

Webengine

The World Congress on Engineering and Computer Science (WCECS'14)

San Francisco, USA, 22-24 October, 2014

Matthew Harris

 Lawrence Livermore
National Laboratory

LLNL-CONF-656083

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC



Software Used

- Hyde

- Hyde is a static website generator written in python. While Hyde took life as awesome Jekyll's evil twin, it has since been completely consumed by the dark side and has an identity of its own.
- Hyde desires to fulfill the lofty goal of removing the pain points involved in creating and maintaining static websites.

- Jekyll

- Jekyll is a simple, blog-aware, static site generator. It takes a template directory containing raw text files in various formats, runs it through Markdown (or Textile) and Liquid converters, and spits out a complete, ready-to-publish static website suitable for serving with your favorite web server. Jekyll also happens to be the engine behind GitHub Pages, which means you can use Jekyll to host your project's page, blog, or website from GitHub's servers for free.

- Webshooter

- Webshooter is a command-line assistant to the static website generator Hyde or Jekyll. It produces Hyde or Jekyll ready websites in the user's choice of layout and automates page creation.

- Bootstrap

- Bootstrap is the most popular HTML, CSS, and JS framework for developing responsive, mobile first projects on the web.

Content Management Systems

- Drupal
 - Drupal is an open source content management platform powering millions of websites and applications. It's built, used, and supported by an active and diverse community of people around the world.
- Zope
 - Zope is a free and open-source web application server written in the object-oriented programming language Python.
- Plone
 - Plone is among the top 2% of all open source projects worldwide, with 340 core developers and more than 300 solution providers in 57 countries. The project has been actively developed since 2001, is available in more than 40 languages, and has the best security track record of any major CMS.
 - It is owned by the Plone Foundation, a 501(c)(3) not-for-profit organization, and is available for all major operating systems.
- Moin Moin
 - MoinMoin is an advanced, easy to use and extensible WikiEngine with a large community of users. Said in a few words, it is about collaboration on easily editable web pages.

Generated HTML vs Markdown

- Output from Drupal

```
1 <div class="link-wrapper">
2   <ul class="links inline">
3     <li class="node-readmore first">
4       <a href="/?q=node/11" rel="tag" title="ArticleTwo"> Read more
5         <span class="element-invisible"> about ArticleTwo </span>
6       </a>
7     </li>
8     <li class="comment_forbidden last">
9       <span>
10        <a href="/?q=user/login&destination=node/11%23comment-form">Log in</a> to post comments
11      </span>
12    </li>
13  </ul>
14 </div>
```

- Output from Webengine (markdown)

```
1 * [Read more](/?articleid=11) about ArticleTwo
2 * [Log in](/?login=user) to post comments
```

webengine.py

- url_gatherer: Takes a site-index.html page and generates a list of URLs
- file_gatherer: Fetches (WGET) all the pages in the URL list
- site_gatherer: Takes a site URL and fetches (WGET) the whole site
- file_extractor: Recursively reads all site files, and saves a little of the content as possible
- file_corrector: Reads all new files changing href and src to user input
- html_table_2_markdown: Parses all files changing html tables to GitHub flavored markdown
- file_converter: Converts the remaining html to markdown (plain text – Arron Swartz html2text.py)
- image_gatherer: Creates a unique list of all site images and downloads them
- bold_cleanup: Removes all ** and __ in all files
- head_adder: Adds a Hyde or Jekyll header and footer to each page

webshooter.py

- Now with Jekyll Support!
- Creates a Hyde or Jekyll static generating website with one of three site styles.
- Customizable, easy to add new styles.

LLNL One Lab Bootstrap

Analytics and Informatics Management Systems 

[Mission](#) [Contributors](#) [Documents](#) [Projects](#) [Climate Game](#) [Calendar](#) [Wiki](#)

The AIMS Group

AIMS is a leader and visionary architect for all aspects of data discovery and knowledge integration.

[Learn more »](#)



