Sources of news by country [2015]

	UK	Germany	Spain	Italy	France	Denmark	Finland	USA	Urban Brazil	Japan	Ireland	Australia
TV	75%	82%	82%	78%	80%	75%	75%	64%	81%	73%	76%	72%
Radio	37%	50%	40%	23%	28%	50%	45%	26%	39%	17%	50%	41%
Printed Newspapers	38%	38%	47%	38%	19%	33%	49%	23%	33%	44%	49%	39%
Online (inc. social media)	73%	60%	86%	81%	71%	85%	90%	74%	91%	70%	83%	85%

From Reuters Institute *Digital News Report 2015*. http://www.digitalnewsreport.org/survey/2015/sources-of-news-2015/

Main source of news by age – young prefer online – all countries

Legend: 18–24, 25–34, 35–44, 45–54, 55+

Social media as a source of news – all countries

From Reuters Institute *Digital News Report 2015*. http://www.digitalnewsreport.org/survey/2015/sources-of-news-2015/
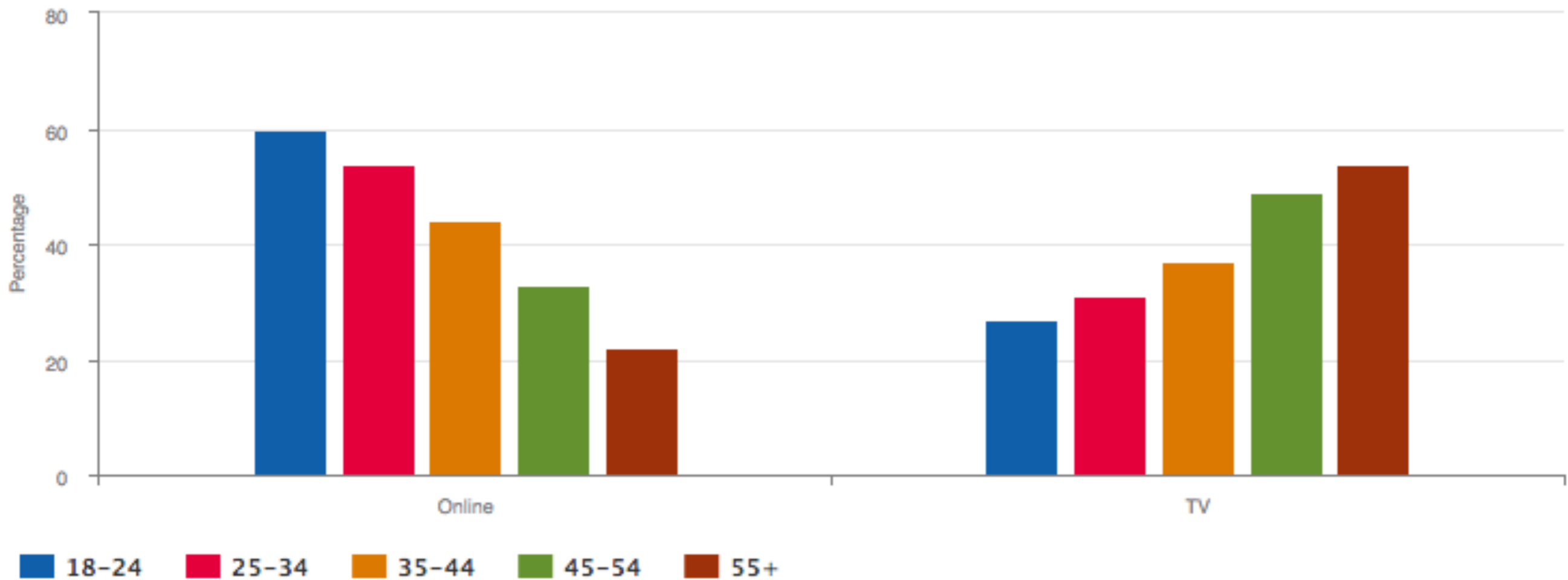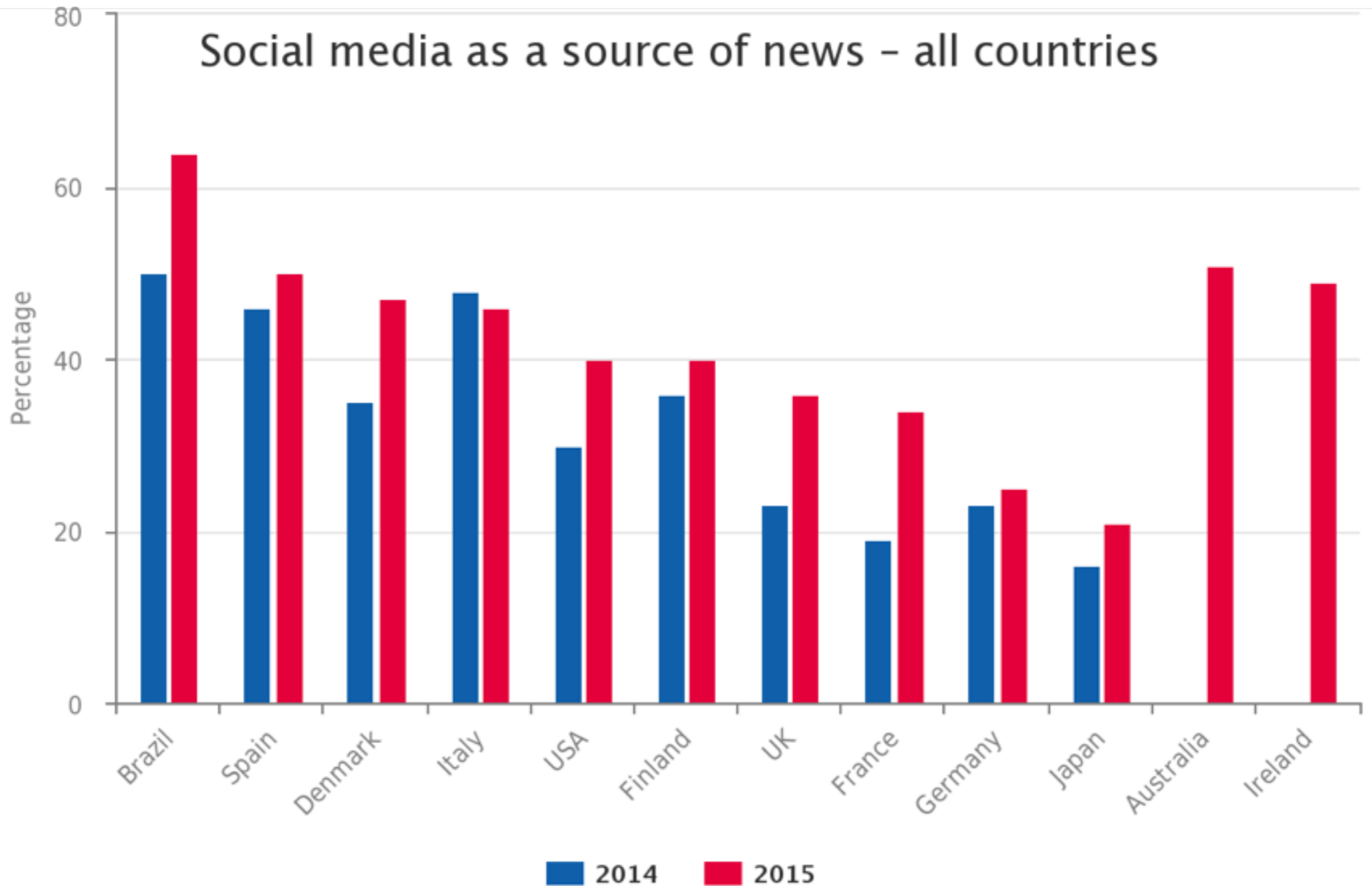
February's Most Shared Sites

Facebook Shares

Site	Shares
The Huffington Post	4,631,287
BuzzFeed	3,354,310
Fox News	3,043,063
The Guardian	2,905,813
The New York Times	2,805,191
NBC	2,607,382
BBC	2,565,922
Washington Post	2,432,556
Daily Mail	2,245,179
CNN	1,984,541

