RÚIAN Downloader

*Příloha k průběžné zprávě za rok 11/2013 – 5/2014*

Číslo projektu: TB01CUZK004

Název projektu: Výzkum uplatnění závěrů projektu *e*Content*plus* s názvem EURADIN v podmínkách RÚIAN

Předkládá:

*Název organizace: Výzkumný ústav geodetický, topografický   
 a kartografický, v.v.i*.

*Jméno řešitele: Ing. Radek Augustýn*

**Obsah**

[Základní informace 3](#_Toc400881894)

[Instalace 4](#_Toc400881895)

[Nároky a systém 4](#_Toc400881896)

[Modul RUIANDownloader 4](#_Toc400881897)

[Konfigurace 5](#_Toc400881898)

[Adresář se staženými daty DataDir 5](#_Toc400881899)

[Rozbalování archivů po stažení UncompressDownloadedFiles 5](#_Toc400881900)

[Nastavení rozsahu stahování DownloadFullDatabase 5](#_Toc400881901)

[Spouštění importu do databáze RunImporter 5](#_Toc400881902)

[Automatické spouštění jednou denně AutomaticDownloadTime 5](#_Toc400881903)

[Zamezení načítání historických dat IgnoreHistoricalData 5](#_Toc400881904)

[Příklad nastaveného konfiguračního souboru 6](#_Toc400881905)

[Nastavení parametrů v příkazovém řádku 6](#_Toc400881906)

[Spouštění modulu RUIANDownloader 7](#_Toc400881907)

[Automatické stahování aktualizací databáze RÚIAN 8](#_Toc400881908)

[Operační systém Microsoft Windows 8](#_Toc400881909)

[Operační systém Unix, (Linux) 8](#_Toc400881910)

[Informace o výsledcích a průběhu stahování 10](#_Toc400881911)

[Index.html 10](#_Toc400881912)

[RUIANDownlaoder.log 11](#_Toc400881913)

[Použité zkratky 12](#_Toc400881914)

# Základní informace

RÚIAN Downloader je jednoduchý nástroj určený k automatickému stahování dat z Veřejného dálkového přístupu (dále jen VDP) v pravidelných intervalech. Aplikace VDP umožňuje bezplatný přístup k datům registru územní identifikace, adres a nemovitostí. Pomocí jednoduché konfigurace umožňuje stahovat jak plnou datovou sadu, tak pouze aktualizační data od poslední aktualizace. Stažená data jsou rozbalena z archivu.gz do souboru XML ve výměnném formátu RÚIAN a připravena k importu do geodatabáze. Modul může být nainstalován a provozován samostatně, případně nakonfigurován ke spouštění dalšího nástroje RÚIAN Importer. V této konfiguraci poskytuje ucelenou funkcionalitu zabezpečující aktuálnost lokální kopie databáze RÚIAN v prostředí internetu. Jestliže je kopie databáze umístěna v oddělené síti, umožňuje po ukončení stahování zaslat upozornění na E-mail o vhodnosti přenést data a provést aktualizaci.



Obrázek Automatické stahování dat RÚIAN Downloaderem

# Instalace

Knihovna RÚIAN Downloader je součástí sady nástrojů RÚIAN Toolbox, jejíž aktuální verze je dostupná na <https://github.com/vugtk21/RUIANToolbox/archive/master.zip>, po stažení a rozbalení archivu ji najdeme ve složce RUIANDownloader.

## Nároky a systém

|  |  |
| --- | --- |
| Hardwarové požadavky: | nejsou |
| Operační systém: | Nejsou, kterýkoli OS podporující instalaci prostředí Python (Windows, Linux, iOS, Android atd) |
| Softwarové požadavky: | Python 2.7 a vyšší (testováno na verzi 2.7.5) |
| Paměťové nároky: | nejsou |

## Modul RUIANDownloader

|  |  |
| --- | --- |
| Soubory | Obsah |
| RUIANDownlaoder.py | Vlastní skript modulu umožňující stahování z VDP |
| RUIANDownlaoder.cfg | Konfigurační soubor pro nastavení parametrů stahování |
| RUIANDownlaoder.log | Soubor s výstupy a hlášeními průběhu stahování. |

Pozn. Knihovnu je možné používat jako samostatný celek, stačí vytvořit kopii složky RUIANDownloader.

# Konfigurace

Konfigurace parametrů se provede nastavením jednotlivých hodnot v konfiguračním souboru *RUIANDownlaoder.cfg*.

|  |  |
| --- | --- |
| Parametr | Význam |
| DataDir | Adresář se staženými z VDP. |
| UncompressDownloadedFiles | True, jestliže chceme archiv po stažení rozbalit |
| DownloadFullDatabase | True, jestliže chceme pokaždé stahovat obsah celé databáze |
| RunImporter | Zapíná spouštění importu do databáze po stažení dat. |
| AutomaticDownloadTime | Nastavuje čas automatického stahování dat. |
| DownloadURLs | Parametry stahování z VDP |
| IgnoreHistoricalData | Pokud je nastaveno na True, při načítání stavových dat ignoruje data historická data. |

## Adresář se staženými daty DataDir

Pomocí parametru DataDir určíme adresář, do kterého budou uloženy stahované soubory z VDP a další pomocné soubory pro zabezpečení stahování aktualizací.

