DoelenCalendar Documentation

Release 0.0.1.dev20131014160300

Gertjan van den Burg

October 14, 2013

CONTENTS

CHAPTER

ONE

CORE

1.1 Calendar

dedoelen.core.calendar.**make_calendar**(*voorstellingen*)

Create a `icalendar.Calendar` instance from a list of `dedoelen.core.models.Voorstelling` instances. The timezone information is set for the Amsterdam timezone (DST is incorporated).

Parameters **voorstellingen** (*list*) – events that need to be in the calendar

Returns calendar instance

Return type `icalendar.Calendar`

1.2 Conf

This file contains the configuration options of the program and the code necessary to create a `Settings` instance.

class dedoelen.core.conf.**Settings**(***entries*)

A `Settings` instance contains the configuration for the program.

Parameters

- **AMSTERDAM** (`pytz.timezone`) – timezone information for Europe/Amsterdam
- **FEED_URL** (*str*) – address of the RSS feed
- **OUTFILE** (*str*) – path of the output file

1.3 Convert

This file contains convertion functions from internal `dedoelen.core.models.Voorstelling` instances to `icalendar.Event` instances and vice versa.

dedoelen.core.convert.**event2voorstelling**(*event*)

Convert a `icalendar.Event` instance to an internal `dedoelen.core.models.Voorstelling` instance.

Parameters **event** (`icalender.Event`) – event in the calendar

Returns event in De Doelen

Return type `dedoelen.core.models.Voorstelling`

dedoelen.core.convert.**voorstelling2event**(*voorstelling*)

Convert a dedoelen.core.models.Voorstelling instance to an icalendar.Event instance. The uid of the event is the hash of the event URL. Hence, we assume that if the event changes, the URL stays the same. This allows us to update an event in the calendar by changing everything, but keeping the same UID ensures the calendar application can still identify the event as merely changed (not added new). When an event is changed the sequence must be increased.

Parameters **voorstelling** (dedoelen.core.models.Voorstelling) – event in De Doelen

Returns ICS event for the event

Return type icalender.Event

1.4 Models

class dedoelen.core.models.**Voorstelling**(*title=''*, *link=''*, *tstart=None*, *tend=None*, *room=''*, *desc=''*, *seq=0*)

Een voorstelling object bevat alle informatie voor een voorstelling in De Doelen.

Parameters

- **title** (*str*) – titel van de voorstelling
- **link** (*str*) – www link voor het evenement
- **tstart** (datetime.time) – start tijd van het evenement
- **tend** (datetime.time) – eind tijd van het evenement
- **room** (*str*) – zaal waarin het evenement plaatsvind
- **description** (*str*) – beschrijving van het evenement
- **sequence** (*int*) – sequence van het evenement (wordt gebruikt wanneer het evenement wordt geupdate)

1.5 Parser

dedoelen.core.parser.**html2voorstelling**(*page*)

Convert a scraped html page to a dedoelen.core.models.Voorstelling instance. Pages that can not be parsed properly are ignored, and None is returned. Text entries are converted to ascii, and special characters are ignored. Events that have no start time are also ignored (None is returned).

Parameters **page** (*tuple*) – tuple of the page url and the page html

Returns representation of the event

Return type dedoelen.core.models.Voorstelling

1.6 Scraper

In this file all functions related to scraping are defined.

dedoelen.core.scraper.**get_doelen_page**(*browser*, *url*)

Get a single page from the web. Several warnings can occur in this process. An httplib.IncompleteRead exception occurs when the request to the server is not completed, in this case a 5 second wait is allowed and the error is logged. After the wait the attempt is repeated. An urllib2.URLError can also occur, it is handled

in the same manner. Finally, a SPAM-bot page can be shown, which can happen when too many requests are made to the server. Luckily, this page includes the cookies required for a valid request, so when this page is encountered the next attempts do not suffer. To be fair, a 10 second wait is included after this error. When a page is retrieved succesfully, the loop breaks.

Parameters

- **browser** (`mechanize.Browser`) – initialized browser instance
- **url** (*str*) – url of the web page we want

Returns webpage

Return type str

`dedoelen.core.scraper.`**`scrape_html`**(*urls*)

Get all pages given by the `urls`. If the server returns an error when too many requests are placed, a 10 second wait is performed. If an error occurs when getting the page a 5 second wait is performed and a warning is logged. All pages are returned as (url, page) tuples. A progress bar is printed to show the total number of pages retrieved.

Parameters **urls** (*list*) – urls of De Doelen events

Returns tuples of html pages and the corresponding url

Return type list

`dedoelen.core.scraper.`**`scrape_rss`**()

Scrape the RSS feed for urls. All urls are collected and returned.

Raises ValueError when no urls are found in the feed.

Returns urls scraped from RSS feed

Return type list

1.7 Update

`dedoelen.core.update.`**`update_voorstellingen`**(*voorstellingen*)

Update events in the calendar file. This update is based on a newly collected set of events, and previous events read from the calendar. Old events are those that are in the past. These events are in the event list from the previous calendar file, but not necessarily in the list of new events. Updated events are in the list of new events, and their link is in the list of previous events (link is assumed constant). New events are those which link is not in the list of previous links. Deleted events are those that are in the list of events from the previous calendar file, but are not in the new list and are not updated or old.

Parameters **voorstellingen** (*list*) – `dedoelen.core.models.Voorstelling` instances that are newly retrieved.

Returns updated list of `dedoelen.core.models.Voorstelling` instances

Return type list

CHAPTER

TWO

USER INTERFACE

2.1 Command Line

CHAPTER

THREE

UTILS

3.1 Localize

dedoelen.utils.localize.**set_locale**()
Stel de taal van het programma in op Nederlands.

3.2 Log

dedoelen.utils.log.**init_logger**(*flushlevel*)

3.3 Progress

class dedoelen.utils.progress.**AdaptiveETA**(*format='Elapsed Time: %s'*)

update(*pbar*)

3.4 Version

dedoelen.utils.version.**get_git_changeset**()

PYTHON MODULE INDEX

d

INDEX