Finding the Machine [processes]: Applying the Open Data Project to Research Services

Charleston Conference 2015 (Session with Kamran Naim and Robin Champieux)

Jason A. Clark Associate Professor Head of Library Informatics & Computing Montana State University (MSU) Library www.jasonclark.info

pinboard.in tag

pinboard.in/u:jasonclark/t:open-machines/

twitter as channel (#hashtag)

@jaclark #chs15

Open has Consequences

Almost always good ones...

Overview

- Why Open?
- "Open" and Research Signals
- Machine Processes + Tools
- Questions

What do you mean by "open"?

 An open work must satisfy requirements around open licenses, open access, and open formats. It can be summed up in the statement that:

Put most succinctly:

Read the full Open Definition »

"Open means anyone can freely access, use, modify, and share for any purpose (subject, at most, to requirements that preserve provenance and openness)."

"Open data and content can be freely used, modified, and

shared by anyone for any purpose"

http://opendefinition.org/od/

"A piece of data or content is open if anyone is free to use, reuse, and redistribute it - subject only, at most, to the requirement to attribute and/or share-alike."

Open Data + Research Signals? Part of a Library Continuum

Machine Harvesting

HTML

Application: Web Harvesting of PubMed

www.lib.montana.edu/~jason/files/crawler/ www.lib.montana.edu/~jason/files/crawler/index.txt

Given web page, return article titles and article URLs

Web Crawler Research Signals

https://www.ncbi.nlm.nih.gov/pubmed/26539112

Crawl

Author information

Abstract

Amyotrophic lateral sclerosis (ALS) is a lethal neurodegenerative disease causing death of motor neurons. This study investigated

the roles of energy metabolism in the pathogenesis of ALS in the SOD1(G93A) transgenic mouse model. Control and SOD1(G93A)

mice were administered with shcontrol or shPGC-1α in combination with PBS or thiazolidinedione (TZD) for 8 weeks. Gene

expression was analyzed by quantitative real-time PCR and Western blot. ROS and fibrosis were assessed with a colorimetric kit

and Sirius staining, respectively. Inflammatory cytokines were measured using ELISA kits. The levels of tissue ROS and serum

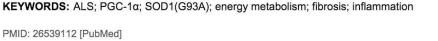
inflammatory cytokines were significantly higher in SOD1(G93A) mice compared to control mice, and knocking down peroxisome proliferator-activated receptor gamma coactivator 1-alpha (PGC-1α) drastically increased cytokine levels in both control and











exacerbate the pathological changes caused by SOD1 mutations to promote the pathogenesis of ALS.

SOD1(G93A) mice. Muscle fibrosis was much severer in SOD1(G93A) mice, and worsened by silencing PGC-1α and attenuated by

TZD. The expression levels of PGC-1α, SOD1, UCP2, and cytochrome C were substantially reduced by shPGC-1α and increased

by TZD in muscle of both control and SOD1(G93A) mice, whereas the level of NF-kB was significantly elevated in SOD1(G93A)

mice, which was further increased by PGC-1α silencing. These data indicated that disruption of energy homeostasis would

Recent Activity



Search

Peroxisome proliferator activator receptor gamma coactival [Mol Neurodegener. 2011]

Disruption of skeletal muscle mitochondrial

network genes and mi [Neurobiol Dis. 2013]

and PGC [J Neuropathol Exp Neurol. 2012]

disease progression ii [Muscle Nerve. 2011]

PKC and Ryk in the s [Dev Neurobiol. 2014]

See reviews...

Turn Off Clear

See all...

Decreased mRNA expression of PGC-1a

PGC-1α protects neurons and alters

Review Altered expression of atypical

Help

_













Web Crawler Research Signals

Chemicals & Bioassays #chemicals-bioassays

BioSystems/

PubChem BioAssay /pcassay

PubChem Compound

/pccompound

PubChem Structure Search http://pubchem.ncbi.nlm.nih.gov/search/search.cgi

PubChem Substance

/pcsubstance

/guide/chemicals-bioassays/

DNA & RNA

#dna-rna

BLAST (Basic Local Alignment Search Tool) http://blast.ncbi.nlm.nih.gov/Blast.cgi

All Chemicals & Bioassays Resources...