0K   1,000K   2,000K   3,000K   4,000K   5,000K
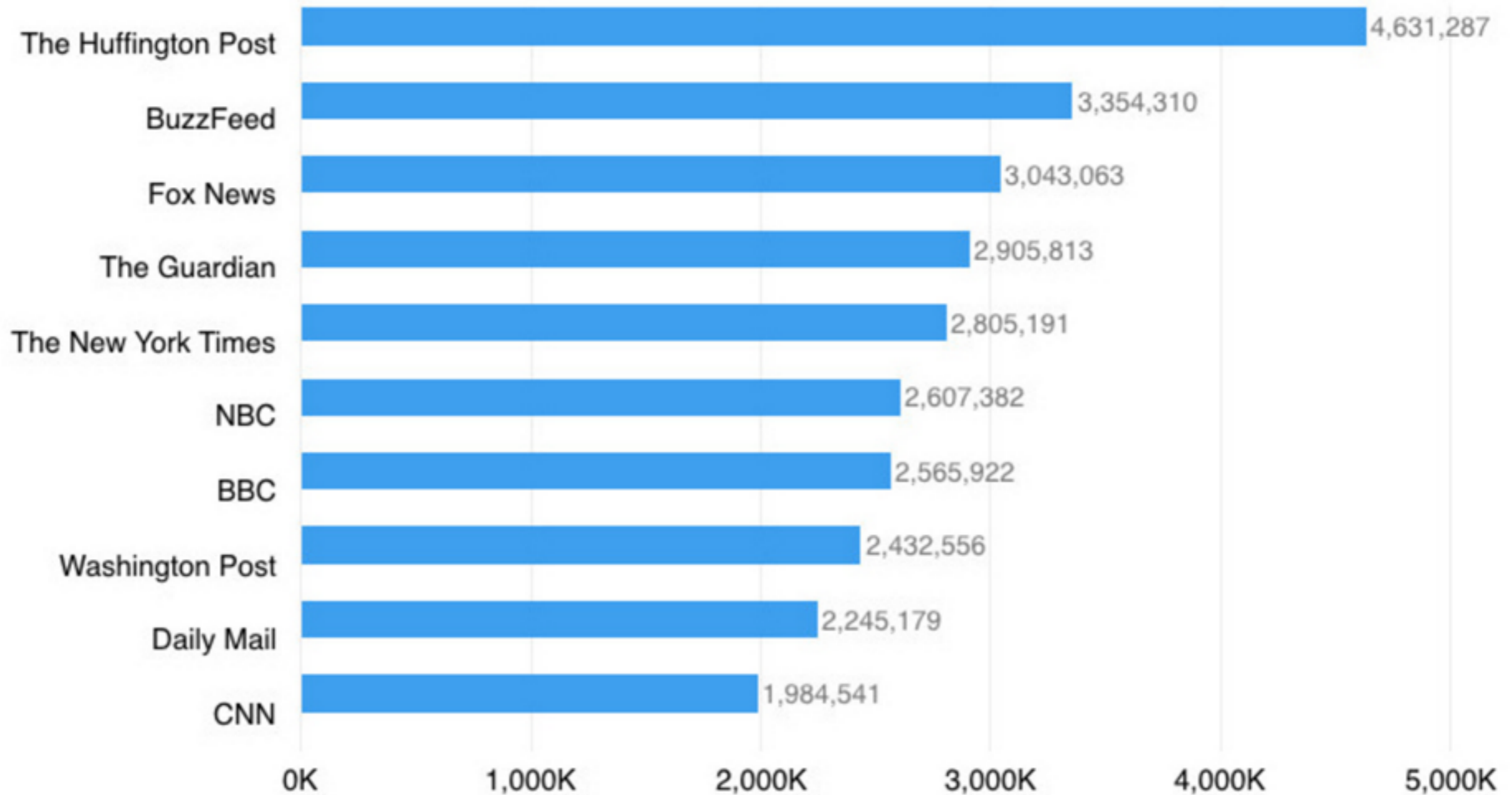
Based on content published in Feb. 2016 only.

NEWS WHIP

NewsWhip. https://www.newswhip.com/2016/03/biggest-facebook-publishers-february-2016

Digital news metadata: Publisher and library perspectives

Frederick Zarndt
IFLA Governing Board
Digital Divide Data
frederick@frederickzarndt.com
@cowboyMontana

Digital news metadata: Publisher ~~and library~~ perspectives

Frederick Zarndt
IFLA Governing Board
Digital Divide Data
frederick@frederickzarndt.com
@cowboyMontana

Why
digital news metadata?

"Once I know WHY, I can figure out what I can do about that ... or if I CAN do anything about that."

Overview

- IPTC: The Methuselah of digital news
- Google News
- schema.org
- Facebook's open graph protocol and instant news
- RSS: Methuselah's younger sibling
- Metadata crosswalk
- Issues vs. stories vs. websites

IPTC International Press Telecommunications Council: Methuselah of Digital News Standards

Digital News, Enabled.

The IPTC is the global standards body of the news media. We provide the technical foundation for the news ecosystem.

OUR STANDARDS   FOR DEVELOPERS

Photo Metadata

Media Topics

NewsML-G2

RightsML

From IPTC https://iptc.org/

I can't imagine being a professional photographer without using IPTC Photo Metadata. It makes my life easy: I can search my database efficiently and I don't lose track of photos.

Hardy Klahold, Photographer and Owner, Hardy Klahold Photography

● ● ○ ● ●

From IPTC https://iptc.org/

The IPTC had significant impact and leadership in the development of the W3C ODRL specifications that is the underlying framework for RightsML. Their insight and experience has driven both standards to be world-leading in the domain of rights management.

Renato Iannella, Semantic Identity and Chair, W3C ODRL Community Group

• • • ○ •

Our success is largely connected to the worldwide adoption of the IPTC Standards, upon which our products are heavily dependent. They allow us to connect to and integrate with anyone and anything - a fundamental requirement for succeeding with software products.

Christopher Frenning, CEO of FotoWare

• • • • ○

From IPTC https://iptc.org/

> IPTC was instrumental in developing the rNews standard, now part of the Schema.org standard, and Parse.ly uses this markup as the basis for our content analytics platform. As a result, major news publishers globally have integrated semantic data into their websites, and the entire web is better for it!

Andrew Montalenti, Co-Founder & CTO, Parse.ly

• ○ • • •

From IPTC https://iptc.org/

"

The IPTC brought openness, flexibility, tenacity and deep domain expertise to their participation in the Schema.org effort, and our cooperation set the standard for all others that followed.

R.V. Guha, Google Fellow and Co-founder, Schema.org

○ • • • •

From IPTC https://iptc.org/

IPTC Standards

		Released	Last update
IPTC 7901		1979	1995
IPTC Information Interchange Model (IIM)		1991	2014
IPTC NITF [XML]		early 1990s	2012
IPTC Photo Metadata		early 1990s	2014
IPTC Subject Codes / Media Topics [taxonomy]	~1400 terms	early 1990s	2015
IPTC NewsML 1 [XML]		2001	2015
IPTC NewsML-G2 [XML]		2011	2015
IPTC SportsML-G2 [XML]		2008	2011
IPTC RightsML [XML]		2013	2013
IPTC rNews [HTML]		2011	2013
IPTC ninjs [JSON]		2014	

IPTC standards are released under Creative Commons Attribution 3.0 (https://creativecommons.org/licenses/by/3.0/)

IPTC Photo Metadata

Data and information is entered into the image file by users or by automated capture from cameras or scanners.

- There are 31 IPTC Core metadata fields, 46 Extension fields

- IPTC photo metadata is primarily descriptive, administrative, and rights

- Metadata is stored in two main places

  - Internally – embedded in the image file in formats such as JPEG or TIFF.

  - Externally – outside the image file in a digital asset management system (DAM) or by a "sidecar" file, such as XMP, or an external XML-based news exchange format file.

IPTC NITF
News Industry Text Format

- Facilitates the exchange of text news. Multimedia "attachments" are supported
- Published as a XML-based standard in 1998
- Released version 3.4 in May 2007 as DTD and XML schema
- Supports identification and description of news
  - Who owns the copyright to the item, who may republish it, and who it's about.
  - What subjects, organizations, and events it covers.
  - When it was reported, issued, and revised.
  - Where it was written, where the action took place, and where it may be released.
  - Why it is newsworthy, based on the editor's analysis of the metadata.

nitf

head

- title: document title
- meta: construct for sending generic metadata
- tobject: subject code identifying type of news based on Subject Code system
- im: IIM record data container
- docdata: container for document metadata
- pub-data: metadata about how news object was used
- revision-history: creative history of document including audit trail

body

body.head
- hedline: headline and sub-headlines
- rights: information on rights, rights holder, rights type, rights geography, etc
- byline: container for structured or unstructured byline information
- dateline: date and location where story was created
- abstract: abstract summary or synopsis of document

body.content
- p: paragraph
- media: media object (audio, video, image, etc) with caption and credit

body.end
- tagline: byline at end of story

IPTC rNews

... news providers have created feeds to supply news using IPTC formats such as NITF and NewsML-G2. ... there are an increasing number of consumers of news who only want to work with "pure" web technologies, i.e. HTML rather than XML. ...IPTC has been looking at the two major paths to represent metadata in HTML - microformats and RDFa.

- **hNews** is the microformat for news that was adopted by the [news] community in late 2009. It builds upon hAtom by adding a few news-specific fields.

