen vlastní diskusní prostor, který se člení na vlákna. Uživatelé si vlákna definují podle libosti. S výjimkou *skupiny uživatel*, je diskuse veřejná. U skupiny uživatel se tak mohou do diskuse zapojit pouze členové skupiny. Moderátorem diskuse je vždy automaticky nastaven správce entity.
- Pro zjednodušení práce s inzeráty je na stránce každého uživatelem vytvořeného inzerátu tlačítko pro přímé zaslání zprávy správci inzerátu. Uživatel může případně jménem skupiny, jíž je správcem, zaslat odpověď na inzerát poptávající skupinu.
- Podobně je také pro účely usnadnění navázání komunikace mezi uživateli na stránce každého uživatele tlačítko pro zaslání pozvánky uživateli do projektu/skupiny.
- Klasické zasílání zpráv mezi uživateli portálu.

### 3.6 Data na portálu

Portál Síť IT je předem vyplněn obsahem, shromážděným v rámci projektu SoSIREČR, který je platný v den spuštění portálu pro odbornou veřejnost. Aktuálně jsou na portálu tato data: 65 firem, 449 institucí a univerzit, 241 odborných časopisů, 224 projektů, 705 studijních oborů a 23 poskytovatelů grantů.

Jeho uživatele bude jistě zajímat, jakým způsobem bude probíhat aktualizace těchto informací. Vývojový tým FEL ČVUT se snaží o to, aby, kde to bude možné, příslušné informace mohly být aktualizovány automaticky prostřednictvím datových rozhraní s využitím podpůrných konfigurovatelných modulů, nebo alespoň, aby bylo možné automaticky informovat provozovatele systému (např. prostřednictvím zprávy na



portálu nebo e-mailem) v takových případech, kdy např. na portálu odkazované stránky již nejsou platné či dostupné. A v neposlední řadě zbývá pasivní způsob informování správců systému zasláním e-mailu ke zjednání nápravy. Pokud se podíváme konkrétně na jednotlivé části obsahu, aktuální stav a naše představy jsou následující:

- **Uživatelé** – jedná se aktivní entity portálu, sami registrovaná uživatelé zodpovídají za aktuálnost údajů uvedených na svých stránkách.
- **Skupiny** – aktivní entity portálu, správci skupin odpovídají za aktuálnost informací.
- **Projekty** – informace získané o projektech podporovaných z veřejných prostředků ČR – možnost automatické aktualizace zde zatím analyzujeme. Vedle toho uživatelé mohou zakládat vlastní entity dalších projektů. Udržování jejich stránek je pak zodpovědností těchto uživatelů – správců.
- **Univerzity a vysoké školy** – implementujeme automatizovanou podporu pro aktualizaci údajů v rámci ČVUT, další subjekty budeme analyzovat.
- **Instituce, poskytovatelé grantů** – zde nepředpokládáme příliš dynamické změny, proto prozatím aktualizaci nerealizujeme.
- **Firmy** – tyto entity mohou zakládat a spravovat sami registrovaní uživatelé, ti pak také zodpovídají za aktuálnost vložených údajů. Entity firem vložené při spuštění systému bude možné předat pod správu oprávněným uživatelům vystupujícím za danou společnost.
- **Studijní obory** – možnosti aktualizace zde analyzujeme.
- **Časopisy** – možnosti aktualizace analyzujeme především u dynamičtějších informací, jako je jejich impakt faktor a jeho vývoj.
- **Inzeráty** – probíhá denní automatická aktualizace platných inzerátů poptávky po relevantních profesích úřadů práce v ČR. Vedle toho uživatelé sami mohou vkládat vlastní inzeráty typů nabídka osoby či skupiny a poptávka po osobě či po skupině. V tomto případě sami zodpovídají za aktuálnost vložených dat.

## 4 Realizace uživatelských scénářů

V této části práce uvedeme, jakým způsobem lze použít portál Síť IT k uspokojení uživatelů ve smyslu uvedených uživatelských scénářů.