Linked Data Assignments

Linked Pages + Resources

Application: GetSemantic

www.lib.montana.edu/~jason/files/getsemantic/ github.com/jasonclark/getsemantic

Given a linked resource, return

linked data entities + keywords

getSEMantic [beta] jason a. clark

Feed me a URL

Get Analysis



Key: Demo App What is this? View Code @jaclark Credit

getSEMantic [beta] jason a. clark

Mmmmm... Semantic Yumminess

Ranked Entities

Engineering

relevance: 0.944351

concepts & linked data resources:

- http://dbpedia.org/resource/Engineering
- http://rdf.freebase.com/ns/m.02ky346

Social sciences

relevance: 0.795123

concepts & linked data resources:

- http://dbpedia.org/resource/Social_sciences
- http://sw.opencyc.org/concept/Mx4rwauQZNc8Qdibm_jCfd9X4g



getSEMantic [beta] jason a. clark

Mmmmm... Semantic Yumminess

Ranked Keywords

Research Databases

relevance: 0.919159

State University Library

relevance: 0.911417

MSU Library Accessibility

relevance: 0.840938

Policy Mobile Site

relevance: 0.81602

Open access journal

relevance: 0.810361



Understanding the Sharing Economy

API

Application: Watching Reuse in Social Networks

www.lib.montana.edu/~jason/files/social-signal/www.lib.montana.edu/~jason/files/social-signal/index.txt





Can Tweets Predict Citations? Metrics of Social Impact Based on Twitter and Correlation with Traditional Metrics of Scientific Impact

Gunther Eysenbach^{1,2,3}, MD, MPH, FACMI

¹University Health Network, Centre for Global eHealth Innovation & Techna Institute, Toronto, ON, Canada

²Institute for Health Policy, Management, and Evaluation, University of Toronto, Toronto, ON, Canada

³JMIR Publications Inc., Toronto, ON, Canada

J Med Internet Res 2011;13(4):e123

DOI: 10.2196/jmir.2012

PMID: 22173204

PMCID: PMC3278109

□ Copy Citation to Clipboard

№ Export Metadata

END for: Endnote

BibTeX for: BibDesk, LaTeX

RIS for: RefMan, Procite, Endnote,

RefWorks

Add this article to your Mendeley library Add this article to your CiteULike library

△ Download:

Eysenbach G. Can Tweets Predict Citations? Metrics of Social Impact Based on Twitter and Correlation with Traditional Metrics of Scientific Impact. J Med Internet Res 2011;13(4):e123m DOI: 10.2196/jmir.2012

Given a linked resource, return

Altmetrics references ("citations")

Twitter or Facebook APIs ImpactStory Plum Analytics Altmetric

Social Networks Research Signals

Enter a DOI (E.g., 10.1186/s12971-015-0060-9)

Check Social

Social Networks Research Signals

10.1080/01930826.2014.946778

Check Social

Social Networks Research Signals

Demonstrating Library Value at Network Scale: Leveraging the Semantic Web With New Knowledge Work

http://dx.doi.org/10.1080/01930826.2014.946778

journal: Journal of Library Administration

doi: 10.1080/01930826.2014.946778

open access?:

facebook:

twitter: 15

readers (mendeley, citeulike, connotea).: 11

n 12

Complete details from Altmetrics.com

Filtering Research

API

Application: MSU Research Citations

arc.lib.montana.edu/msu-research-citations/github.com/msulibrary/msu-research-citations

Leila Sterman - People at MSU Library



Leila Sterman

Title Scholarly Communication Librarian

[Assistant Professor]

Department Collection Development,

Montana State University (MSU)

Phone 406-994-4519

Email leila.sterman@montana.edu

Room 149

Guides Subject and Course Help

Liaison Role(s)