- **rNews** is a proposal for a semantic markup vocabulary for news. ... rNews and hNews are similar in intent (news-specific metadata in HTML) but somewhat different in approach. Whereas hNews went through the microformats process, a semantic markup vocabulary can be created by anyone. The IPTC recently adopted rNews 1.0, based somewhat on the NewsML-G2, NITF and hNews models but extending beyond those standards ...

- **rNews** is an approved standard for using semantic markup to annotate news-specific metadata in HTML documents.

- **rNews** specifies the terminology and data model required to embed news specific metadata into HTML documents

- **rNews** uses RDFa triples or Microdata

IPTC rNews

- **rNews** provides a basic semantic markup of the content of a web page.

- **rNews** apply to the portion of feed workflows where publishers target primarily the end user but also any business which makes use of the public presentation of content on the web [search engines].

- **rNews** provides no or only a very limited set of metadata for managing content. [Managing content] is the domain of the G2-Standards NewsML-G2 and EventsML-G2, of NITF and NewsML 1.

- **rNews** is not made for delivering packages of structured news items in a business-to-business context, but it can reflect the composite content of a web page.

Google news

News

Deutschland-Ausgabe ▾

Schlagzeilen

Schlagzeilen »

Borussia Mönchengladbach

Griechenland

Thomas Schaaf

TSG 1899 Hoffenheim

Tatort

Panama

ZDF

Novak Đoković

Borussia Dortmund

Aserbaidschan

Phnom Penh, Phnom Pe...

Vorschläge für Sie

International

Deutschland

Wirtschaft

Wissen/Technik

Unterhaltung

Sport

Gesundheit

Panama Papers : Diese Staats- und Regierungschefs tauchen in den Dokumenten auf

FOCUS Online - vor 52 Minuten   G+ ✉ 🐦 f ✉

Es ist ein Paukenschlag: Durch ein Datenleck bisher einmaligen Ausmaßes sind Briefkastenfirmen von Politikern und Prominenten auf der ganzen Welt aufgeflogen. Darunter mehrere amtierende oder ehemalige Staats- und Regierungschefs. Das Wichtigste ...

FOCUS Online

Echtzeit-Berichte

Staatsanwaltschaft ermittelt zu "Panama Papers" DIE WELT

"SZ" deckt Geflecht von Briefkastenfirmen in Panama auf Reuters Deutschland

Häufig zitiert: Panama Papers - Geheimgeschäfte von Hunderten Politikern und Prominenten in ...
Süddeutsche.de

Meinung: "Panama Papers" alarmieren Ermittlungsbehörden Deutsche Welle

Ähnlich
Panama »
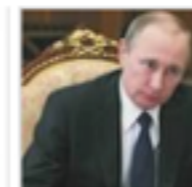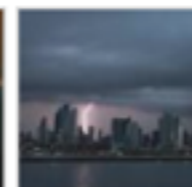
euronews   FAZ - Frank...   n-tv.de NAC...   FOCUS On...   FOCUS On...   Süddeutsch...   FOCUS On...

+++ Flüchtlingskrise im News-Ticker +++: Griechenland beginnt mit ...

FOCUS Online - vor 1 Stunde

Die Türkei soll einen zentralen Beitrag bei der Bewältigung der Flüchtlings-Krise leisten. Dafür hat die EU Ankara einige Zugeständnisse gemacht. Doch nun muss sich zeigen, ob die Türkei in der Lage ist, den Zustrom zu begrenzen. Lesen Sie die ...
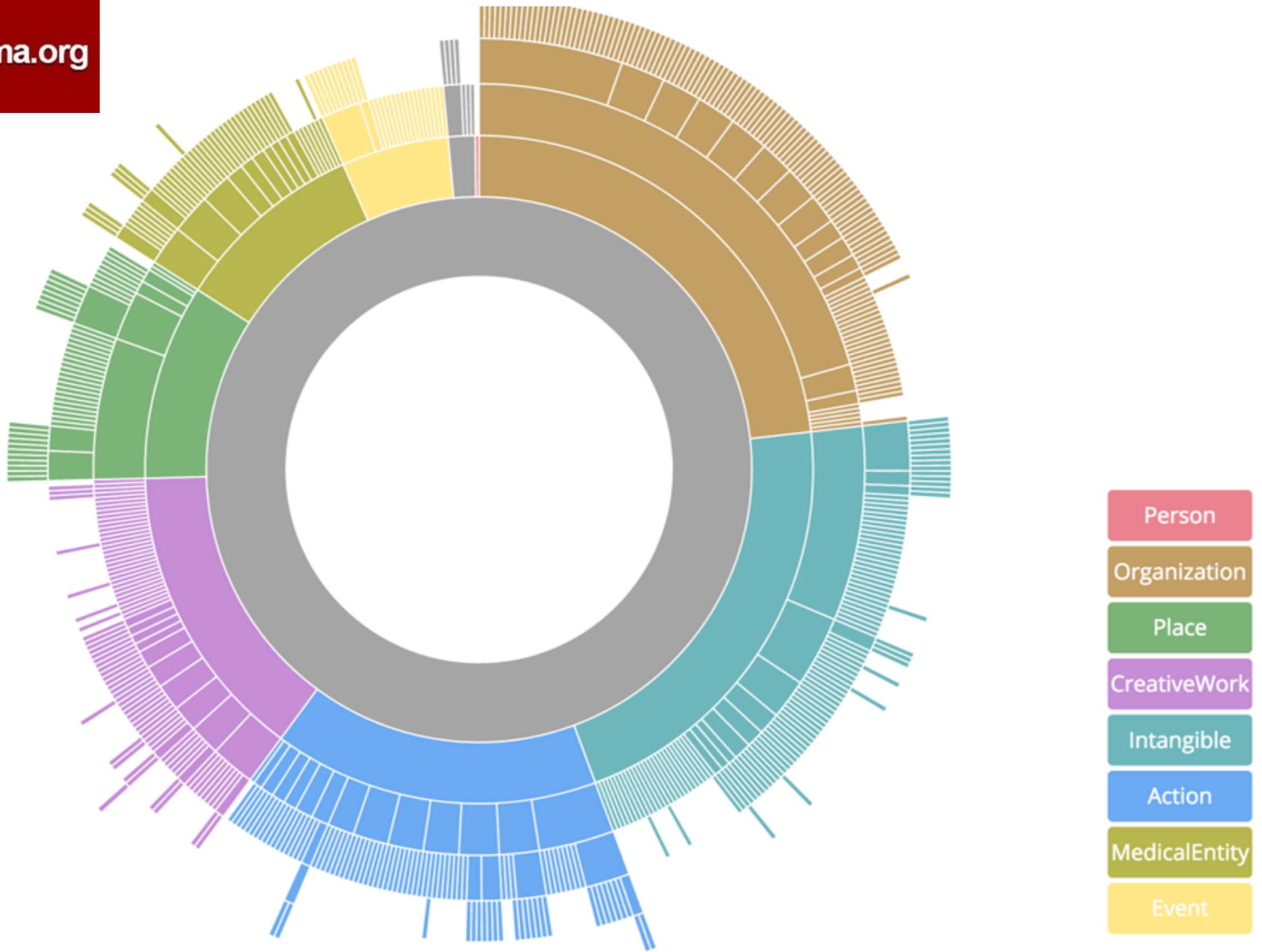
FOCUS Onli...

Stretegiewechsel: Stoiber fordert von Rechtsruck von der CDU

FOCUS Online - vor 13 Minuten

Nach den CDU-Verlusten bei den jüngsten Landtagswahlen hat der CSU-Ehrenvorsitzende Edmund Stoiber von der Schwesterpartei einen grundlegenden Strategiewechsel gefordert. «Ja, es geht auch um eine Kursänderung», sagte der 74-Jährige dem ...

FOCUS Onli...

Google news

- Mark up your articles using the following properties of the **https://schema.org/Article** type. Be sure to mark up your page with the most specific applicable schema.org type. For example, a news article should be marked up as **https://schema.org/NewsArticle** and a blog post as **https://schema.org/BlogPosting**.

- Be unique. Each of your pages that display an article's full text needs to have a unique URL. ...

- Be permanent. ... In order to ensure that our links to articles function properly, each article on your news site needs to be associated with one unique URL, and that URL must be permanent (i.e., it can't be recycled).

- Additionally, do not republish previously published articles under a new URL.

- Google News RSS feeds are available in RSS 2.0 format.