### 4.1 Vyhledávání výzkumných partnerů

Regionální firma *ContractsOnline* pro nalezení informací potřebných k realizaci jejich informačního systému může použít funkce **fultextového vyhledávání**. Konkrétně následující dotazy lze realizovat vyhledáním příslušných klíčových slov s omezením hledaných výsledků na požadované entity.

- Jaké skupiny či osoby v ČR mají know-how v oblasti strojového zpracování textů, webových technologií a sémantického webu?
  - Postupné vyhledání klíčových slov: „zpracování textu“, „webové technologie“, „sémantický web“ s omezením hledaných entit na „uživatelé“ a na „uživatelské skupiny“.
- Jaké skupiny či osoby v ČR spolupracují nebo se účastní iniciativy *OpenGov.eu*?
  - Vyhledání klíčového slova: „*OpenGov*“ s omezením hledaných entit na „uživatelé“ a na „uživatelské skupiny“.
- Jsou nebo byly v ČR řešeny projekty v uvedených oblastech?



- Postupné vyhledání klíčových slov: „zpracování textu“, „webové technologie“, „sémantický web“, „OpenGov“ s omezením hledaných entit na „projekty“.

Ve všech případech nabídne portál seznam relevantních výsledků vyhledávání, ve kterém lze jedním klikem přejít na stránky příslušných osob, skupin či projektů. Zde lze z uvedených popisů usoudit, zda se skutečně jedná o hledané odborníky či projekty. Po nalezení vhodných entit může firma zahájit konzultaci svých potřeb buď zasláním zprávy přímo na portálu, nebo s využitím některého z uvedených kontaktních údajů.

Alternativně je možné dotazy lze formulovat také pomocí **vyhledávání znalostních profilů**, které nejlépe vystihují hledané znalosti. V takovém případě portál nabídne seznam entit s nejlepší shodou hledaných kategorií znalostního profilu.

## 4.2 Hledání lidských zdrojů

Katedra softwarového inženýrství na Vysoké škole informatiky může využít vyhledávání odborníků nebo inzerátů s nabídkou osob nebo skupin odborníků. Konkrétně následující dotazy lze realizovat vyhledáním příslušných klíčových slov s omezením hledaných entit.

- Kdo se zkušenostmi z práce na grantových projektech a/nebo v oblasti programování webových aplikací nabízí své služby jako J2EE programátor?
  - Vyhledání klíčových slov: „programátor J2EE“ s omezením hledaných entit na „uživatelé“ a na „uživatelské skupiny“. Nejrelevantnější výsledky budou nabídnuty mezi prvními.
- Jací jsou v ČR výzkumníci v oblasti databázového zpracování RDF dat s publikacemi v oblasti na významných konferencích, případně navíc se zkušenostmi z EU či národního výzkumného projektu v oblasti zpracování RDF dat?
  - Vyhledání klíčových slov: „databáze RDF“ s omezením hledaných entit na „uživatelé“. Nejrelevantnější výsledky budou nabídnuty mezi prvními. U nich lze proklikem na stránkách příslušných osob nalézt např. odkazy na jejich významné publikace v německém portálu DBLP<sup>1</sup>, dokument s životopisem jako přílohu nebo alespoň odkaz na osobní stránku, kde mohou být uvedeny doplňující informace.

Na rozdíl od katedře dostupných informací a běžných webových vyhledávačů lze na portálu Síť IT velice snadno vyhledat kontakty na zde registrované regionální odborníky v České republice. Přímým odkazem ze seznamu nalezených entit jsou k dispozici příslušné stránky, na kterých lze přímo **vygenerovat pozvánku do projektu**, která se odešle vybranému odborníkovi jako interní zpráva. Dále lze využít uvedených kontaktních údajů k zahájení komunikace i dalšími kanály, které jsou k dispozici (např. e-mail).

