

Forecast Search Wizard: A Tool to Search NOAA Text-Forecasts



https://github.com/allenea/Forecast_Search_Wizard

Eric Allen, allenea@udel.edu

Department of Geography; College of Earth, Ocean, & Environment;
University of Delaware; Newark, DE, USA

Overview

- Forecasters at the National Weather Service do an excellent job creating accurate text-based forecasts every day.
- The Forecast Search Wizard is a tool developed to analyze archived NOAA text products to obtain valuable information.
- Searches for keywords in NOAA/NWS text products.
- Identifies Case Studies, Weather Events, Other Specific Information

Benefits

- Convenient, time-saving, and accurate
- One-of-a-kind tool developed to search NOAA/NWS text-products
- Robust & efficient search algorithm
- New source of big data: A forgotten source of information
- Improve depth & quality of research
- Could be used as a tool to parse and archive NOAA text products to a database.

Limitations

- Human-Input Text-Products
 - Abbreviations, misspellings and typos
 - Retrieving date, time, and time zone information
- Speed vs Effectiveness and Accuracy

***** Examples *****

- In the analysis of all NHC forecasts from 1996-Present there are only 24 cases where BOTH "lightning" AND "intensification" were used in the same forecast (20 sec). Looking for "lightning' AND "rapid intensification" resulted in 7 cases (10 sec).
- The keyword "Chupacabra" has been mentioned twice in AFDs. The search took less than 12 minutes.

09-02-2012 16:49 SJT Chupacabra
09-16-2018 01:28 EWX Chupacabra

NWS Text-Product Format^[1]

FXaaii cccc ddhhmm
AFDxxx

AREA FORECAST DISCUSSION
NATIONAL WEATHER SERVICE city state
time am/pm time_zone day mon dd yyyy

WMO heading
AWIPS ID

NWS Product Name
Issuing Office or Agency
Issuing Time/Date

Figure 1: Sample format of the Area Forecast Discussion (AFD). Unfortunately, these protocols are not always followed. The Forecast Search Wizard is dynamic. It is capable of resolving most dates and times that do not follow this format or that contain incomplete information. Approximately 13,400 out of 10 Million+ possible forecasts are unattainable.

