# Increasing access to free and open knowledge for speakers of underserved languages on Wikipedia

• • •

Lucie-Aimée Kaffee
@frimelle

# Find the slides at

http://tinyurl.com/ArticlePlaceholder-FOSSCOMM16

label

# Douglas Adams (Q42)

English writer and humorist

description

Douglas Noël Adams | Douglas Noel Adams

► In more languages

## Statements

| property | | |
|---|---|---|
| spouse | Jane Belson | value |
| | start time | 25 November 1991 |
| claim | end time | 11 May 2001 | qualifiers |
| | ► 1 reference | |
| | | + add (statement) |

rank     collapsed reference
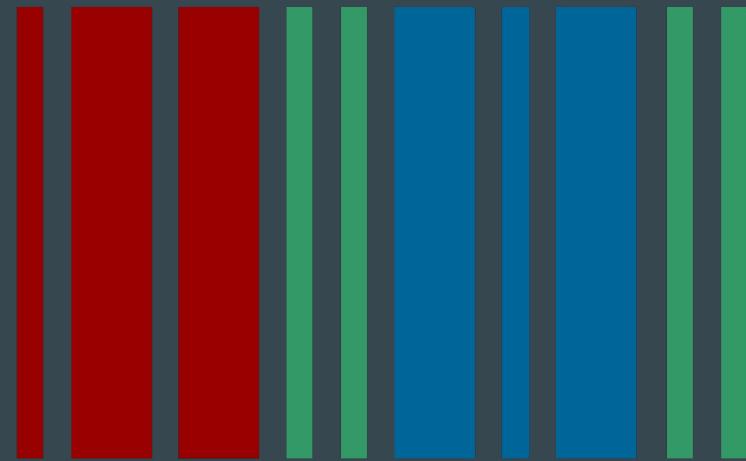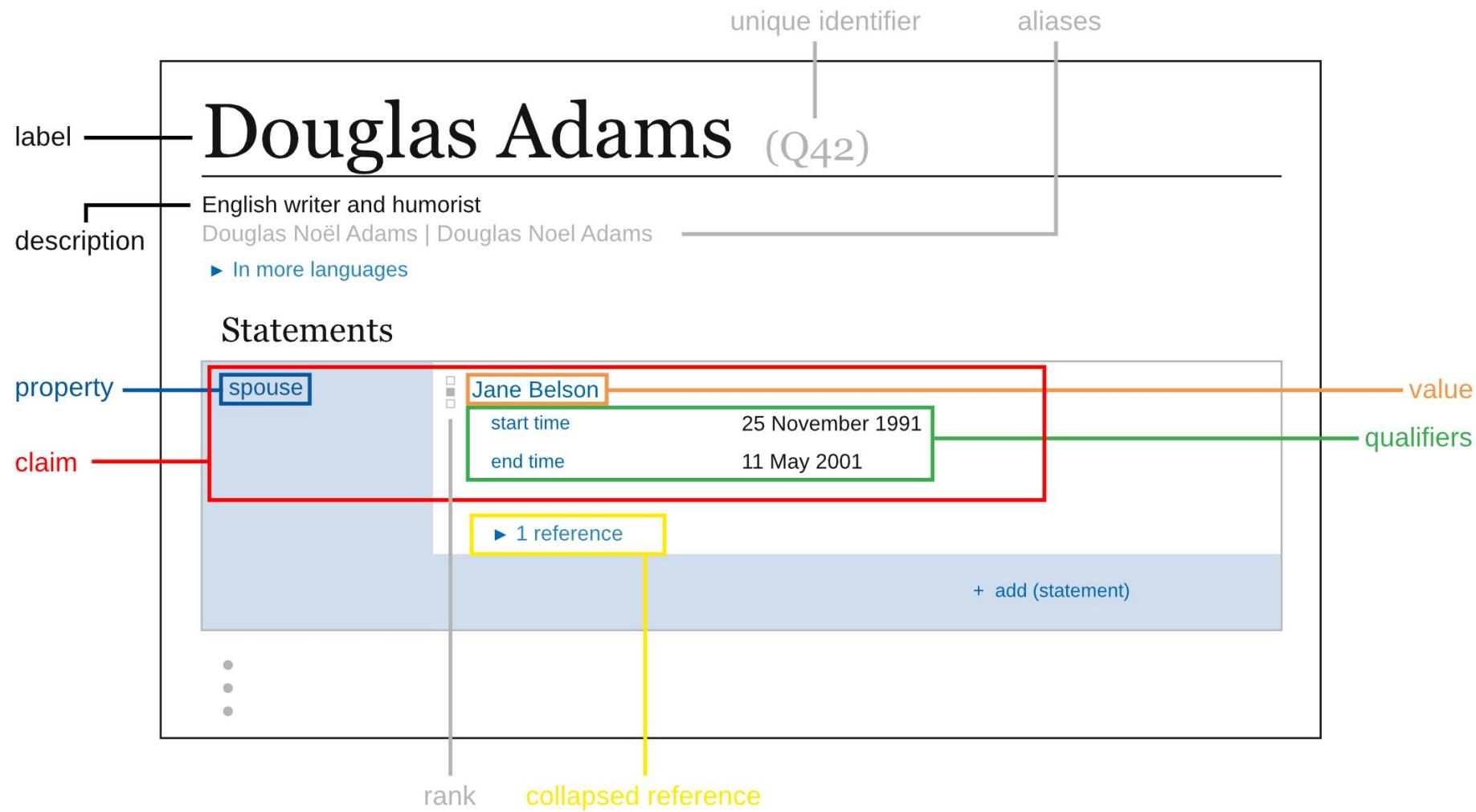
# State of languages online today

