

České vysoké učení technické v Praze
Fakulta elektrotechnická

Household Intelligent Assistant



Autoři: Jiří Burant
Jakub Drápela
Martin Klučka
Petr Kovář
Jakub Konrád
Pavel Trutman

Studijní obor: Kybernetika a robotika

Datum vypracování: 16. března 2016

Popis projektu, motivace

Jako lidi si klademe různé otázky. Abychom mohli ve světě rozumě fungovat potřebujeme znát odpovědi na ně. Od té doby, co se začaly informace zaznamenávat lze odpovědi nalézt v záznamech. V nedávné historii lidé prahnoucí po informacích ve velkém kupovali knihy, noviny, jízdní řády, prostě média se žádaným obsahem.

Doba pokročila. My jsme se posunuly do doby, kde jsou téměř všechny informace uloženy v elektronické podobě. Je běžné k nim přistupovat pomocí moderních přístrojů. Téměř každý aktivní člověk vlastní alespoň jednu z věcí, jako je počítač, notebook, chytrý mobilní telefon, tablet, iPad a jiné. S využitím toho vybavení máme značně usnadněnou cestu k získání potřebné informace. Stačí mít u sebe správnou aplikaci, přístup na internet a například dopravní spojení nalezneme do minutky. Limituje nás ovšem schopnost manipulace, což při troše nešikovnosti dokáže člověka dost rozčílit.

My se snažíme jít o krok vpřed. Ruční manipulaci obejít a informace získávat pouze pomocí našeho vlastního hlasu, což ovšem obsahuje širokou škálu požadavků.

Detailní specifikace

V našem projektu se zaměříme na vytvoření dialogového systému pro použití v běžné domácnosti či kanceláři. Tedy systému, který dovede sám odpovídat na kladené otázky z limitované oblasti počasí, dopravy, obecných informací atd. Systém bude neustále čekat na aktivační slovo a potom bude schopen zodpovědět určitou škálu otázek. K realizaci projektu budeme využívat již existující aplikace.

V současné době existuje řada aplikací, které fungují na podobném principu. Uveďme například aplikaci **Cortana** od firmy Microsoft. Cortana, podobně jako další aplikace od firmy Apple Siri, je označována jako osobní asistent. Lze ji použít pro širokou škálu úkonů. Například hlasové ovládání přijímače nebo vyhledávání informací na internetu. Její nevýhodou je použití pouze pod operačními systémy Windows.

Jako další lze uvést open-source aplikaci **Jasper** v jazyce Python. Tato aplikace umožňuje využití hlasu pro získání informací. Můžeme ji také použít v inteligentním bydlení a v dalších situacích. Jelikož je volně dostupná, lze k ní libovolně přidávat nové moduly a její použití ještě rozšířit. Stává se tak pro nás vhodnou inspirací.

Za zmínku ještě určitě stojí projekt **Amazon echo** od firmy Amazon. Tento osobní asistent z titulní strany může přehrávat muziku, odpovídat

na otázky, vyhledávat zprávy, zjišťovat počasí, vytvářet seznamy a mnoho dalších věcí.

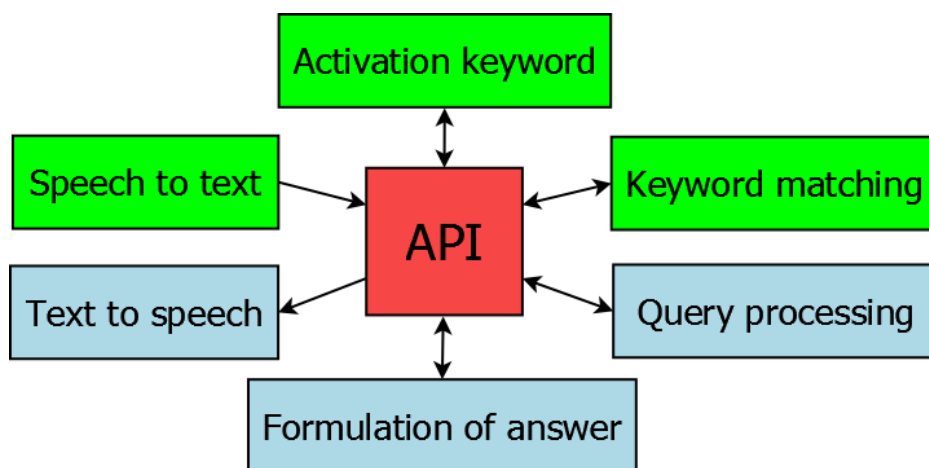
My se budeme snažit o to sestavit vlastní systém s podobnými vlastnostmi jako již vzniklé projekty.

Popis navrženého řešení

Při popisu návrhu řešení aplikace budeme postupovat tak, jak bude rekce na mluvené slovo postupně prostupovat mezi moduly, až nakonec vznikne finální řečená odpověď. Blokové schéma aplikace je zobrazeno na obrázku 2.