## 4.3 Propagace výzkumného zaměření a výsledků

Výzkumná skupina XRG na MFF UK může představit svůj nástroj pro navrhování a údržbu sady XML schémat vložení inzerátu na portálu Síť IT s nabídkou know-how skupiny XRG. Nejprve se její vedoucí či některý z jejích členů zaregistruje na portálu. Tím získá svou vlastní stránku, kde může umístit svou osobní prezentaci, a také může zakládat nové entity. Pro uskutečnění požadovaných úloh provede následující kroky:

- Zveřejnění nabídky svého know-how a nástroje

<sup>1</sup> The DBLP Computer Science Bibliography, <http://www.informatik.uni-trier.de/~ley/db/>

- Založí novou entitu inzerátu za skupinu XRG. V případě, že na portálu již existuje entita výzkumné skupiny XRG (např. na nejnižší úrovni organizačních složek Univerzity Karlovy), může při založení inzerátu nastavit odkaz na tuto skupinu pro informování potenciálních zájemců. V těle inzerátu lze uvést potřebné předpoklady spolupráce a požadavky na případné zájemce, nastavit dokdy bude nabídka platná atd.
- šíření nabídky potenciálním zájemcům z řad akademických pracovníků i firem
  - Pro dlouhodobější šíření povědomí o know-how je vhodnější založit novou entitu projekt se vztahem ke skupině XRG. Vedle popisných textů, klíčových slov, odkazů atp. je zde také možnost kategorizovat projekt znalostním či vědeckým profilem. Díky tomu budou výsledky dotazů potenciálních zájemců o daná témata obsahovat také odkaz na tento projekt.

Zástupce skupiny XRG může dále mezi institucemi prezentovanými na portálu procházet různá centra pro transfer technologií, která jsou nápomocna při uplatňování výsledků v praxi a vybrat si tak odpovídající centrum, se kterým se rozhodne navázat kontakt.

Také v případě jednotlivce, který nabízí svoji odbornost jiným skupinám či projektům lze použít uvedené kroky. Svou nedostatečnou sítí osobních kontaktů tak může obohatit za pomoci odborného portálu Síť IT s působností v dané oblasti, kde může velice snadno vyhledat potenciální partnery či dostávat zprávy přímo od relevantních zájemců o spolupráci.

## 5 Zamýšlená rozšíření portálu

Současná verze portálu není ještě finální. Plánujeme další rozšíření portálu, například v těchto směrech:

- Posílení regionální orientace portálu možností výběru regionu při vyhledávání z mapky.
- Integrace fulltextového vyhledávání s vyhledáváním ve znalostním/vědeckém profilu pro větší variabilitu a zpřesňování pokládaných dotazů.
- Propojení novinek s jinými blogovacími systémy.
- Možnost exportu vlastního profilu do podoby RDF dle ontologie typu FOAF (Friend of a Friend).

## Literatura

- [1] Matoušek, K. - Doležal, J.: Analýza portálů sociálních sítí pro vědu, výzkum a inovace. In Systems Integration 2010. Praha: Oeconomia, 2010, s. 214-223. ISBN 978-80-245-1660-8.
- [2] Nečaský, M.: Analýza potřeb členů sociální sítě informatiků v regionech ČR. Zpráva projektu SoSIREČR, 2011.
- [3] J. Voříšek, P. Doucek, O. Novotný. Konkurenceschopnost absolventů IT oborů VŠ a VOŠ na trhu práce v ČR. Hlavní výsledky projektu. 15.5.2007. Vysoká škola ekonomická v Praze.  
[http://www.vse.cz/media/konkurenceschopnost\\_it.pdf](http://www.vse.cz/media/konkurenceschopnost_it.pdf)
- [4] The ACM Computing Classification System (1998). ©2010.  
<http://www.acm.org/about/class/ccs98-html>. [verze ze 4. srpna 2010]
- [5] Matoušek, K., Kubalík, J., Nečaský, M.: Portál pro podporu sociální sítě informatiků v regionech ČR: Popis stavu systému. In *Hovory s informatiky 2011*. Praha: Ústav Informatiky AV ČR, v.v.i., 2011, s. 177-178. ISBN 978-80-87136-11-9.
- [6] Ziegler, C.N., Lausen, G.: Propagation Models for Trust and Distrust in Social Networks. *Information Systems Frontiers*, Volume 7, Issue 4-5, Pages 337 – 358, December 2005.