

**Název projektu:**

## System pro správu a učení preferencí PreSys

**Vedoucí:** Alan Eckhardt a Peter Vojtáš (eckhardt,vojtas@ksi.mff.cuni.cz)

**Počet řešitelů:** 4-5

**Jazyk, OS:** Preferovaná Java, ale může být i jiná.

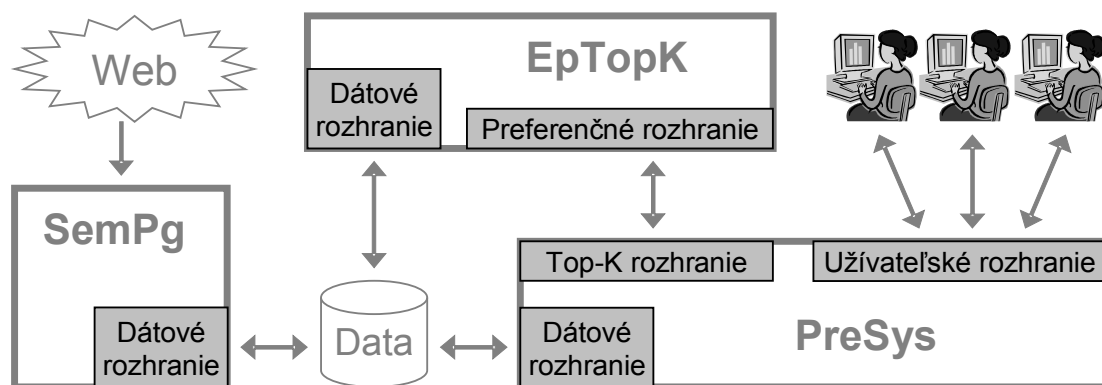
**Termín dokončení:** Za 9 měsíců od vypsání projektu.

### Motivace:

V souvislosti s výzkumem Sémantického webu a především myšlenkou jeho postupné *sémantizace* [1] potřebujeme experimentální platformu (dále W2U – Web to User) pro:

1. vývoj nástrojů pro učení se modelu uživatelských preferencí a
2. aplikaci těchto nástrojů v systému prezentace a hledání dat na webu.

Souběžně s tímto projektem jsou vypsány dva navazující, SemPg a EpTopK, které zajišťují další části celého systému W2U. Bude nutné poskytovat rozhraní pro poskytování uživatelských preferencí pro projekt EpTopK. V případě souběžného plnění je možná spolupráce jednotlivých projektů.



Obr.1 Schéma celého systému W2U

### Popis projektu:

Preferenční systém PreSys je součástí systému W2U, který zpracovává data od webu až k uživateli. PreSys bude získávat data od uživatele, data z databáze (třeba i sémantická data) a jeho úkolem bude vykonstruovat obecný model preferencí, pro každého uživatele zvlášť. Tento obecný model bude schopný říci o každém objektu, jak se bude uživateli líbit či nelíbit.

### Vstupy:

- 1) Explicitní preference od uživatele  
Uživatel zadá přímo, že chce velké a levné notebooky
- 2) Explicitně zadaná hodnocení  
Uživatel řekne, že notebook A se mu líbí na 5 hvězdiček, B pouze na 3
- 3) Implicitní preference  
Uživatel se pohybuje na stránkách a vykonává běžné akce

Tyto vstupy jsou zpracovány a z nich je vykonstruován obecný model.

Výstupy:

Preferenční systém bude umět

- 1) Poskytnout model uživatelských preferencí pro získání dat prezentovaných uživateli (navazující projekt)
- 2) Poskytnout informace o jednotlivých preferencích pro úpravu uživatelského rozhraní

Úkolem řešitelů bude vymyslet modulární strukturu rozhraní mezi jednotlivými částmi systému, implementovat zpracování vstupů i generování výstupů. Součástí projektu bude webové rozhraní pro zadávání všech tří druhů vstupů. Řešitelé by měli v co největší míře využít existující volně dostupné softwarové komponenty.

Systém by tedy měl na základě kombinací vstupů 1-3 vytvořit obecný model preferencí, který bude schopen „počítat“ preference pro nové objekty. To je typicky vnímáno jako „učení preferencí“, jako jistá odrůda strojového učení nebo dolování znalostí z dat.

