Introducing the Open Online Newspaper Initiative

Jessica Dussault
jdussault@unl.edu
University of Nebraska, United States of America

Laura Weakly
lweakly2@unl.edu
University of Nebraska, United States of America

Karin Dalziel
kdalziel2@unl.edu
University of Nebraska, United States of America

Jeremy Echols
jechols@uoregon.edu
University of Oregon, United States of America

Karen Estlund
kme20@psu.edu
Pennsylvania State University
United States of America

Andrew Gearhart
andrew@psu.edu
Pennsylvania State University
United States of America

Sheila Rabun
srabun@clir.org
IIIF Consortium, United States of America

Greg Tunink
techgique@unl.edu
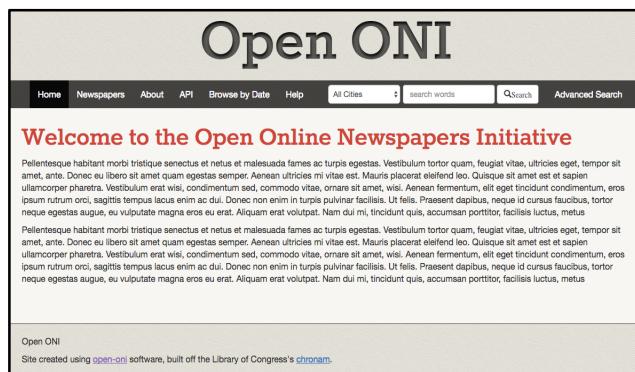University of Nebraska, United States of America

The Open Online Newspaper Initiative (Open ONI) is an open source collaboration whose goal is to lower the entrance bar for libraries, archives, historical societies, and other cultural heritage institutions to display digital newspaper content. Open ONI was formed in response to a need for free, easily deployed, flexible, plug-and-play software that is useful for collections large and small, local and national.

Open ONI's code base was forked from the Library of Congress newspaper application, chronam (Library of Congress, 2016a). chronam was created to support the National Digital Newspaper Program (NDNP), a national project supported by the Library of Congress and the National Endowment for the Humanities, which seeks to digitize and add titles to a searchable, online collection at Chronicling America (Library of Congress, 2016d; 2016b). State entities often have newspaper content which they are unable to put online through Chronicling America and must seek a different solution to present their digitized periodicals (Library of Congress, 2016c; Center for Research Libraries Global Resources Network, 2015). The Library of Congress released the source code for chronam to help address this problem. The software currently available on GitHub requires skilled technical staff to customize the application and update related code packages, which can be a barrier for small institutions. Open ONI was born in 2015 when a group of librarians, project managers, and developers working on their own chronam installations gathered to discuss a friendly fork of the software. In the process of setting up, deploying, and customizing chronam for their own newspaper sites, Open ONI members had identified many shared interests, from fixing common bugs to building new features, and decided that pooling their resources and efforts to develop an application to fit the majority of their needs would benefit them all. Though we are currently working towards meeting our own implementation goals, we are keeping in mind how this might be applicable to others in the future.

Since beginning the initiative, Open ONI developers have made substantial changes and improvements to the initial chronam code. The web framework, Django, has been updated to the latest long term support version, and many supporting libraries and command line tools have been updated or replaced, when existing libraries and tools were deprecated or no longer available. Perhaps the largest improvement is the incorporation of the RAIS image server, an International Image Interoperability Framework (IIIF) compliant image server developed by the University of Oregon Library to deliver JP2s nearly as fast as chronam delivered TIFFs, but using a fraction of the RAM (Echols and Krech, 2016). RAIS is a 100% open source alternative to other JP2 image servers. With contributions from the Maryland Institute for Technology in the Humanities (MITH) at the University of Maryland, and sponsored by the IIIF-Consortium, Open ONI offers IIIF compatible URLs to allow the substitution of other image servers. Beyond existing interoperability formats such as MARC and

linked data, future plans could include IIIF manifests and other metadata formats/downloads, depending on developer time. Open ONI's developer environment relies on Docker containers to separate the different components of the software and provide a one-line setup command. The application is configurable through settings files and personalizable, with a default theme that users may copy to get started on their own look and feel. Additionally, Open ONI offers several plugins that users can incorporate into their site such as featured content, randomly selected newspaper pages, and a customizable map. A recent feature allows for defining copyright statements by publication and date range, and showing these statements directly below each page.

Three states have plans to redeploy their newspaper sites with Open ONI in 2017: The University of Nebraska, University of Oregon, and Penn State University. Features that will be developed before all three launch their sites include extending the advanced search features for pages and newspaper titles, reinforcing the current test suite, updating the search engine from Solr 4 to Solr 6, and developing documentation to help organizations migrate from chronam to Open ONI.

Our poster will cover the above improvements of Open ONI, as well as the road map for future work. We will also have an instance of Open ONI available at the conference which will demonstrate the speed of the RAIS image server for JPG 2000s, default theme, new advanced search, and plugins. We hope that Open ONI offers organizations with newspaper collections of all types a reliable and customizable option for presenting and searching their collections. This poster presentation is also outreach to gain support and input from the archives, library, historical society, developer communities and others.

Bibliography

Center for Research Libraries Global Resources Network (2015). The 'State of the Art': A Comparative Analysis of Newspaper Digitization to Date. https://www.crl.edu/sites/default/files/d6/attachments/events/ICON_Report-State_of_Digitization_final.pdf (accessed 24 October 2016).

Echols, J. and Krech, D. (2016). "RAIS Image Server." GitHub. https://github.com/uoregon-libraries/rais-image-server (accessed 24 October 2016).

Library of Congress (2016a). "chronam." GitHub. https://github.com/LibraryOfCongress/chronam (accessed 11 October 2016)

Library of Congress (2016b). "Chronicling America." http://chroniclingamerica.loc.gov/ (accessed 24 October 2016).

Library of Congress (2016c). "Content Selection." http://www.loc.gov/ndnp/guidelines/selection.html (accessed 24 October 2016)

Library of Congress (2016d). "National Digital Newspaper Program." http://www.loc.gov/ndnp/ (accessed 24 October 2016).

Open Online Newspaper Initiative (2016). "Open ONI." GitHub. https://github.com/open-oni/open-oni (accessed 11 October 2016).