## Rozbalování archivů po stažení UncompressDownloadedFiles

Parametr UncompressDownloadedFiles umožní nastavit, že budou data z VDP po stažení rozbalena z archivu .gz do souboru XML ve výměnném formátu RÚIAN.

## Nastavení rozsahu stahování DownloadFullDatabase

Data RÚIAN nejsou svým objemem příliš rozsáhlá, nicméně při stahování pomocí pomalého připojení je vhodné využít možnosti stahovat aktualizační balíčky.

Rozsah stahování nastavíme pomocí parametru DownloadFullDatabase. jestliže je jeho hodnota nastavena na True, tak se každým spuštěním modulu RUIANDownloader stáhne kompletní obsah databáze RÚIAN z VDP. Jestliže je hodnota nastavena na False, je nejdříve stažena celá databáze, následně při každém dalším spuštění modulu RUIANDownloader se stáhne aktualizační balíček od poslední aktualizace, případně dojde ke stažení celé databáze.

## Spouštění importu do databáze RunImporter

Tento parametr, pokud je nastaven na True způsobí, že se po stažení souborů vyvolá program RÚIAN Importer, který načtená data importuje do databáze dle nastavení.

## Automatické spouštění jednou denně AutomaticDownloadTime

Data RÚIAN se aktualizují jednou denně, proto je vhodné databázi jednou denně aktualizovat. Pokud není hodnota tohoto parametru nastavena, nebude se aktualizace provádět.

## Zamezení načítání historických dat IgnoreHistoricalData

Z VDP je možno kromě aktuálních stavových dat stahovat také data za poslední dva měsíce. Pokud je IgnoreHistoricalData nastaveno na True, načítají se pouze aktuální stavová data.

## Příklad nastaveného konfiguračního souboru

Stažená data budou ukládána do složky C:\Users\raugustyn\Desktop\Euradin\, po stažení rozbalena a stažený soubor vymazán. Budou pouze stahovány aktualizační balíčky ve dvě hodiny v noci a po importu bude spouštěn import do databáze.

RUIANDownlaoder.cfg:

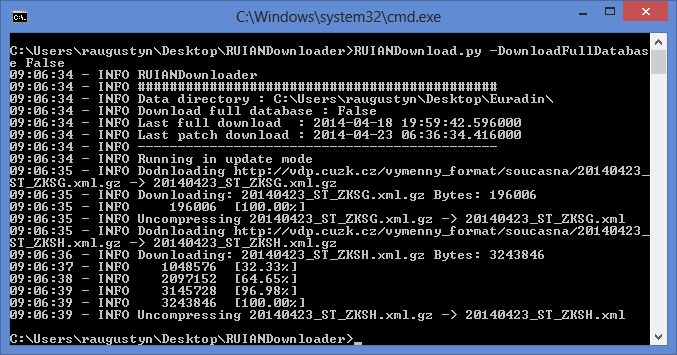
DataDir=C:\Users\raugustyn\Desktop\Euradin\  
UncompressDownloadedFiles=True  
DownloadFullDatabase=False  
RunImporter=False  
AutomaticDownloadTime=02:00:00  
IgnoreHistoricalData=True  
DownloadURLs=http://vdp.cuzk.cz/vdp/ruian/vymennyformat/vyhledej?vf.pu=S&\_vf.pu=on&\_vf.pu=on&vf.cr=U&vf.up=ST&vf.ds=K&\_vf.vu=on&vf.vu=G&\_vf.vu=on&vf.vu=H&\_vf.vu=on&\_vf.vu=on&search=Vyhledat

## Nastavení parametrů v příkazovém řádku

Pokud chceme modul RUIANDownlaoder.py využívat jinými knihovnami a nastavovat parametry z příkazového řádku, mají stejné názvy a hodnoty jako obsah souboru RUIANDownlaoder.cfg. Pokud parametr není nalezen v příkazovém řádku, bude použita hodnota nastavení v konfiguračním souboru.