Adapted from https://support.google.com/news/publisher

Person
Organization
Place
CreativeWork
Intangible
Action
MedicalEntity
Event

From http://blog.schema.org/2015/11/schemaorg-whats-new.html

schema.org

schema.org is a collaborative, community activity with a mission to create, maintain, and promote schemas for structured data on the Internet, on web pages, in email messages, and beyond.

schema.org vocabulary can be used with many different encodings, including RDFa, Microdata and JSON-LD. These vocabularies cover entities, relationships between entities and actions, and can easily be extended through a well-documented extension model. Over **10 million sites** use schema.org to markup their web pages and email messages. Many applications from Google, Microsoft, Pinterest, Yandex and others already use these vocabularies...

Adapted from https://schema.org/

Thing>CreativeWork>Article>NewsArticle

Used in 10,000 to 50,000 domains

Property	Expected Type	Description
Properties from NewsArticle		
dateline	Text	The location where the NewsArticle was produced.
printColumn	Text	The number of the column in which the NewsArticle appears in the print edition.
printEdition	Text	The edition of the print product in which the NewsArticle appears.
printPage	Text	If this NewsArticle appears in print, this field indicates the name of the page on which the article is found. Please note that this field is intended for the exact page name (e.g. A5, B18).
printSection	Text	If this NewsArticle appears in print, this field indicates the print section in which the article appeared.
Properties from Article		
articleBody	Text	The actual body of the article.
articleSection	Text	Articles may belong to one or more 'sections' in a magazine or newspaper, such as Sports, Lifestyle, etc.
pageEnd	Text or Integer	The page on which the work ends; for example "138" or "xvi".
pageStart	Text or Integer	The page on which the work starts; for example "135" or "xiii".
pagination	Text	Any description of pages that is not separated into pageStart and pageEnd; for example, "1-6, 9, 55" or "10-12, 46-49".
wordCount	Integer	The number of words in the text of the Article.
Properties from CreativeWork		
about	Thing	The subject matter of the content.
accessibilityAPI	Text	Indicates that the resource is compatible with the referenced accessibility API (WebSchemas wiki lists possible values).
accessibilityControl	Text	Identifies input methods that are sufficient to fully control the described resource (WebSchemas wiki lists possible values).
accessibilityFeature	Text	Content features of the resource, such as accessible media, alternatives and supported enhancements for accessibility (WebSchemas wiki lists possible values).
accessibilityHazard	Text	A characteristic of the described resource that is physiologically dangerous to some users. Related to WCAG 2.0 guideline 2.3 (WebSchemas wiki lists possible values).

From https://schema.org/NewsArticle

schema.org

rNews is here. And this is what it means.
By EVAN SANDHAUS   FEBRUARY 16, 2012 12:26 PM

On January 23rd, 2012, The Times made a subtle change to articles published on nytimes.com. We rolled out phase one of our implementation of rNews – a new standard for embedding machine-readable publishing metadata into HTML documents. Many of our users will never see the change but the change will likely impact how they experience the news.

As of April 2016, NewsArticle ...

• ... has 5 unique (not inherited) properties

• ... has 83 properties inherited from Article, CreativeWork, and Thing

• ... may appear as an AssociatedArticle property on MediaObjects.

From https://schema.org/NewsArticle
From http://open.blogs.nytimes.com/2012/02/16/rnews-is-here-and-this-is-what-it-means

schema.org/NewsArticle Applied

On 22 March 2016, the Washington Post published this story on its website @ https://goo.gl/dEoPTR

schema.org/NewsArticle Applied

<!DOCTYPE html> <html itemscope itemtype="**http://schema.org/NewsArticle**" lang="en" class="blog"> <head>

...

<h1 itemprop="**headline**">The simple, brilliant plan to change macaroni and cheese that duped us all</h1>

...

itemprop="author" itemscope itemtype="**http://schema.org/ Person**">By Roberto A. Ferdman

...

Follow schema.org

Welcome to Schema.org

Schema.org is a collaborative, community activity with a mission to create, maintain, and promote schemas for structured data on the Internet, on web pages, in email messages, and beyond.

Schema.org vocabulary can be used with many different encodings, including RDFa, Microdata and JSON-LD. These vocabularies cover entities, relationships between entities and actions, and can easily be extended through a well-documented extension model. Over 10 million sites use Schema.org to markup their web pages and email messages. Many applications from Google, Microsoft, Pinterest, Yandex and others already use these vocabularies to power rich, extensible experiences.

Schema.org is sponsored by Google, Microsoft, Yahoo and Yandex. The vocabularies are developed by an open community process, using the public-schemaorg@w3.org mailing list and through GitHub.

A shared vocabulary makes it easier for webmasters and developers to decide on a schema and get the maximum benefit for their efforts. It is in this spirit that the sponsors, together with the larger community have come together, to provide a shared collection of schemas.

We invite you to get started!

View our blog at blog.schema.org or see release history.

Follow schema.org on Github
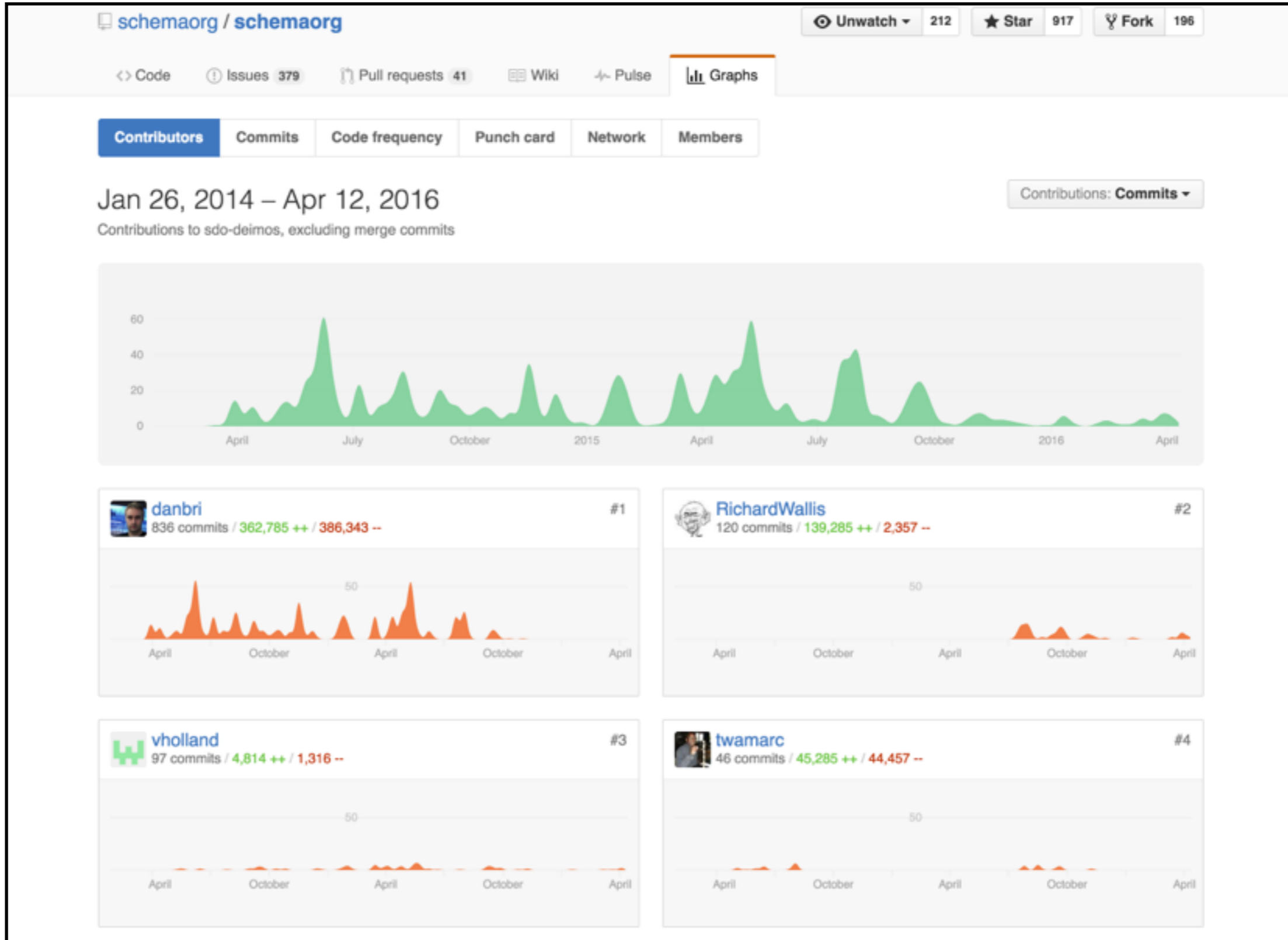
Facebook is the No.1 Social Traffic Source for News Websites