Projects

ESGF



UV-CDAT



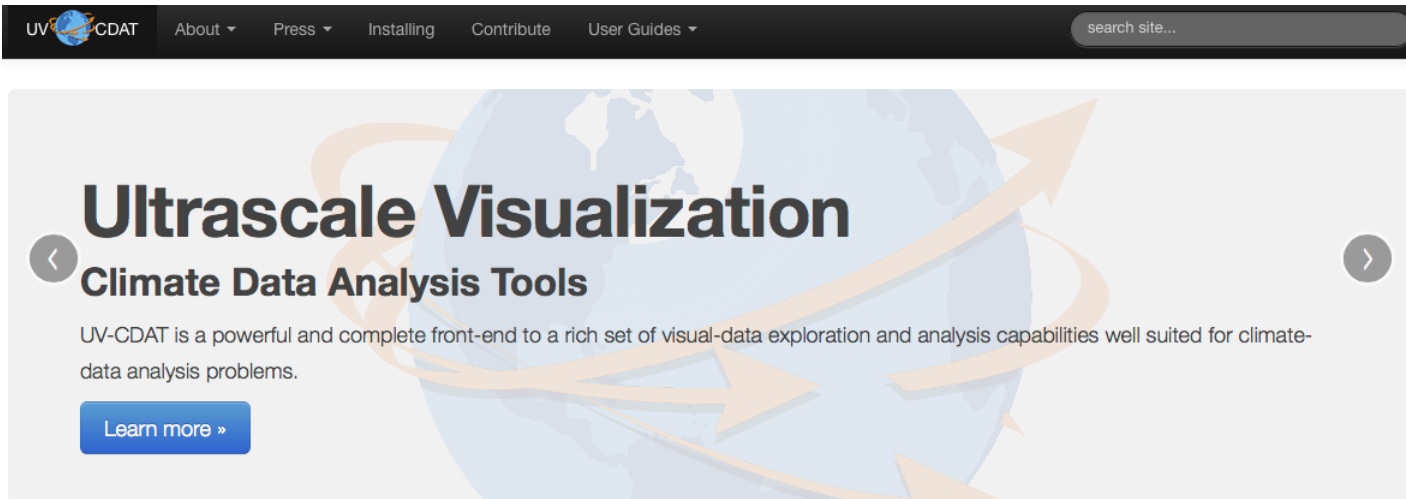
ACME



Webshooter



Basic Bootstrap



The screenshot shows the top navigation bar of the UV-CDAT website with links for 'About', 'Press', 'Installing', 'Contribute', and 'User Guides'. A search bar is located on the right. The main banner features a globe with orange arrows and the text 'Ultrascale Visualization Climate Data Analysis Tools'. Below the title, it states: 'UV-CDAT is a powerful and complete front-end to a rich set of visual-data exploration and analysis capabilities well suited for climate-data analysis problems.' A blue button labeled 'Learn more »' is positioned at the bottom left of the banner.

The UV-CDAT Project

UV-CDAT builds on the following key technologies:

1. The Climate Data Analysis Tools (CDAT) framework developed at LLNL for the analysis, visualization, and management of large-scale distributed climate data;
2. ParaView: an open-source, multi-platform, parallel-capable visualization tool with recently added capabilities to better support specific needs of the climate-science community;
3. VisTrails, an open-source scientific workflow and provenance management system that supports data exploration and visualization;
4. VisIt: an open-source, parallel-capable, visual-data exploration and analysis tool that is capable of running on a diverse set of platforms, ranging from laptops to the Department of Energy's largest supercomputers.

These combined tools, along with others such as the R open-source statistical analysis and plotting software and custom packages (e.g. vtDV3D), form UV-CDAT and provide a synergistic approach to climate modeling, allowing researchers to advance scientific visualization of large-scale climate data sets. The UV-CDAT framework couples powerful software infrastructures through two primary means:

1. Tightly coupled integration of the CDAT Core with the VTK/ParaView infrastructure to provide high-performance, parallel-streaming data analysis and visualization of massive climate-data sets (other tightly coupled tools include VCS, VisTrails, DV3D, and ESMF/ESMP);
2. Loosely coupled integration to provide the flexibility of using tools quickly in the infrastructure such as VISUS, VisIt, R, and MatLab for data analysis and visualization as well as to apply customized data analysis applications within an integrated environment.

Within both paradigms, UV-CDAT will provide data-provenance capture and mechanisms to support data analysis via the VisTrails infrastructure.

LinkedIn tshirt



ESGF

Earth System Grid Federation

installer · node-manager · publisher · idp · orp · security · search · drslib · stager · fetcher · toolbox · ...

esgf
design
download
quickstart
developer info
performance
repo
[code docs](#)
dep-report
unit tests
coverage
ci
faq
mailing list
bugs
wiki
blog

Web Front-End - How we present ourselves on the web...

SEE and manipulate your data! :-)

Browser Support

Experience

- Time savings
- Creating updated websites
- Dynamically created static webpages
- Hyde vs Jekyll
- Power of Markdown

Future Work

- Automated Testing
- Corner Cases / One off error handling
- Interactive webengine and command line options
- webshooter command line options
- Automation of style updating and changing

Conclusion

- The Webshooter repository started with one file in June 2013.
- I hope others will find it useful and even contribute to it.

Acknowledgments

I would like to thank Sam Fries for his time debugging, testing, and contributing to the code base. I would also like to recognize Ben Carlsson, Dean Williams for their support and contributions. Special thank you to Katie Walter, and Brian Harris for their editorial skills.

Questions?

<http://github.com/webshootertk>

<http://github.com/mattben>

harris112@llnl.gov