Předávání informací o akcích uživatele by mělo být co možná nejkompexnější, aby zachycovalo důvody, proč uživatel dělá to, co dělá. Např. se může hodit informace o tom, které objekty jsou právě zobrazené (nebo jiný kontext), protože to také ovlivní uživatelské preference.

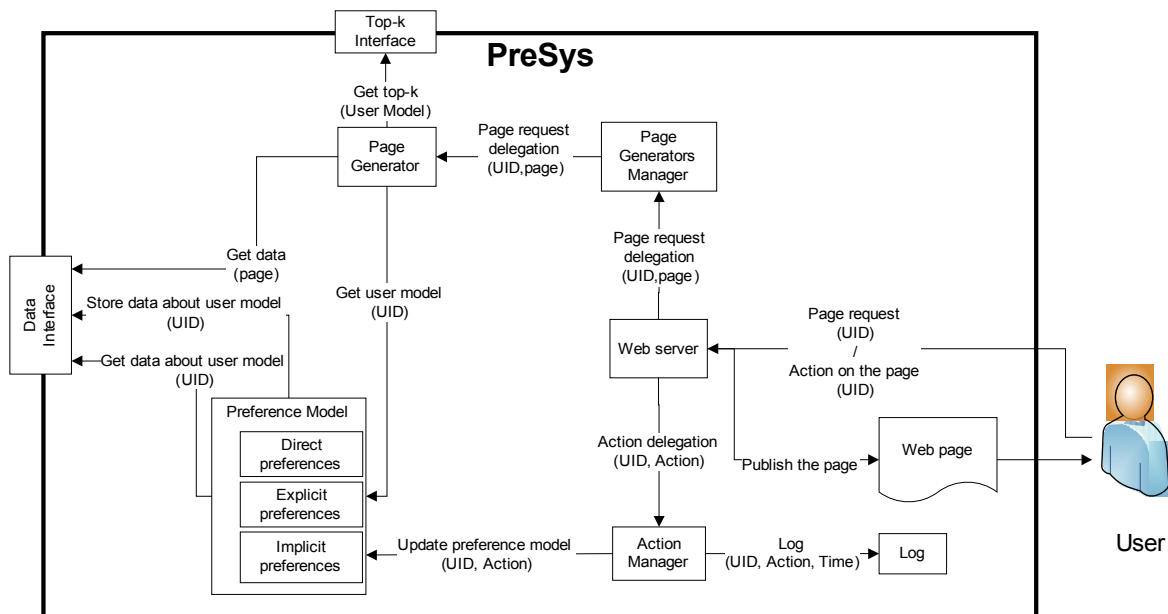
Pomocí obecného modelu poté budou nalezeny objekty, které uživateli nejvíce vyhovují.

Součástí projektu bude také implementace uživatelského rozhraní v podobě webové stránky se seznamem objektů. Na této stránce bude možno vybírat konkrétní objekty a zobrazovat jejich detaily, filtrovat množinu objektů, nechat si doporučit objekty, hodnotit objekt pomocí nějaké škály a explicitně zadávat preference.

Pro tuto funkčnost bude třeba vyrobit specifikaci Datového rozhraní ke zdroji dat.

Doporučování objektů bude zajišťovat naivní implementace některého top-k algoritmu. Získávání dat bude také implementováno cestou nejmenšího odporu naivní implementací nějakého rozhraní. Tyto části systému budou pokryty jinými projekty.

Pro top-k algoritmus bude ovšem třeba vyrobit rozhraní, za které se schová buď naivní implementace nebo projekt EpTopK



Obr.2 Schéma podsystému PreSys

Na obrázku jsou znázorněny jednotlivé komponenty systému a jejich vzájemná komunikace spolu s parametry, které si budou komponenty předávat.

Toto schéma není fixní – v prvotní fázi bude třeba jej analyzovat do detailů a upravit případné nekonzistence. Funkčnost jednotlivých komponent bude přesněji podrobně popsána až ve fázi specifikace, nicméně základní funkčnost by měla být zřejmá z prvních konzultací.

#### Reference:

[1] Jan Dědek, Alan Eckhardt, Leo Galamboš, Peter Vojtáš: **Sémantický Web**, in DATAKON 2008, Brno, ISBN 978-80-7355-081-3, pp. 12-30, 2008.