

## Odborné profily: nástroj sofistikovaného vyhledávání a sběru uživatelských požadavků

V rámci projektu SOSIREČR vzniká profesní sociální síť informatiků (tzv. portál), která nabídne kromě tradičních funkcí sociální sítě také svou hlavní „zbraň“ – tzv. **odborné profily**. Konceptem odborných profilů se portál bude lišit od tradičních sociálních sítí typu FaceBook, ale i od profesních sítí jako např. LinkedIn. Proto na odborné profily klademe v rámci projektu SOSIREČR mimořádný důraz.

Odborné profily půjde použít k široké škále popisu a vyhledávání osob, projektů a dalších entit v rámci portálu. Také poslouží samostatně ke sběru uživatelských požadavků a pro jinou zpětnou vazbu, např. formou jednoduchých webových formulářů. V tomto dokumentu chceme představit možné způsoby využití odborných profilů a zahájit veřejnou diskusi k jejich uplatnění. V kapitole 1 představíme OP po technické stránce, v kapitole 2 potom jeho potenciální praktické využití v rámci portálu i mimo něj.

### 1. Co je to odborný profil (OP)

Neformálně řečeno, odborný profil je strojově čitelný popis rozsahu znalostí v IT. Krapet formálněji, odborný profil je instance strukturovaného datového typu, kterým je nějaká hierarchie kategorií (klasifikační strom). Jako základní kategorizaci vhodnou pro popis odbornosti v IT uvažujeme ACM klasifikaci, která je všeobecně uznávaným standardem klasifikace IT oborů. Například se jí anotují veškeré publikace vydávané ACM, viz <http://www.acm.org/about/class/>. Do budoucna bychom nicméně rádi využili i jiných klasifikací, vyvinutých v rámci projektu SOSIREČR. OP je tedy navržen obecně, aby bylo možno do budoucna použít libovolný klasifikační strom. Do diskuse o nich chceme zapojit širší odbornou veřejnost, zejména prostřednictvím regionálních koordinátorů projektu.

OP slouží k numerickému ohodnocení jednotlivých kategorií dané klasifikace. Máme-li např. ACM klasifikaci - strom, OP vyrobíme tak, že každou kategorii ohodnotíme číslem 1-5. Jednička znamená „velmi málo“, pětka „nejvíce“. Konkrétní význam stupnice je samozřejmě závislý na použití, viz případové studie dále. Ohodnotit můžeme libovolný počet kategorií, resp. neohodnocené kategorie chápeme jako ohodnocené číslem 0 – „vůbec“. Protože schéma OP je strom, můžeme hodnotit kategorie nejen na nejvyšší úrovni (u ACM kategorie A-K), ale můžeme se libovolně vnořovat a ohodnocovat podkategorie na jakékoliv úrovni hierarchie, a to nezávisle na hodnocení kategorie přímo nadřazené.

*Příklad:*

*Chci-li podle ACM klasifikace popsat odborným profilem znalosti vědce, mohu ho ohodnotit jako: velmi zdatného (4) v informačních systémech (kategorie H. Information 1systems) a dobrého (3) v softwaru (kategorie D. Software). Chci-li svoje volby specifikovat podrobněji, např. druhé ohodnocení (D. Software) zjemním ohodnocením v rámci podstromu, kde zvolím 3 pro D.2 softwarové inženýrství, 4 pro D.4 operační systémy a 2 pro D.m „různé“.*



evropský  
sociální  
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,  
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání  
pro konkurenceschopnost