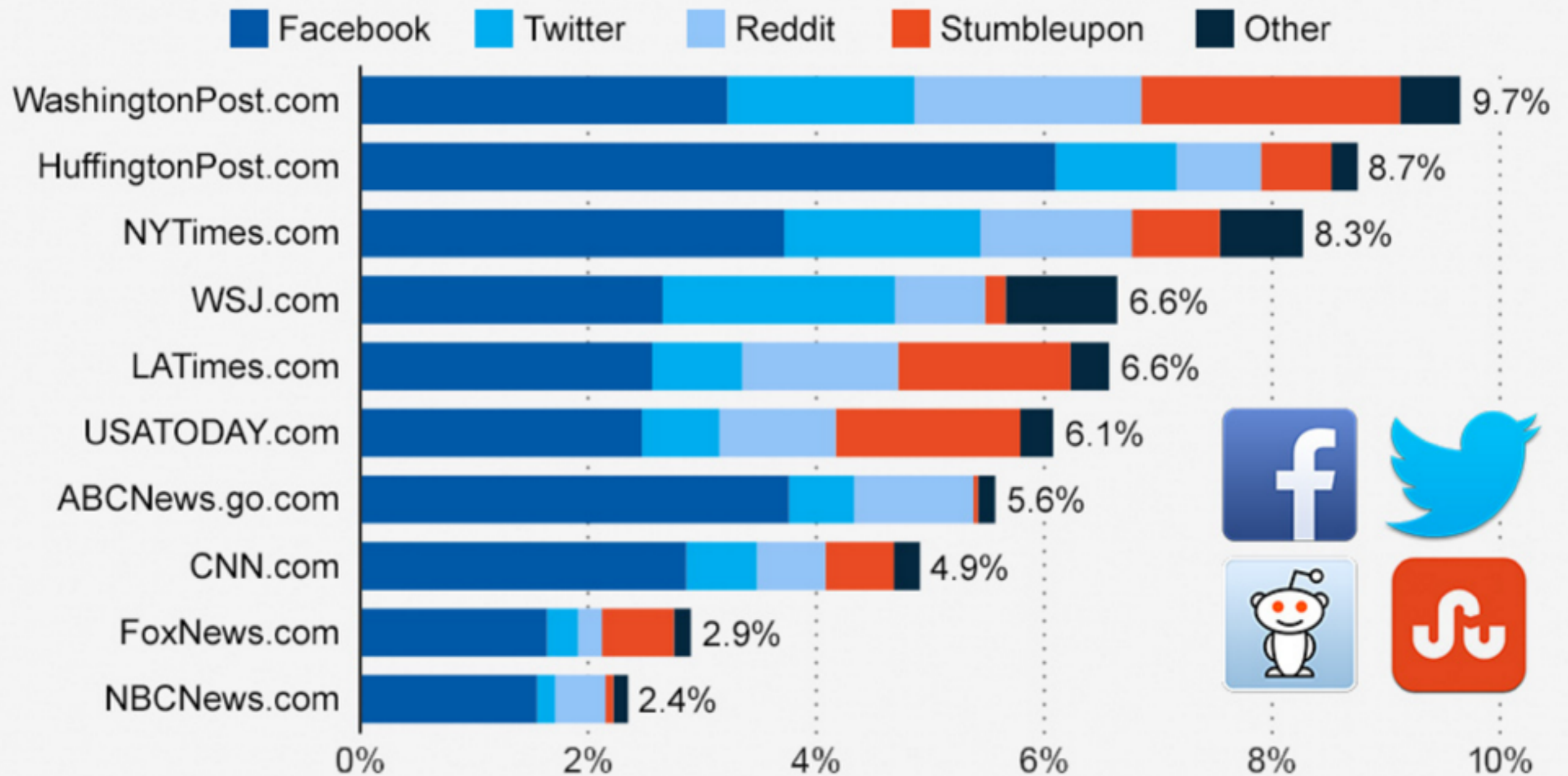Percentage of website traffic coming from social media sites in the past three months

Legend: Facebook | Twitter | Reddit | Stumbleupon | Other

Website	Total
WashingtonPost.com	9.7%
HuffingtonPost.com	8.7%
NYTimes.com	8.3%
WSJ.com	6.6%
LATimes.com	6.6%
USATODAY.com	6.1%
ABCNews.go.com	5.6%
CNN.com	4.9%
FoxNews.com	2.9%
NBCNews.com	2.4%

Axis: 0% 2% 4% 6% 8% 10%

statista
The Statistics Portal

Mashable

Source: SimilarWeb

From Statista the Statistics Portal. 2013. https://www.statista.com/chart/1324/social-media-traffic-of-news-websites/

Facebook Is a Must for Brands Around the World

Percentage of brands present on the following platforms (June 2014)

Platform	Percentage
facebook.	100.0%
Twitter	96.3%
YouTube	94.2%
Instagram — Fast beautiful photo sharing	90.8%
g+	79.1%
Pinterest	75.4%
新浪微博 weibo.com	59.7%
WeChat	44.0%
youku 优酷 .com	39.5%
tumblr.	29.3%

n=382 brands

@StatistaCharts   Source: L2

statista

From Statista the Statistics Portal. 2014. https://www.statista.com/chart/2942/brand-use-of-social-media-platforms/
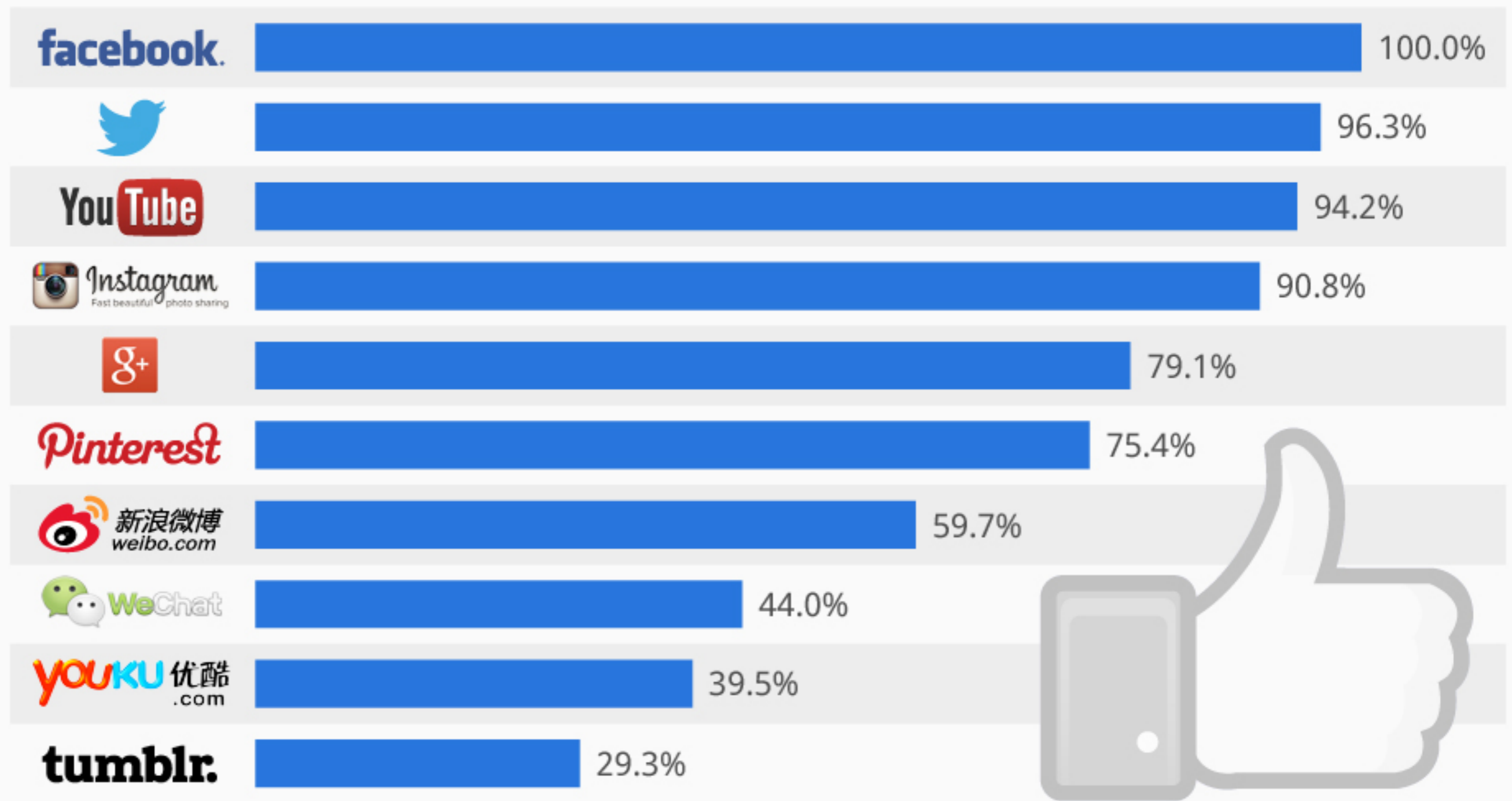
TOP SOCIAL NETWORKS*

Facebook	40%
YouTube	16%
Twitter	11%
Google+	5%
Reddit	4%

*used weekly for news

DIGITAL PARTICIPATION

32% share a news story via email or social media. US is = **4th out of 12** in overall participation index.

