

Разработка сервиса получения архивов климатических данных

Евгений Юмашев ¹

¹ Институт динамики систем и теории управления им. В.М. Матросова СО РАН, ул. Лермонтова, д. 134, 664033, Иркутск, Россия

Аннотация

В статье представлена реализация WPS-сервиса получения архивов климатических данных с погодного сайта rp5 по названиям или географическим координатам населенных пунктов за определенный период. При реализации WPS-сервиса использованы специализированные библиотеки языка программирования Python, такие как requests, bs4, json, haversine, ruwps-flask и др. Разработанный сервис обеспечивает автоматизацию получения и обработки больших массивов климатических данных, снижая трудозатраты конечных пользователей.

Ключевые слова

Архив климатических данных, WPS-сервис, Python, rp5

1. Введение

Климатические данные играют ключевую роль в различных научных исследованиях, прикладных проектах и принятии решений в таких областях, как сельское хозяйство, городское планирование, энергетика и многих других. Правильное понимание климатических условий и их изменений позволяет более эффективно управлять ресурсами и минимизировать риски, связанные с экстремальными погодными явлениями.

Однако доступ к климатическим данным зачастую затруднен из-за значительных временных затрат на их сбор и обработку. Для решения этой проблемы появилась необходимость в разработке автоматизированного сервиса, способного оперативно предоставлять необходимые данные из открытых источников. В связи с этим в данной статье рассматривается создание WPS-сервиса, который предназначен для получения архивов климатических данных с погодного ресурса на основании заданных пользователем названий или географическим координатам населенных пунктов.

Цель данной работы заключается в реализации этого WPS-сервиса и демонстрации его возможностей по автоматическому извлечению архивов климатических данных в заданном диапазоне дат. Этот подход позволит повысить доступность и удобство получения климатической информации для широкого круга пользователей.

2. Функциональные и системные требования к разрабатываемому WPS-сервису

К разрабатываемому WPS-сервису предъявляются следующие функциональные и системные требования:

- открытый WPS-сервис;
- получение архива климатических данных по названию населенного пункта;



- возможность получения климатических данных по географическим координатам местности;
- получение данных за выбранный период;
- выдача результата в формате Excel;
- преобразование Excel-файла к необходимому виду.

Реализация вышеперечисленных требований позволит разработать сервис для получения архивов климатических данных для любых населенных пунктов.

3. Разработка WPS-сервиса

Для получения архива климатических данных пользователю необходимо выполнить следующие шаги: зайти на погодный сайт, выбрать населенный пункт, перейти по ссылке на форму скачивания архива, задать начальную и конечную даты расчетного периода, затем скачать архив и разархивировать его. Трудность заключается в том, что это приходится делать вручную, так как нет средства для автоматического получения данных. Кроме того, не имеется возможности по географической широте и долготе получить архив данных ближайшей метеостанции.

Для решения вышеописанных проблем реализованы сервис-ориентированные методы в виде WPS-сервиса. Схема работы WPS-сервиса представлена на рисунке 1.