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

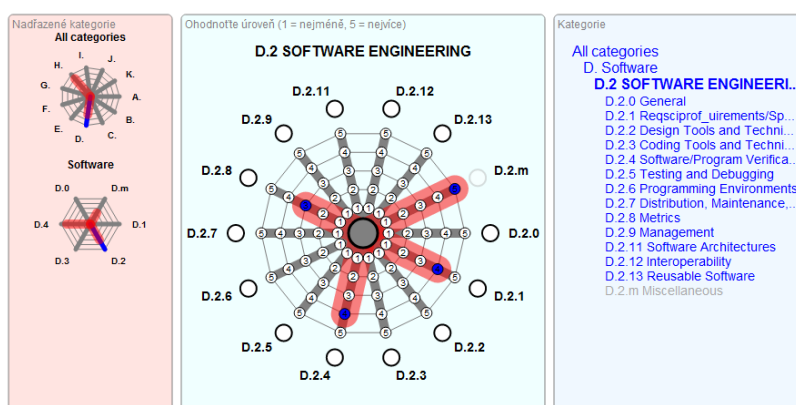
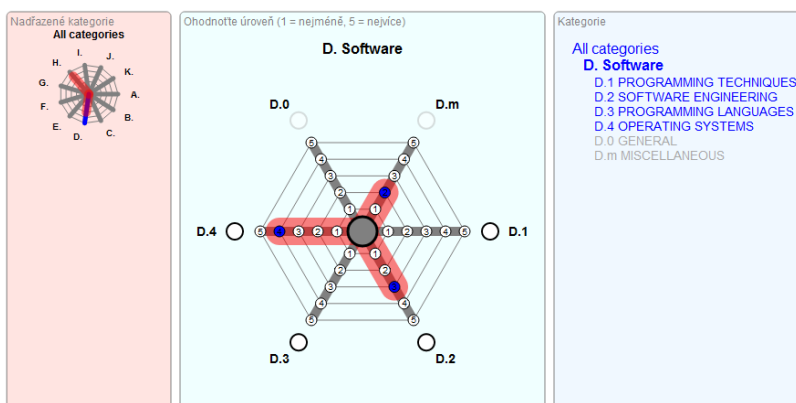
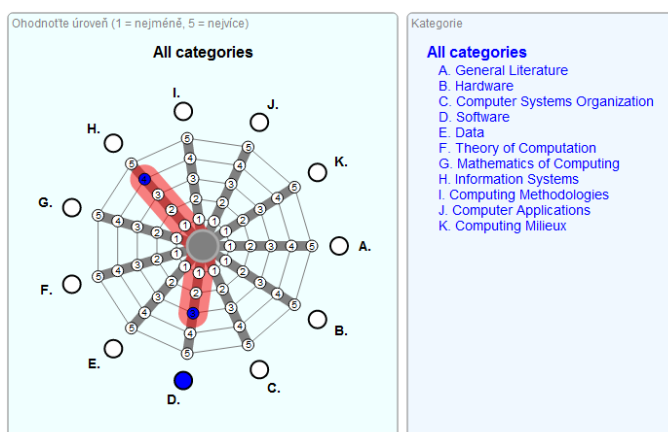
## Editace OP a vizualizace

Jelikož OP vznikají jako součást zmiňovaného portálu, tedy webové aplikace, byla vyvinuta implementace komponenty pro vizualizaci a editaci OP na webu (řešeno jako Flash aplikace). OP je vizualizován po částech – vždy je vidět pouze jeden uzel daného klasifikačního stromu. Graficky je tento uzel znázorněn jako jistý druh sloupcového grafu uspořádaného do kruhu (resp. chcete-li pavučinový graf), kde každý sloupec představuje podkategorii stromu v daném uzlu, kterou lze ohodnotit 1-5. Zároveň se na pravé straně zobrazuje textový popis k tomuto uzlu, tj. názvy kategorie a podkategorií. Aby měl uživatel přehled o ohodnocení přímo nadřazených kategorií, v levé části komponenty jsou vykresleny malé grafy příslušné všem přímo nadřazeným kategoriím. Vizualizace OP je velmi přímočará, zobrazuje se pouze to, co uživatel specificky „naklikal“, nikoliv např. agregace hodnocení podkategorií do nadřazené kategorie. Tento minimalistický přístup má svá pro a proti

– např. ohodnocení pouze podkategorií bez nutnosti hodnocení nadřazené kategorie může sloužit k „defenzivnímu“ hodnocení, kdy jsem si jist jistým specifickým podoborem, ale netroufám si hodnotit nadřazený, širší, podobor. Na druhou stranu, nerozpoznáme nadřazenou kategorii, kterou uživatel pouze zapomněl ohodnotit. Zde je jistě prostor k diskusi, jak vizualizaci a editační funkce vylepšovat, nicméně na datový typ OP tyto změny nemají vliv.