TOP SOCIAL NETWORKS*

|	
Facebook	23%
YouTube	13%
WhatsApp	9%
Google+	6%
Twitter	4%

*used weekly for news

Germans are less interested in news-related participation via social media than people in other countries. Facebook is still the biggest network for news while Twitter attracts media coverage but has struggled to appeal to the wider public. WhatsApp has been growing fast in Germany over the last few years and some local newspapers have been experimenting with it for distributing their stories.

From Reuters Institute *Digital News Report 2015*. http://www.digitalnewsreport.org/survey/2015/sources-of-news-2015/

Open Graph Protocol (OG)

- OG enables any web page to become an object in a social graph.

- News articles, blog posts, videos, audio clips, websites, and more can be described by OG.

- OG is used or recognized by Facebook, Google, Pinterest, Linkedin, Twitter, WordPress, and others.
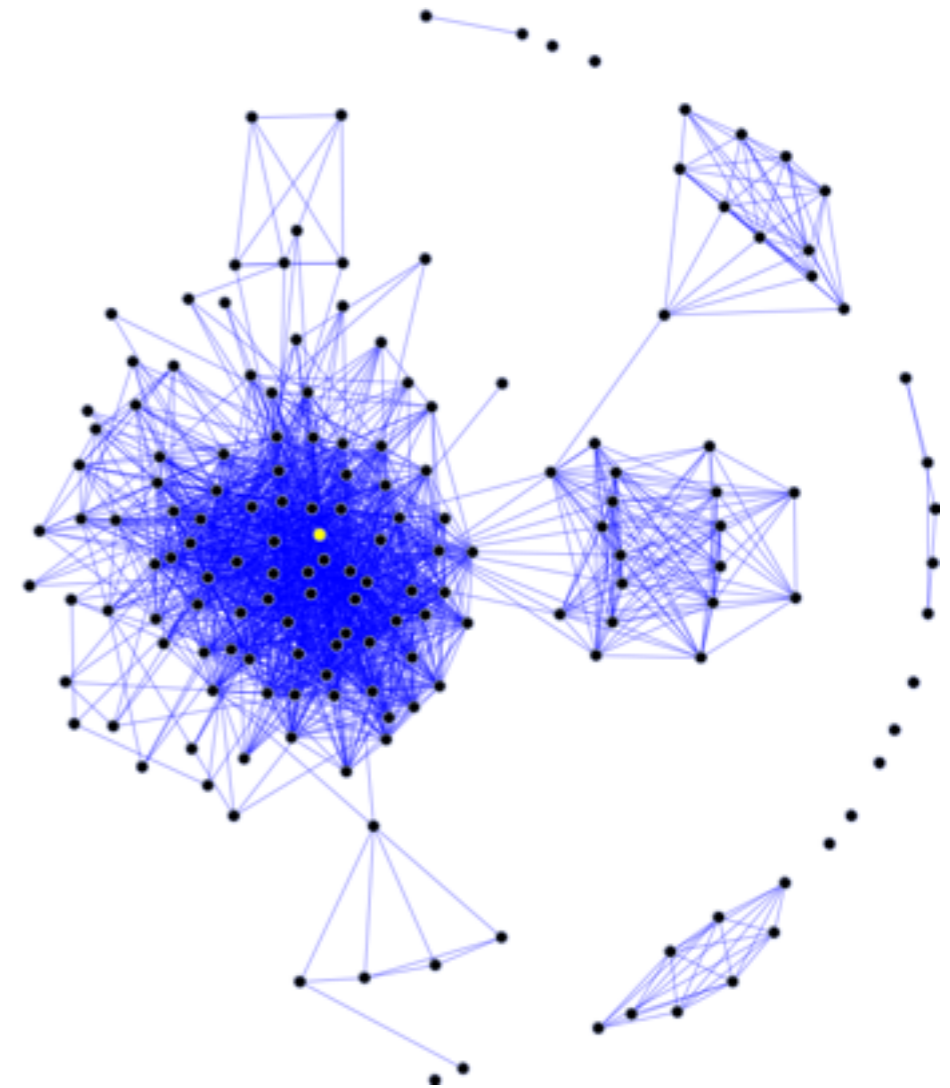
Basic OG Tags

The four required metadata properties for every OG object

- **og:title** - The title of the object as it should appear within the graph, e.g., "The Rock".

- **og:type** - The type of the object, e.g., "video.movie". Depending on the type specified, other properties may also be required.

- **og:image** - An image URL which should represent the object within the graph.

- **og:url** - The canonical URL of the object. This will be used as its permanent ID in the graph, e.g., "http://www.imdb.com/title/tt0117500/".

Other Tags

Other metadata properties that may be attached to an OG news story

- **article:author** - Array of Facebook profile URLs or IDs of the authors for this article

- **article:expiration_time** - A time when the article expired (or will expire)

- **article:modified_time** - A time when the article was last modified

- **article:published_time** - A time when the article was published

- **article:publisher** - A Facebook page URL or ID of the publishing entity

- **article:section** - The section of your website to which the article belongs, such as 'Lifestyle' or 'Sports'

- **article:tag** - An array of keywords relevant to the article

Adapted from http://ogp.me/#type_article

FaceBook Instant News

- An Instant Article is a HTML5 document optimized for fast mobile performance, rich storytelling capabilities, branded design and customized visual display.

- Every article published as an Instant Article must be published on a news publisher's website as well. That means when someone shares an Instant Article by email or on Twitter, they also share a link to the publisher's website.

From https://developers.facebook.com/docs/instant-articles

Instant News Applied

On 22 March 2016, the Washington Post published this story on its website @ https://goo.gl/dEoPTR

Wonkblog

The simple, brilliant plan to change macaroni and cheese that duped us all

A | 🖶 | 💬 72 | 🔖 Save for Later | ☰ Reading List

By **Roberto A. Ferdman** March 22 at 1:16 PM ✉ 🐦 Follow @robferdman

(Flickr user The Marmot on a Creative Commons license)

Last April, Kraft announced that it would eventually change its famous macaroni and cheese product. The plan was to get rid of artificial preservatives and replace artificial dyes with a trio of spices: paprika, annatto and turmeric. But there was no timetable — at least none shared publicly.

Most Read

1 Why smart people are better off with fewer friends

2 We had all better hope these scientists are wrong about the planet's future

3 Why it seems impossible to buy your first home

4 What we're doing to the Earth has no parallel in 66 million years, scientists say

5 The simple, brilliant plan to change macaroni and cheese that duped us all

Market Watch

DJIA **-0.23%**   NASDAQ **0.27%**

Co. name or symbol   Get quote

Last Update: 4:31 PM 03/22/2016(DJIA&NASDAQ)

Instant News Applied

As shared on Facebook, the story had the properties below (and many others, some OG, some schema.org, some HTLM5) https://goo.gl/dEoPTR

<meta property="og:description" content="Macaroni and cheese has changed--and nobody noticed."/>

<meta property="og:type" content="article"/> <meta property="og:site_name" content="Washington Post"/>

<meta property="og:title" content="The simple, brilliant plan to change macaroni and cheese that duped us all"/>

<meta itemprop="image" property="og:image" content="http://www.washingtonpost.com/blogs/wonkblog/files/2016/03/15426457201_3f7f18a168_o.jpg"/>

<meta property="article:publisher" content="https://www.facebook.com/washingtonpost">

<meta property="og:url" content="https://www.washingtonpost.com/news/wonk/wp/2016/03/22/this-brilliant-plan-to-change-macaroni-and-cheese-that-duped-us-all/"/>

Instant News Applied

On the Washington Post website

On Facebook

From Washington Post 22 March 2016: https://goo.gl/dEoPTR

Our latest Freakonomics Radio *episode is called "The Economics of Sleep, Part 1." (You can subscribe to the podcast at* **iTunes** *or* **elsewhere***, get the* **RSS feed***, or listen via the media player above.)*

The gist: Could a lack of sleep help explain why some people get much sicker than others?

Below is a transcript of the episode, modified for your reading pleasure. For more information on the people and ideas in the episode, see the links at the bottom of this post. And you'll find credits for the music in the episode noted within the transcript.

*   *   *

This week we're bringing you an episode from our archives. It's called, "The Economics of Sleep, Part 1." And yes, that means that there's a Part 2, which you'll hear next week. We thought it was time to replay these episodes because they are two of our most popular episodes ever. Now, why is that? I think it may be because as much as people tend to focus on nutrition and exercise as the vital inputs in maintaining the human machine, sleep often gets overlooked. So, let's stop overlooking it, yeah? Hope you enjoy, and I hope you learn as much about sleep as we did in making this episode.