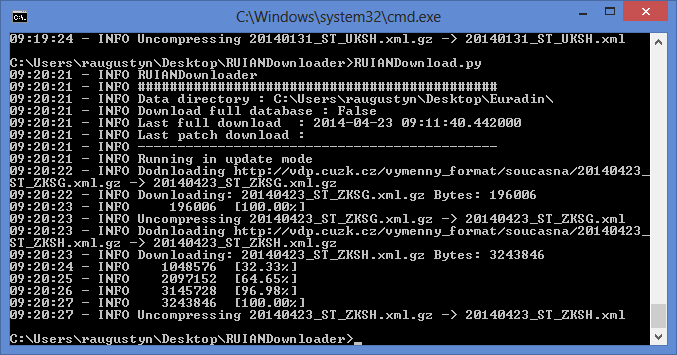
|  |  |
| --- | --- |
| Parametr | Význam |
| DataDir | Adresář, do kterého budou data RÚIAN stahována. Může obsahovat celou cestu, případně pouze relativní cestu k adresáři RUIANDownloader.  Implicitní hodnota je ..\DownloadedData, tj. adresář DownloadedData v RUIANToolbox |
| UncompressDownloadedFiles | Data RÚIAN jsou za účelem urychlení stahování komprimována do archivu GZ. Jestliže chceme archiv po stažení rozbalit, nastavíme hodnotu UncompressDownloadedFiles na True.  Implicitní hodnota je False |
| DownloadFullDatabase | Jestliže chceme pokaždé stahovat obsah celé databáze, nastavíme na True. Hodnota False znamená, že se budou stahovat aktualizace od posledního stažení stavových dat, případně posledního stažení aktualizací.   Jestliže je adresář DataDir prázdný, je nejdříve stažen obsah celé databáze a zároveň jsou staženy všechny aktualizace. |
| RunImporter | Tento parametr umožní import dat do databáze po stažení dat, pokud je nastaven na True.  Nejdříve je nutné oživit modul RÚIAN Importer, viz dokumentace. |
| DownloadURLs | Maska pro stahování stavové databáze a aktualizací, oddělené středníkem. Nejjednodušší je vyhledat data na http://vdp.cuzk.cz/vdp/ruian/vymennyformat/vyhledej a zkopírovat vygenerovanou adresu URL. |
| IgnoreHistoricalData | Tato volba zabezpečí, že při stahování stavových dat (celé databáze) se bude stahovat pouze nejnovější stav a archivní databáze budou ignorovány. |

Příklad stažení aktualizace nastavením parametru v příkazovém řádku:



# Spouštění modulu RUIANDownloader

Modul spouštíme přímo pomocí manažera souborů, v příkazové řádce, případně můžeme vytvořit zástupce na ploše spuštěním modulu **RUIANDownloader.py**



## Obsah staženého adresáře

Do adresáře se staženými daty jsou kromě dat RÚIAN uloženy také stavové informace pro další moduly RÚIAN Toolboxu.

Například po stažení dat a aktualizací 9.11.2014 se vytvoří:

|  |  |
| --- | --- |
| Soubor | Obsah |
| Download\_2014.11.09.txt | Seznam stavových dat stažených dle nastavení. |
| info.txt | Podrobné informace o stavu stahování. |
| Patch\_2014.11.09.txt | Seznam aktualizací stažených dle nastavení. |
| Soubory \*.gz nebo \*.xml | Stažená, případně rozbalená data RÚIAN |



# Automatické stahování aktualizací databáze RÚIAN

K databázi RÚIAN jsou publikovány aktualizační balíčky jednou denně, je proto vhodné je stahovat automaticky každý den, případně zároveň importovat do databáze. K tomu postačí nastavit plánovač operačního systému tak, aby jednou denně spouštěl modul RÚIAN Downloader.

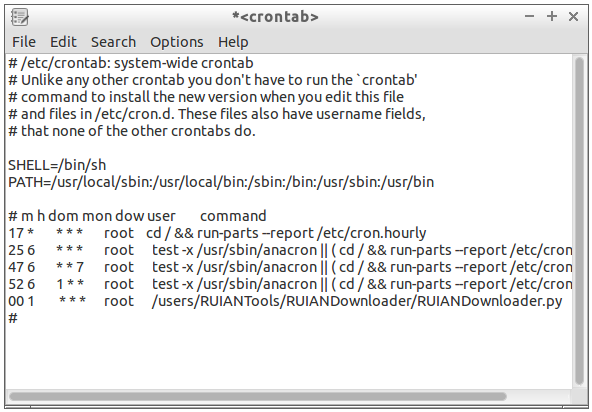
### Operační systém Microsoft Windows

1. Vyvoláme položku **Nastavení** | **Ovládací panely** | **Systém a zabezpečení** | **Nástroje pro správu** | **Plánovač úloh**[[1]](#footnote-1)
2. Z nabídky zvolíme položku **Akce** | **Vytvořit základní úlohu**
3. Vyplníme položky v zobrazeném průvodci tak, aby odpovídal našim požadavkům
4. Stiskneme **Dokončit**.



### Operační systém Unix, (Linux)

Přidáme příslušné nastavení do tabulky *etc/crontab* programu *cron* dle našich požadavků, například:

*00 1\* \* \* root run-parts /users/raugustyn/RUIANTools/RUIANDownloader/RUIANDownloader.py*

# Informace o výsledcích a průběhu stahování

## RUIANDownlaoder.log

Tento soubor je generován v adresáři RUIANDownloader a obsahuje podrobný výpis stahování, včetně případných chybových hlášení. Tento soubor se po opětovném spuštění stahování a stahování aktualizací nemaže, ale je průběžně doplňován.



# Použité zkratky

|  |  |
| --- | --- |
| [RÚIAN](http://www.cuzk.cz/ruian/RUIAN.aspx) | Registr územní identifikace, adres a nemovitostí |
| [VDP](http://vdp.cuzk.cz/) | Veřejný dálkový přístup |
| [OpenSource](http://opensource.org/) | OpenSource Software Initiative |
|  |  |
|  |  |

1. Platí pro Microsoft Windows 8.0 [↑](#footnote-ref-1)