Výše zmíněný příklad zadávání znalostí vědce je zobrazen na screenshotech s komponentou

pro editaci profilů (vpravo). První obrázek ukazuje ohodnocení na kořenové úrovni, další dva pak na podúrovňích příslušejících podkategoriím D a D.2.



## Demo

Vyzkoušet si editor OP můžete na adrese [www.sosirecr.cz/forms/sciprof/](http://www.sosirecr.cz/forms/sciprof/)

## Interpretace OP a vyhledávání

Zatím jsme OP popsali pouze jako datový typ a jeho vizualizaci, bez bližší motivace jeho uplatnění. Obecně si slibujeme použití OP v mnoha aplikacích, a to jak v rámci portálu, tak mimo něj.

V první řadě, OP lze použít jako nástroj k popisu odbornosti nějaké existující entity. Entitou máme na mysli konkrétní osobu studující/pracující v IT, IT studijní program, konkurs na IT pozici ve firmě, téma diplomové práce, výzkumný tým, apod. Ke všem těmto entitám lze přiřadit jejich OP, a to jak objektivně, tak mnohokrát subjektivně, jako hodnocení uživateli.

V druhé řadě, OP lze použít k hodnocení odbornosti dané entity, pro kterou již existuje referenční OP. Mohu například definovat OP popisující které **další odbornosti** by entita (např. studijní program) měla obsahovat vzhledem k referenčnímu OP. Nebo mohu vyloženě význam stupnice OP obrátit, a **negativně reagovat** na referenční OP, tj. stupně hodnocení by pak znamenaly 0 – nekomentuji, 1 – téměř souhlasím, 5 – zcela nesouhlasím.

V obou případech lze OP použít jako nástroj ke sběru informací, resp. zpětné vazby od uživatelů.

Co s nasbíranými OP? Uvažujeme dvě možnosti, které jsou nyní stále předmětem diskuse. První je **statistické zpracování** OP a vizualizace, resp. agregace více OP do jednoho reprezentativního. Druhou je **vyhledávání podobných OP** (již uvnitř portálu). Až s touto druhou funkcí se význam OP projeví naplno, neboť se s ní částečně naplní sny o sémantickém webu. Sémantický web původně sliboval významové vyhledávání v plných textech, tj. strojové porozumění informacím na webu. To se ovšem ukázalo jako cíl zcela nerealizovatelný bez pomoci samotných tvůrců webových stránek. Vyhledávání pomocí OP je sice oproti “chimérickému” sémantickému webu úzce zaměřené, nicméně zároveň opravdu vnáší do vyhledávání právě potřebnou sémantiku.

## 2. Případové studie užití odborných profilů

V této kapitole nastíníme několik příkladových studií poukazujících na potenciál odborných profilů v praxi (na portálu).

### Student

Na portálu bude seznam vysokých škol poskytujících studijní programy v informatice a IT. Zájemce o studium může:

- 1) jednoduše vizualizovat jednotlivé obory pomocí OP a opticky si je srovnat
- 2) srovnat obory (nástrojem na portálu) tak, aby viděl, který obor je v čem jedinečný
- 3) přímo vyhledat obor, který je nejpodobnější jeho představě (nakliká si vlastní OP)

Pokud už studuje, může si stejným způsobem vybrat téma diplomové práce. Pokud absoluuje jako Bc a nechce pokračovat ve studiu na stávající škole, může si vybrat navazující studium na jiné škole na základě svého Bc profilu a požadavků na Mgr program na jiné škole.