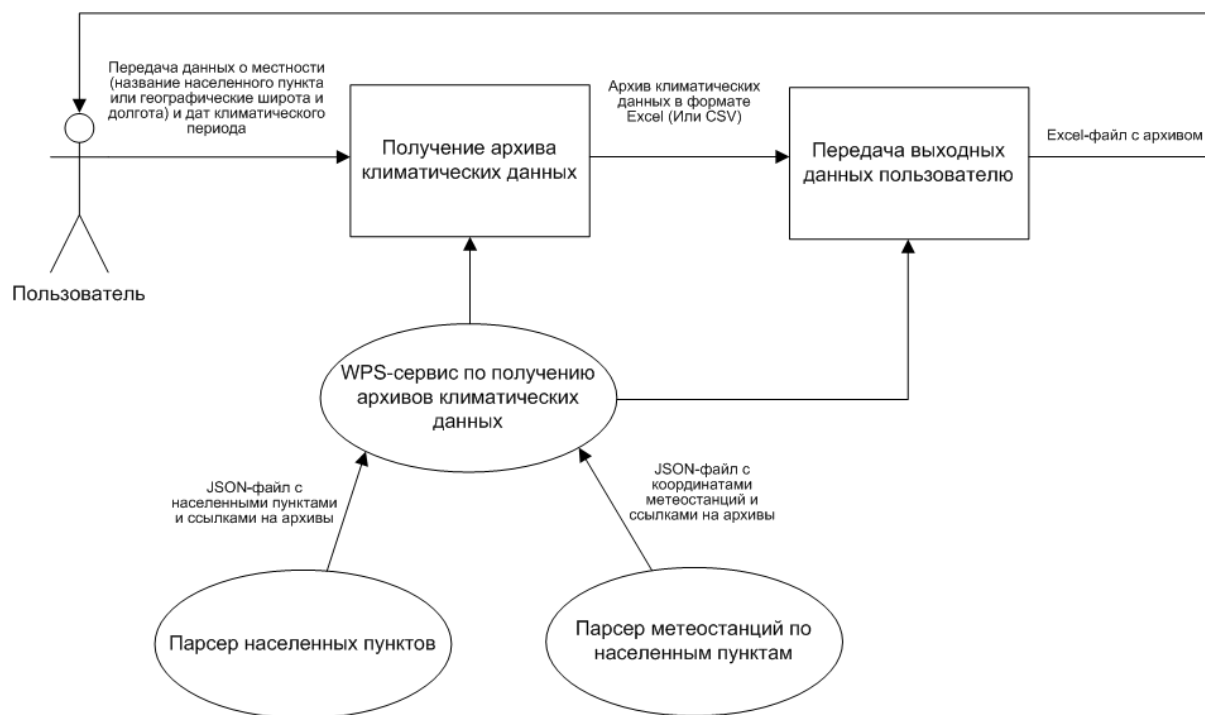


Рисунок 1: Схема работы WPS-сервиса

Пользователь подает на вход WPS-сервиса либо название населенного пункта, либо географическую координаты (широту и долготу), а также период, по которому он хочет получить данные. Для поддержки этих запросов сформированы два JSON-файла: один файл содержит названия более 15 тысяч населенных пунктов и соответствующие ссылки на архивы, другой файл – координаты ближайших метеостанций и их архивы.

Для получения архивов климатических данных используется сайт gr5.ru [1], который предоставляет развернутую информацию о таких показателях, как температура, давление, влажность, состав облаков, фактическая облачность и другие параметры. Создание формируемых JSON-файлов осуществлялось с помощью парсеров веб-страниц. Парсер населенных пунктов собирает названия населенных пунктов и соответствующие ссылки на архивы, а парсер метеостанций получает координаты и ссылки на архивы. Также файлы

обработаны с целью устранения некорректных данных. Оба парсера реализованы при помощи библиотек bs4 [2], requests и json.

Для работы сервиса реализованы методы, которые осуществляют извлечение архива с сайта gr5.ru [3], распаковку и сохранение Excel-файла, а также его преобразование (удаление ненужной информации). Эти методы разработаны с использованием библиотек requests, bs4, haversine [4], openpyxl [5], gzip и datetime. Для создания и тестирования WPS-сервиса использованы библиотеки ruwps, os и ruwps-flask.

В результате разработки сервиса успешно решены задачи автоматического получения архивов климатических данных, как по названию населенного пункта, так и по географическим координатам. На рисунке 2 представлены результаты работы сервиса с получением архива для города Иркутска за период с 2018 по 2023 годы в формате Excel.