Cell Biology Neuroscience

Microbiology

Immunology

Expertise and Skill(s)

Scholarly Communication

Institutional Repository

Open Access

Intellectual Property

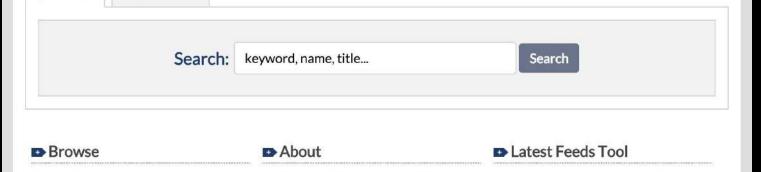


Monitor scholarship, collect/curate to understand and analyze scholarly

output

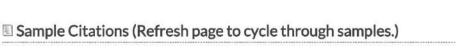
Using the Google Feed API and Data Feeds from Publishers and Aggregators

Montana State University Research Citations - Montana State University (MSU) Library Montana State University Research Publication Database Advanced Search **Quick Search** Search: keyword, name, title... Search



About the MSU Research Publication

Online Database



Keywords

Subjects

Cloninger

Multivalent scaffolds induce galectin-3 Ethnic Identity, Drinking Motives, and Alcohol Consequences Among Alaska aggregation into nanoparticles Native and Non-Native College Students Authors: Goodman, Candace K, Mark L

pea-barley forage Wofenden, Pratima Nangia-Makker, Authors: Skewes, Monica C., and Arthur Authors: Ragen, Devon L., Erin E. Nix, Anna K. Michel, Avraham Raz, and Mary J. W. Blume Rachel L. Endecott, Patrick G. Hatfield, Mark K. Petersen, and Janice G. P. Date: December 2014

Check on latest research output from

Individual mineral supplement intake by

ewes swath grazing or confinement fed

MSU

Montana State University Research Citations - Montana State University (MSU) Library

Latest Montana State University Research Citations

ScienceDirect Search: Montana State University

Title: Solar image parameter data from the SDO: Long-term curation and data mining

Author:

Published Date: Fri, 30 Oct 2015 09:18:56 -0700

Content Snippet: Publication date: Available online 29 October 2015 Source: Astronomy and Computing Author(s): M.A. Schuh, R.A. Angryk, ...

Content: Publication date: Available online 29 October 2015

Source: Astronomy and Computing

Author(s): M.A. Schuh, R.A. Angryk, P.C. Martens
The Solar Dynamics Observatory (SDO) mission captures thousands of images of the Sun per day,

motivating the need for efficient and effective storage, representation, and search over a massive repository of data. This work investigates the general-purpose image parameter data produced by the SDO Feature Finding Team's trainable module, which operates at a fixed six minute cadence over

all AIA channels. The data contains ten numerical measures computed for each image cell over a 64 × 64 grid for each image. We analyze all available data and metadata produced over the first three years and present comprehensive statistics and outliers while validating the cleanliness and usability of the data source for future research. We then utilize a database of automated solar event reports to create large-scale region-labeled datasets available to the public. We highlight the new-found potential for data-driven discovery by presenting several best-case labeling scenarios that establish a

baseline for comparing machine learning classification and attribute (image parameter) evaluation results. Future work focuses on continued dataset curation and spatiotemporal data mining.

Select a Feed

Science Direct
PLOS Collections

Medline via Web of Science

Gale Power Search

Show All Citations

Auto-populate Database

Manage Feeds

_

Open, Networked Research Signals



New Research Services

Questions?

twitter.com/jaclark www.lib.montana.edu/~jason/talks.php

Acknowledgements

I would like to thank Montana State University and my excellent colleagues at MSU Library for supporting my research.

In particular, I would like to thank Leila Sterman and James Espeland for their work in conceiving and building the MSU Research Citations Application.

I am distributing this work under the <u>Attribution-Noncommercial-Share Alike</u> <u>3.0 Unported license</u>.