## Absolvent

Na webu existují inzertní servery nabízející nabídky zaměstnání, jako např. [www.jobs.cz](http://www.jobs.cz). Vyhledávání v těchto inzerátech je ale nepřesné, a vyžaduje opakovaná hledání tak, jak se nabídky objevují. Kategorie zde sice existují, ale jsou příliš hrubé (např. jen IT). Náš portál bude rovněž podporovat inzerci a management lidských zdrojů, nicméně pracovní nabídky v IT budou popsány (kromě textového popisu) také OP, což absolventovi umožní velmi přesné vyhledávání.

Absolvent nemusí explicitně vyhledávat pozice opakovaným kladením dotazu, stačí, když si do svého osobního profilu na portálu uloží svůj OP a přihlásí se k odběru perspektivních pracovních nabídek. Ty mu budou chodit automaticky na základě dostatečné podobnosti OP nabídky s OP absolventa.

## Akademik /Vědec

Ačkoliv vědeckých týmů praktikujících základní výzkum v IT je v České republice jistě méně, než IT firem, či studentů, je s podivem, že mnohé týmy, které se věnují příbuzné problematice o sobě neví. Spolupráce vědeckých týmů v ČR stojí spíše na osobních kontaktech, než na cíleném vyhledávání partnerů. Zde by OP mohly pomoci při hledání partnerského týmu pro základní výzkum. Například předpokládejme, že biologická laboratoř potřebuje do svého projektu zapojit pracoviště, které jí pomůže vyřešit otázky výkonnosti jejich specifických algoritmů. Pomocí odborných profilů zveřejněných jednotlivými týmy může laboratoř takové pracoviště najít snáze, než zdlouhavým prohledáváním textu na ad-hoc vyrobených webových stránkách konkrétních týmů.

Pedagog může pomocí OP „nálepkovat“ svoje předměty, témata diplomových prací, požadavky na studenty do projektů, apod. Zároveň může tytéž záležitosti sledovat u svých kolegů a srovnávat se. Garanti studijních programů na fakultách mohou sledovat profilaci různých IT programů v rámci ČR a reagovat na poptávku ze strany studentů i firem.

## Firma

Podobně jako absolvent hledá práci pomocí OP, může IT firma, resp. její personální oddělení, vyhledávat vhodné zájemce o práci. Místo tradičního pořádání konkurzů, kterými 90% zájemců neprojde, si může firma vytipovat na portálu ty zájemce o práci, kteří se ohodnotili OP odpovídajícím představě firmy. Na osobní pohovor pak firma pozve řádově méně kandidátů, jelikož těch zcela nevhodných se zbavila již předvýběrem na portálu. Pochopitelně že firma může vyhledávat pouze v takovém seznamu osob, které explicitně zveřejní svůj OP a dají tedy souhlas k jeho publikaci.

Dalším užitečným nástrojem pro firmy je hledání akademického partnera pro aplikovaný výzkum – zde je situace obdobná jako v případě hledání partnera pro základní výzkum, viz výše. V současné době se taková firma musí spolehnout na Google a ad-hoc vytvořené stránky výzkumných týmů, případně si zaplatit službu (nebo státní agenturu), která to udělá za ně. Pokud budou mít výzkumné týmy svoje OP zveřejněné na portálu, vyhledávání partnera pro výzkum se tím zjednoduší.

V neposlední řadě mohou s podporou OP zahájit firmy věcný dialog s vysokými školami. Z hlediska hodnocení absolventů jsou to právě firmy, které „vidí“, zda je absolvent použitelný v komeční praxi lépe či hůře. Pokud mají firmy negativní zkušenosti s absolventy ve stále stejných kritériích, mohou poukázat na nedostatky ve studijních programech jednotlivých fakult a samy pomocí OP specifikovat vlastní představu o odborném profilu absolventa. Školám taková zpětná vazba jistě poskytne užitečnou informaci o poptávce po IT vzdělání v konkrétním regionu.