Forecast Search Wizard Workflow

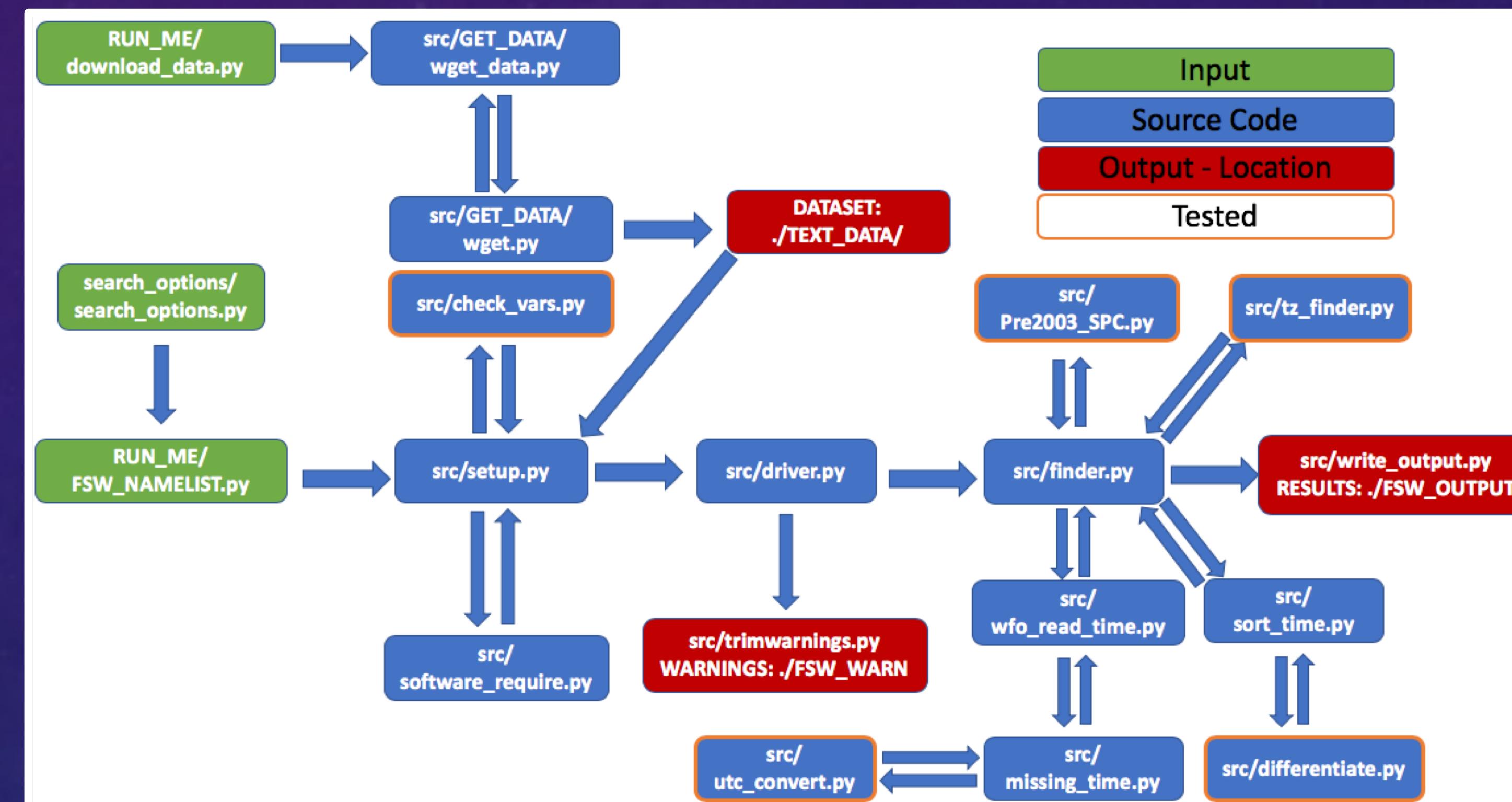


Figure 2: Shows the workflow diagram for the Forecast Search Wizard. First, download the dataset you need by running download_data.py, then configure your search in FSW_NAMELIST.py and run that file. The main component is the finder.py program. It identifies each forecast, the date-time information, and the keywords. The other key component is the wfo_read_time.py which puts that time information into a usable format.

Performance and Results^[3]

Issuing Office(s)	Single Word Search Analysis		Normal Case ("Lightning")		Worst Case ("The")	
	Total Products	Products Searched	Cases	Time	Cases	Time
Single WFO (AFDBGM)	63,619	63,590	1,282	26 sec	59,179	38 sec
All NHC Forecasts	106,602	106,412	952	20 sec	95,026	41 sec
All NCEP Forecasts	1,402,701	1,398,414	28,097	4.8 min	1,270,422	11.1 min
All Area Forecast Discussions	5,079,801	5,072,265	177,131	38.7 min	4,689,592	44.6 min
All Forecasts	10,040,958	10,027,342	609,301	55.1 min	9,124,460	74.2 min

Figure 3: Forecast Search Wizard results from a simple one word GREP-style search from 1996-2019. It provided a .txt file with a list of all forecasts in which the keyword was found.

Data

- Run download_data.py to get the dataset
- Pre-configured and easily adaptable
- All data **MUST** be stored in the TEXT_DATA directory
- 1996 – Present [†]: Iowa Environmental Mesonet (IEM)
- 560+ Products (*Over 10 Million forecasts*)
- 350+ Pre-configured Search Options
- Storage Requirements: 30-35GB
- Supported Types: Area Forecast Discussion, Nowcasts, Local Storm Reports, and 139 forecast products issued by the National Centers for Environmental Prediction

[†]2008 – Present: "Very good data coverage and higher fidelity archiving." For more information visit IEM for Archive Completeness[2]

Running the Forecast Search Wizard

1. User Provides Search Criteria in FSW_NAMELIST.py

- List of Strings
 - Keyword(s)
 - Forecast Product(s) [AWIPS ID]
- Integer
 - Start Year (1996 or later)
 - End Year (Not to exceed the current year)
- Boolean: True / False
 - Enable assumptions
 - Search for ALL or ANY keywords in a forecast
 - Search by FORECAST or by DAY
 - GREP-style search or search for WHOLE WORD

2. Run FSW_NAMELIST.py

3. Get text output file with search information and cases found

Software Requirements: Python3, pytz, re, datetime, os, sys, glob, time, and other standard built-ins

Future

- Support user-provided datasets
- Search-By-Section of the forecast
- Move away from the IEM dataset
- Web-Based Application

References

[1] NOAA/NWS, 2005: NWS Instruction 10-503

[2] Iowa Environmental Mesonet (Iowa State University)

[3] Late 2013 iMac with an 3.2 GHz Intel Core i5 processor and 16GB RAM