*   *   *

RSS

Really Simply Syndication or *Rich Site Summary* (RSS) is a family of standard (http://www.rss-specifications.com/) web feed formats to publish frequently updated information: blog entries, news headlines, audio, video. An RSS document (called "feed", "web feed", or "channel") includes full or summarized text, and metadata, like publishing date and author's name.

RSS was 1st used by Netscape in 1999. RSS is now administered by the RSS Advisory Board (http://www.rssboard.org/)

Atom is a similar, competing syndication format which addresses some of the perceived problems with RSS. According to IETF RFC 4287

Atom is an XML-based document format that describes lists of related information known as "feeds". Feeds are composed of a number of items, known as "entries", each with an extensible set of attached metadata.

RSS definition from https://en.wikipedia.org/wiki/RSS

Atom Syndication Format IETF RFC 4287 https://tools.ietf.org/html/rfc4287

RSS Channel Elements

The three **required** elements for every RSS channel

- **rss:channel:title** - Name of the channel. If the publisher has an HTML website that contains the same information as its RSS file, the title of the publisher's channel should be the same as the title of its website.

- **rss:channel:link** - URL to the HTML website corresponding to the channel.

- **rss:channel:description** - Phrase or sentence describing the channel.

A RSS channel may also have a number of **optional** elements such as **copyright**, **pubDate**, **category**, **language**, **ttl**, etc

RSS Item Elements

The elements for that can be used to describe each item published on the RSS channel. Multiple items may be published on a channel; for each item at least **title** or **description** **must** appear.

- **rss:item:title** - Title of the item. For each RSS item at least title or description

- **rss:item:link** - URL to the HTML website on which the item is published.

- **rss:item:description** - Synopsis of the item.

- **rss:item:author** - Email address of the author of the item.

- **rss:item:guid** - A string that uniquely identifies the item.

A RSS item may also have a number of other **optional** elements such as **category**, **pubDate**, **source**, etc

RSS 2.0

- Many, many news publishers, blogs, etc, publish RSS channels.
- RSS 2.0 is published under Creative Commons Attribution-ShareAlike Generic CC-BY-SA-1.0 (https://creativecommons.org/licenses/by-sa/1.0/).
- The RSS 2.0 specification is owned and maintained by the Berkman Center for Internet & Society at Harvard Law School (http://www.rss-specifications.com/).
- W3C RSS Reference is at http://www.w3schools.com/xml/xml_rss.asp.
- *Lots* of RSS tools and feed readers are available. See for example http://www.rss-tools.com/ and http://www.rss-specifications.com/rss-software.htm.

A Simple RSS Feed

```
<?xml version="1.0" encoding="UTF-8" ?>
<rss version="2.0">
<channel>
 <title>RSS Title</title>
 <description>This is an example of an RSS feed</description>
 <link>http://www.example.com/main.html</link>
 <lastBuildDate>Mon, 06 Sep 2010 00:01:00 +0000 </lastBuildDate>
 <pubDate>Sun, 06 Sep 2009 16:20:00 +0000</pubDate>
 <ttl>1800</ttl>
 <item>
  <title>Example entry</title>
  <description>Here is some text containing an interesting
description.</description>
  <link>http://www.example.com/blog/post/1</link>
  <guid isPermaLink="true">7bd204c6-1655-4c27-
aeee-53f933c5395f</guid>
  <pubDate>Sun, 06 Sep 2009 16:20:00 +0000</pubDate>
 </item>
</channel>
</rss>
```

Metadata Crosswalk

|	MODS	rNews	NITF	schema.org	RSS 2.0
Publication title	MODS:title			CreativeWork:publication	RSS:channel:title
Publication ID	MODS:identifier (ISSN)	NewsItem:identifier	NITF:pubdata	Thing:url	RSS.channel:link
Publisher	MODS:publisher	NewsItem:sourceOrganization	NITF:distributor NITF:bytti	CreativeWork:producer CreativeWork:publisher CreativeWork:sourceOrganizaton	RSS:channel:description
Place of publication	MODS:place	Organization:address NewsItem:dateline	NITF:pubdata	NewsArticle:dateline	RSS.channel:link
Publication date	MODS:dateIssued	NewsItem:dateCreated NewsItem:datePublished NewsItem:dateModified	NITF:date.issue NITF:date.release NITF:dateexpire NITF: dateline	CreativeWork:dateCreated CreativeWork:dateModified CreativeWork:datePublished	RSS:channel:pubDate
Article title	MODS:title	NewsItem:headline	NITF:title	CreativeWork:headline	RSS.item.title
Article author	MODS:name MODS:role	NewsItem:creator	NITF:byline NITF:tagline	CreativeWork:author CreativeWork:creator	RSS:item.author
Article identifier	MODS:genre	NewsItem:contributor	NITF:doc-id	CreativeWork:contributor	RSS.item.guid
Article subject	MODS:subject	NewsItem:about NewsItem:genre	NITF:identified-content NITF:abstract	CreativeWork:genre	RSS.item.category

Digital News Metadata Standards Are Alive
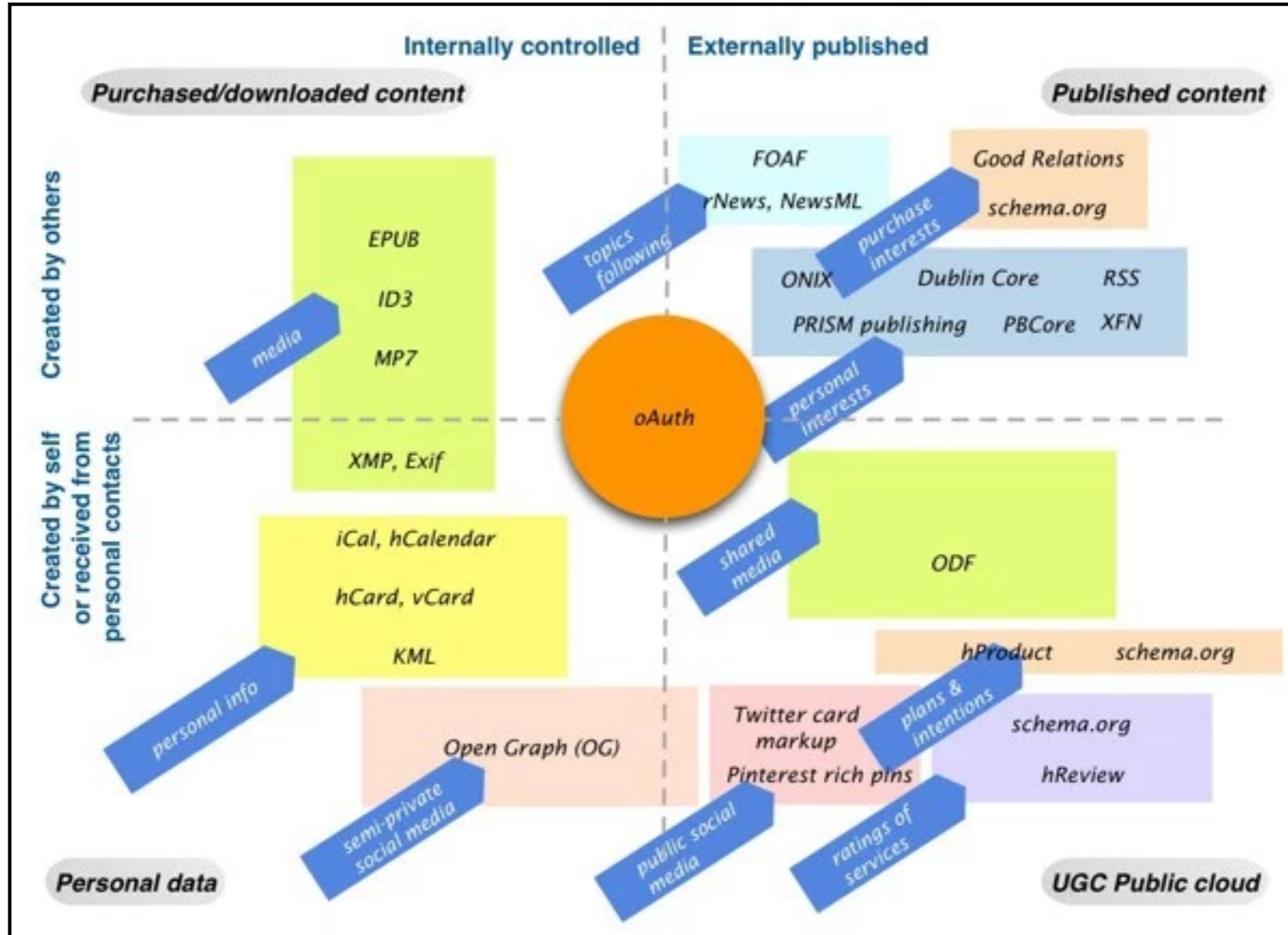
Image from Michael Andrews. Story Needle blog. 2015.
http://storyneedle.com/metadata-standards-and-content-portability/

A Brief Standards Release History

	Released	Last update
IPTC 7901	1979	1995
IPTC Information Interchange Model (IIM)	1991	2014
IPTC NITF [XML]	early 1990s	2012
IPTC Photo Metadata	early 1990s	2014
IPTC Subject Codes / Media Topics [taxonomy]	early 1990s	2015
IPTC NewsML / SportsML / EventsML [XML]	2001	2015
IPTC rNews [HTML]	2011	2015
IPTC RightsML [XML]	2013	2013
schema.org	spring 2012	late 2015
Open Graph Protocol / API / Instant News	2010	fall 2015
RSS	1997	2003
MODS	2002	spring 2015

Feb 24 : Keep Metadata of Photos Alive and Intact

Metadata is of high value for many parties in the photo business: photographers, libraries, agencies, archives. But many consider this only in a very limited context, not across many transitions in a supply chain or for a longer period.