Архив погоды в Иркутске.xlsx																		
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	
1	0	15.07.2018 02:00	15,8	712	Ветер, дуть	1	Облаков	0	0	0	0	0	0,125	0,125	3	0	78	
2	3	15.07.2018 05:00	13,4	711,8	Ветер, дуть	1	40%.	Слоисто-к	Облаков	2500 или	Высококу	Перистые	0,125	0,125	3	3	85	
3	6	15.07.2018 08:00	17,1	712,3	Ветер, дуть	1	70 – 80%.	Слоисто-к	Облаков	2500 или	Высококу	Перистые	0,125	0,125	3	6	71	
4	9	15.07.2018 11:00	22,9	712	Ветер, дуть	1	90 или бс	Слоисто-к	Облаков	2500 или	Высококу	Перистые	0,125	0,125	3	9	51	
5	12	15.07.2018 14:00	26,2	711,2	Ветер, дуть	1	90 или бс	Слоисто-к	Облаков	2500 или	Высококу	Перистые	0,125	0,125	3	12	47	
6	15	15.07.2018 17:00	26	710,8	Ветер, дуть	1	100%.	Слоисто-к	Облаков	2500 или	Высококу	Перисто-с	0,125	0,125	3	15	52	
7	18	15.07.2018 20:00	20,8	711,4	Ветер, дуть	1	100%.	Слоисто-к	40%.	600-1000	Высококу	0	0,125	0,125	3	18	65	
8	21	15.07.2018 23:00	18,4	712,1	Ветер, дуть	2	90 или бс	Слоисто-к	70 – 80%.	600-1000	Высококу	Перистые	0,125	0,125	3	21	70	
9	24	16.07.2018 02:00	16,7	712	Ветер, дуть	1	70 – 80%.	Слоисто-к	40%.	2500 или	Высококу	Перистые	0,125	0,125	3	24	73	
10	27	16.07.2018 05:00	15,4	711,4	Ветер, дуть	2	90 или бс	Кучево-дс	40%.	600-1000	Высококу	Перистые	0,125	0,125	3	27	85	
11	30	16.07.2018 08:00	19,8	711,8	Ветер, дуть	2	70 – 80%.	Слоисто-к	50%.	2500 или	Высококу	Перистые	0,125	0,125	3	30	66	
12	33	16.07.2018 11:00	23,4	711,6	Ветер, дуть	2	90 или бс	Слоисто-к	50%.	2500 или	Высококу	Перисто-к	0,125	0,125	3	33	63	
13	36	16.07.2018 14:00	24,6	711,4	Ветер, дуть	2	100%.	Кучево-дс	70 – 80%.	600-1000	Высококу	0	0,125	0,125	3	36	64	
14	39	16.07.2018 17:00	23	711,6	Ветер, дуть	3	100%.	Кучево-дс	100%.	1000-1500	0	0	0,125	0,125	3	39	67	
15	42	16.07.2018 20:00	21,9	712	Ветер, дуть	2	90 или бс	Кучево-дс	70 – 80%.	600-1000	Высококу	Перистых	0,125	0,125	3	42	68	
16	45	16.07.2018 23:00	19,9	712,3	Ветер, дуть	2	70 – 80%.	Кучево-дс	60%.	600-1000	Высококу	Перистых	0,125	0,125	3	45	75	
17	48	17.07.2018 02:00	18,8	712,6	Ветер, дуть	2	100%.	Кучево-дс	100%.	600-1000	0	0	0,125	0,125	3	48	74	
18	51	17.07.2018 05:00	16,5	713,1	Ветер, дуть	4	100%.	Кучево-дс	50%.	600-1000	Высококу	Перистые	0,125	0,125	3	51	72	
19	54	17.07.2018 08:00	16,5	713,7	Ветер, дуть	3	100%.	Кучево-дс	100%.	600-1000	0	0	0,125	0,125	3	54	72	
20	57	17.07.2018 11:00	17,4	713,8	Ветер, дуть	3	100%.	Кучево-дс	100%.	1000-1500	0	0	0,125	0,125	3	57	69	
21	60	17.07.2018 14:00	18,8	713,8	Ветер, дуть	2	100%.	Кучево-дс	100%.	1000-1500	0	0	0,125	0,125	3	60	70	
22	63	17.07.2018 17:00	20,3	713,2	Ветер, дуть	2	100%.	Кучево-дс	100%.	1000-1500	0	0	0,125	0,125	3	63	69	

Рисунок 2: Результат работы сервиса для города Иркутска

Результаты, полученные при использовании WPS-сервиса, могут в дальнейшем быть использованы в исследованиях по солнечной радиации, электроэнергетике и других научных работах.

4. Заключение

В рамках данной статьи представлена разработка WPS-сервиса, предназначенного для автоматического получения архивов климатических данных с сайта gr5.ru. Процесс получения данных осуществляется по введенным пользователем названиям населенных пунктов или географическим координатам, что значительно упрощает задачу доступа к необходимой информации. Описание архитектуры сервиса, реализация парсеров для формирования JSON-файлов, а также использование библиотек, таких как requests, bs4, openpyxl и ruwps, позволили эффективно автоматизировать сбор и обработку климатических данных. В дальнейшем планируется решение следующих задач:

- оптимизация алгоритмов извлечения и обработки данных для сокращения времени выполнения операций;
- улучшение работы парсеров данных и увеличение числа населенных пунктов и метеостанций в JSON-файлах;
- внедрение разработанного сервиса в геопортал ИДСТУ СО РАН.

5. Благодарности

Работа выполнена в рамках гранта № 075-15-2024-533 Министерства науки и высшего образования РФ на выполнение крупного научного проекта по приоритетным направлениям научно-технологического развития (проект «Фундаментальные исследования Байкальской природной территории на основе системы взаимосвязанных базовых методов, моделей, нейронных сетей и цифровой платформы экологического мониторинга окружающей среды», рег. № 124052100088-3).

6. Список источников

- [1] RP5 – погода в мире: сайт – URL: https://rp5.ru/Погода_в_мире.
- [2] BeautifulSoup 4.12.0 documentation: сайт – URL: <https://www.crummy.com/software/BeautifulSoup/bs4/doc/>.
- [3] GitHub project «rp5_weather»: сайт – URL: https://github.com/KonstantenKomkov/rp5_weather.
- [4] Haversine: сайт – URL: <https://pypi.org/project/haversine>.
- [5] Openpyxl – A Python library to read/write Excel 2010 xlsx/xlsm files: сайт – URL: <https://openpyxl.readthedocs.io/en/stable>.