

## Záměr semestrálního projektu

Účelem projektu je prakticky si vyzkoušet použití Pythonu při realizaci nějakého netriviálního úkolu. Python patří do skupiny tzv. univerzálních programovacích jazyků – jazyků určených pro všeobecné účely, bez omezení domény aplikace. Téma projektu proto může být v podstatě jakékoliv.

### Úkol

Cílem je **formulovat představu o konkrétním úkolu řešeném za použití Pythonu**. Je zbytečné znovu “vymýšlet kolo”, je naopak vhodné využít (a případně si přizpůsobit) již existující implementace řešení v externích knihovnách, které jsou dobře dokumentované – není tedy žádoucí (ani efektivní) používat pouze “čistý” Python.

Obecná témata, do kterých může projekt spadat:

- **Projekt typu aplikace** – vytvoření netriviální aplikace, nástroje/utility jakéhokoliv zaměření
- **Projekt typu práce s daty**
  - *Analýza*
    - využití automatizace v empirickém výzkumu – předpokládá se, že data budou k dispozici v “rozumném” formátu
    - např. může být užitečné pro zpracování části závěrečné práce nebo projektu z předmětu, jehož náplní je provádět experimenty za použití kvantitativních metod
  - *Tvorba datasetu*
    - využití automatizace při sestavení datasetu pro různé účely analýz – sběr a kombinace obtížně dostupných dat z různých zdrojů (např. web scraping)
    - např. může být užitečné pro další zpracování v různých empirických výzkumech, tj. při získávání zajímavých informací o zkoumaném problému (a tedy opět sloužit k řešení části závěrečné práce nebo projektu v jiném studijním předmětu)
  - oba výše zmíněné “podtypy” lze kombinovat

Ve formulaci záměru projektu budou zmíněny podstatné informace s ohledem na typ projektu (viz dále):

- **Krátký název** – stručně vystihující, co bude výsledkem (podrobněji pak v popisu)
- **Členové řešitelského kolektivu** – záměr odevzdává jeden člen řešitelského kolektivu, uvede další členy
- **Popis/návrh projektu**
  - podrobnější specifikace cíle – co bude řešeno, co má být výsledkem
  - zapojení členů týmu – kdo bude řešit jakou část výsledného celku
  - plánovaný seznam knihoven či jiných hotových řešení a způsobu jejich použití

- lze uvést např. i v rámci řešení jakého “většího” úkolu či projektu (např. v nějaké firmě nebo jiné aktivitě) to pomůže nebo na jaké již existující/fungující řešení to navazuje (jak jej vylepšuje, rozšiřuje apod.)

## Odevzdání

Výše uvedené informace je nutné odevzdat ve formě souboru .txt používajícího kódování UTF-8 nebo Windows-1250 prostřednictvím UIS – Odevzdávárna. Poslední možný termín odevzdání je určen termínem uzavření Odevzdávárny, je nutné si dát pozor na to, aby soubor byl opravdu odevzdán, ne pouze nahrán v Odevzdávárně. Po uplynutí termínu uzavření Odevzdávárny není možné již soubory vkládat, ale ani nahrané a neodevzdané soubory odevzdat!

Termíny pro případné opravy, včetně důvodů vrácení k přepracování, zpřesnění atd., budou následně určeny novým datem uzavření odevzdávárny stanoveným při vrácení projektu (tzn. individuálně těm, kteří mají úkol vrácený k přepracování). Neodevzdání návrhu tématu projektu představuje automaticky nesplnění podmínek “malého” zápočtu, nelze se pak přihlásit na zkoušku (resp. závěrečný “velký” zápočet), pokud nebyl získán “malý” zápočet.

## Vybrané projekty z minulých let pro inspiraci

### Projekt typu aplikace

- Implementace vzdáleného ovládání obrazovky (Remote Desktop) – aplikace client-server na komunikaci, využití WebSocket technologie
- Aplikace pro online adopci psů a koček ve frameworku Django
- Aplikace na třídění fotek v adresáři podle různých metadat
- Parser harmonogramu v UIS, možnost uložení události do kalendáře
- Evidence a statistika plnění úkolů
- Digitalizace vlastnoručně udělané stolní mapové strategické hry
- Chatbot, který "radí" s problémy ohledně aplikace MS TEAMS
- Testovací prostředí pro existující REST API ve firmě Phonexia

### Projekt typu práce s daty

- **Analýza**
  - Superliga florbalu (statistiky hráčů, týmů, rozhodčích)
  - Invia – cestovní agentura (porovnání cestovních kanceláří, kam lidi nejvíce jezdí, jaké cestovky nejvíce využívají, kolik ročně utratí za dovolené...)
  - Spotify analýza – Top Spotify Songs 2010–2019. Analýza údajů, jaké žánry jsou nejoblíbenější a jak se oblíbenost mění, nejoblíbenější interpreti v závislosti na žánru, porovnání na základě bpm, danceability atp.
  - Použití metod text miningu na nestrukturovaný text – analýza tématu, četnost slov a jiné...

- Přehled prodejů videoher, vytvoření statistik v textové i grafické podobě (top žánry v zemích, vývoj prodeje napříč časem – podíl platforem, vydavatelů, žánrů, prodejnosti na jednotlivých trzích...)
- Srovnání změn teplot a množství srážek v letech 2000–2010 na území ČR (portál ČHMÚ), určení období s nejvyššími teplotami a nejvyšším počtem srážek.
- **Tvorba datasetu**
  - Scrapping laptopů (ze stránek více elektro e-shopů) a jejich parametrů, zpracování výsledného datasetu a výběr optimálního laptopu podle kritérií
  - Wine Quality scraping
  - Scraping archivu stránek ChallengeTV
  - Sběr dat o florbalových holích + výběr, agregace, zobrazení ukazatelů