The IPTC Photo Metadata Conference on 26 May 2016 in Zagreb (Croatia, Europe) will address how to avoid losing information when images are moved from one person or system to the next one or if they are kept in an archive for a long term.

All interested parties are welcome to participate: free-lance photographers, small and large picture agencies and **libraries**, and trade associations from the photo business.

IPTC News https://iptc.org/news/keep-metadata-of-photos-alive-and-intact/

April 7 : Google Grant Supports Developing EXTRA Classification System

The International Press Telecommunications Council (IPTC) will use a grant from the first round of Google's Digital News Initiative Innovation Fund to build and freely distribute an initial version of EXTRA: The EXTraction Rules Apparatus, a multilingual open-source platform for rules-based classification of news content.

EXTRA will be a classification system for annotating news documents with high-quality subject tags. Such tags will allow publishers to deliver a variety of valuable services including content recommendations, improved advertising targeting and subject-specific content streams, such as alerts and topic pages.

"By creating a freely available rules-based classification engine, IPTC will help publishers to enhance their content with all sorts of metadata services, including enriched search, intelligent recommendations and precise analytics," said Stuart Myles, chairman of IPTC.

EXTRA will provide news publishers with several key capabilities: the ability to automatically categorize documents by subject (for example, terrorism, sports, names of celebrities); the ability to author classification rule sets tailored to existing taxonomies; and the ability to classify documents using the industry standard IPTC Media Topics taxonomy. Taxonomies are used by many news organizations to classify their content. Classification is used in various ways, including improved online news navigation by grouping and linking, to organize editorial workflows and to enrich search.

IPTC News https://iptc.org/news/2016/04/

The 15th International Semantic Web Conference http://iswc2016.semanticweb.org/
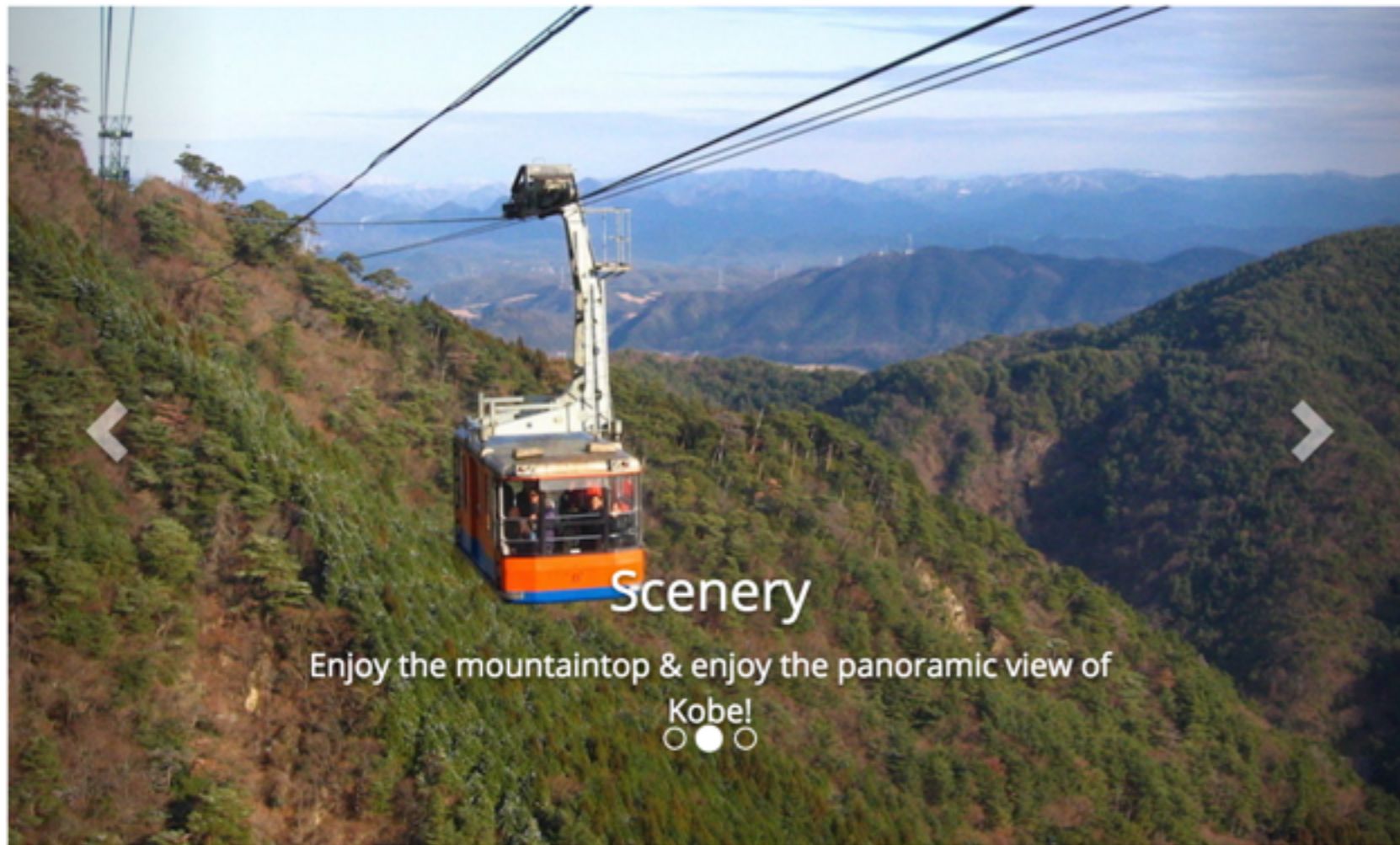
Retroblick

- IPTC schemas (rNews, NITF, NewsML, SportsML, etc), <u>schema.org</u>, and the Open Graph Protocol are **story oriented**.

- IPTC schemas NITF, NewsML-G2, SportsML-G2, Subject Codes, and others facilitate the **exchange** of stories between news publishers.

- rNews (and its predecessor hNews), ninjs, <u>schema.org</u> are designed for **web publishing** and for **web discovery**.

- The Open Graph Protocol OGP enables any object to become part of a **social graph**.

- Library metadata for digitized historical news(papers) is primarily **issue oriented** (some libraries do divide issues into articles).

- Libraries that harvest digital news seem mostly to collect **websites**, not individual articles (exception: National Library of Sweden collects RSS feeds).

References

- Facebook Instant Articles : https://developers.facebook.com/docs/instant-articles
- IPTC Standards : https://iptc.org/standards/
- https://schema.org
- MODS schema and documentations : https://www.loc.gov/standards/mods
- Reuters Institute *Digital News Report 2015*. http://www.digitalnewsreport.org/survey/2015/sources-of-news-2015/
- Biggest Facebook publishers February 2016. NewsWhip. https://www.newswhip.com/2016/03/biggest-facebook-publishers-february-2016
- The Open Graph Protocol. http://ogp.me/
- Roberto Ferdman. *The simple, brilliant plan to change macaroni and cheese that duped us all*. The Washington Post. 23-March-2016. https://goo.gl/dE0PTR
- Evan Sandhaus. rNews is here. And this is what it means. The New York Times. 16-February-2012. http://open.blogs.nytimes.com/2012/02/16/rnews-is-here-and-this-is-what-it-means/
- Atom Syndication Format IETF RFC 4287. https://tools.ietf.org/html/rfc4287
- RSS Specifications etc. http://www.rssboard.org/

Frederick Zarndt
IFLA Governing Board
Digital Divide Data
frederick@frederickzarndt.com
@cowboyMontana