V prvotní situaci bude aplikace v režimu čekání na aktivační slovo. Na rozpoznání aktivačního slova bude vyhrazen jedem samostatný modul. Aktivační slovo by mělo být snadno oddělitelné od ostatních. Ukazuje se vhodné zvolit tříslabičné slovo s výraznými písmeny jako je např. x.

Po zaznění aktivačního slova uživatelem se aktivuje blok pro rozpoznání mluveného slova. Vhodnou robustní aplikací pro tento problém je třeba vybrat. Experimentovat budeme s *wit.ai* a *PocketSphinx*.



Obrázek 1: Blokový diagram navrženého řešení projektu Household Intelligent Assistant

Ze získaného textu je třeba vybrat klíčová slova, která definují otázku v různých formulacích. Klíčová již rozezná například již zmiňovaná aplikace *PocketSphinx*.

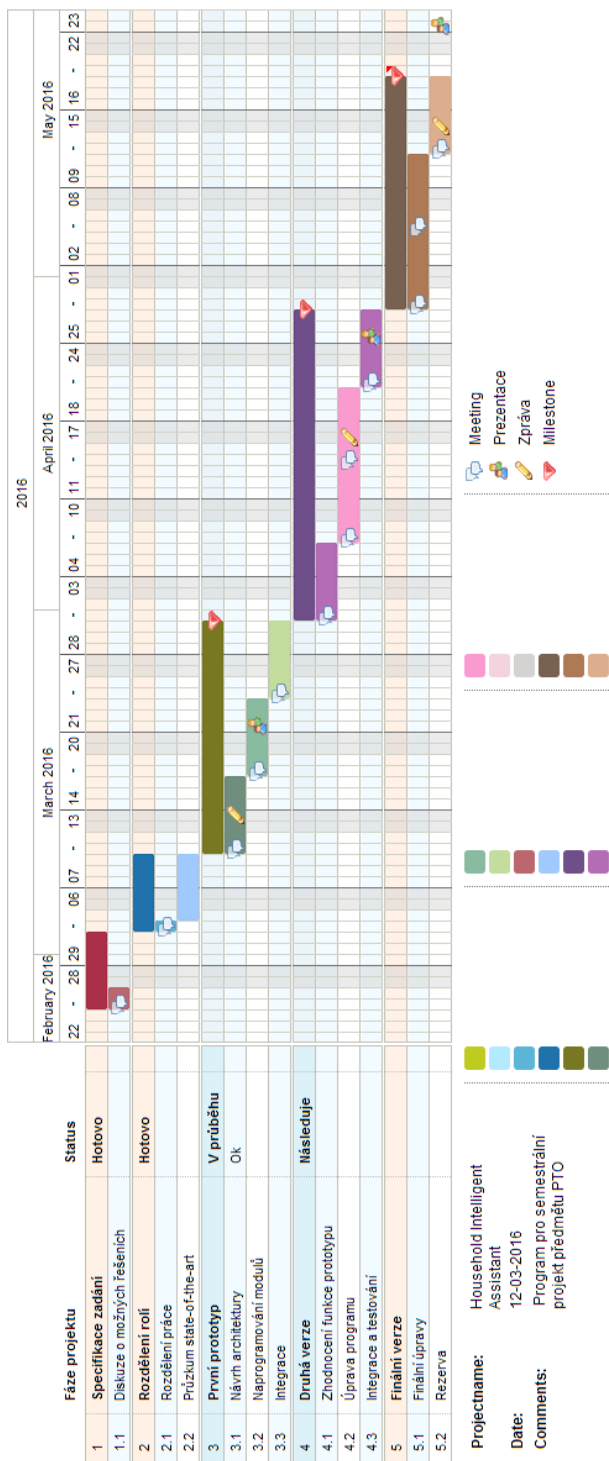
Získaná klíčová slova jsou vstupní hodnotou do modulu, jenž získá potřebné informace z internetu. Úlohy budou rozdělené na jednotlivé sub-

systémy. Jeden subsystém bude například pouze na zpracování odpovědi na počasí přes aplikaci *forecast.io*.

Odpověď kterou dostaneme je třeba přeformulovat do srozumitelné odpovědi podle předem zadaných vzorů.

V poslední řadě využijeme aplikaci k převedení formulované odpovědi do strojově mluvené řeči. Prvotní aplikací pro tento může být jednoduchý "*The Festival Speech Synthesis System*".

Plánování projektu (Ganttův diagram, úkoly, milníky)



Obrázek 2: Ganttův diagram s milníky a úkoly.

Analýza rizik a krizové plány

Předběžné výsledky

Zdroje: <http://static01.nytimes.com/images/2015/06/25/business/GADGETWISE/GADGETWISE-master675.jpg>