

WebAwk



Address	{value_address}
Contact Person	{value_contact_person}
Mobile Number	{value_mobile_number}
Email	{value_email}

This is a proof-of-concept of a tool to automate web browsing / data collection. It works like AWK except that instead of working on files and lines it works on HTML pages and hyperlinks. It is meant to be run as a command line script and includes `base_url` - the URL the script was initially invoked on, `base_path` - root of saved data tree, `url` - current URL being processed, `linked_from` - parent of current URL, and `content` - the actual data corresponding to the current URL.

For more details, please visit <http://scripts.goclix.com/webawk-12905>