# Usage Statistic of Content Languages for Websites

| Language | Percentage |
|---|---|
| English | 53.8% |
| Russian | 6.3% |
| German | 5.7% |
| Japanese | 5.0% |
| Spanish | 4.9% |
| French | 4.1% |
| Portuguese | 2.6% |
| Italian | 2.1% |
| Chinese | 2.0% |
| Polish | 1.9% |
| Turkish | 1.7% |
| Dutch, Flemish | 1.4% |
| Persian | 1.1% |
| Arabic | 0.8% |
| Czech | 0.8% |
| Korean | 0.7% |
| Vietnamese | 0.6% |
| Swedish | 0.5% |
| Indonesian | 0.5% |
| Greek | 0.5% |
| Romanian | 0.4% |
| Hungarian | 0.4% |
| Danish | 0.3% |
| Thai | 0.3% |
| Slovak | 0.3% |
| Finnish | 0.2% |
| Bulgarian | 0.2% |
| Hebrew | 0.2% |

# First-language speakers

| Language | Speakers (millions) |
|---|---|
| Chinese | 1,197 |
| Spanish | 399 |
| English | 335 |
| Hindi | 260 |
| Arabic | 242 |
| Portuguese | 203 |
| Bengali | 189 |
| Russian | 166 |

# On Wikipedia



**English version of Wikipedia gets more than six times as many page views as next largest language version**

*Leading language versions of Wikipedia, by billions of page views in 2015*
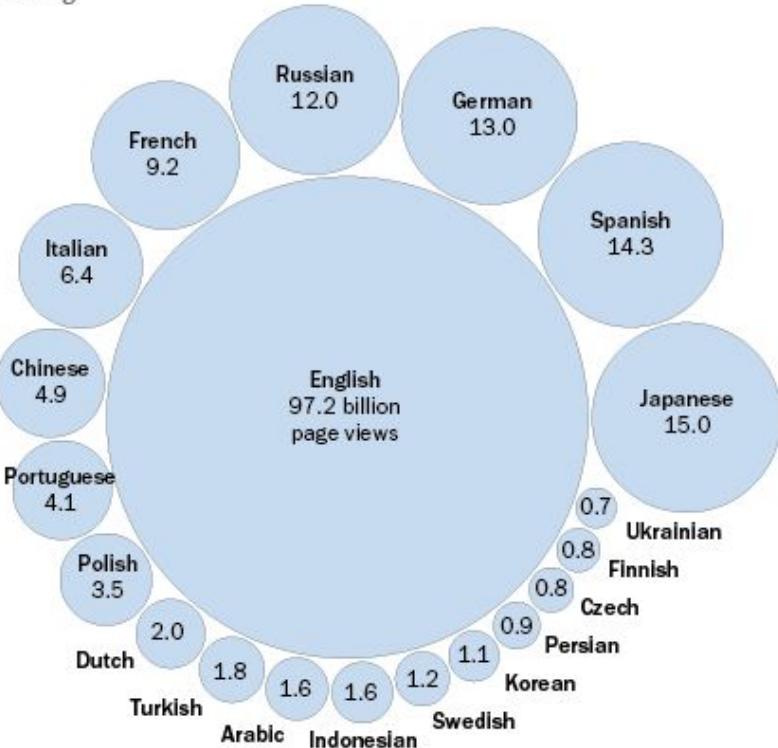
English — 97.2 billion page views

- Japanese 15.0
- Spanish 14.3
- German 13.0
- Russian 12.0
- French 9.2
- Italian 6.4
- Chinese 4.9
- Portuguese 4.1
- Polish 3.5
- Dutch 2.0
- Turkish 1.8
- Arabic 1.6
- Indonesian 1.6
- Swedish 1.2
- Korean 1.1
- Persian 0.9
- Czech 0.8
- Finnish 0.8
- Ukrainian 0.7

| Number of Articles | Number of Wikis |
| --- | --- |
| 5M+ | 1 |
| 1M - 4M | 11 |
| 1K - 1M | 227 |
| 0 - 1000 | 48 |

# Global population covered by number of Wikipedia articles in their primary and secondary language

| | Any number of articles | Over 5000 | Over 20 000 | Over 50 000 | Over 100 000 | Over 1 million |
|---|---|---|---|---|---|---|
| Primary language | 88% | 80% | 74% | 64% | 58% | 20% |
| Primary & secondary language | 90% | 84% | 79% | 71% | 67% | 28% |
| | 280 languages | 158 languages | 104 languages | 74 languages | 55 languages | 12 languages |

■ Primary language   ■ Primary & secondary language

```
                    ┌─────────────┐
                    │ few articles │
                    └─────────────┘
                   ↗               ↘
      ┌─────────────┐           ┌──────────────┐
      │ few editors  │ ←─────────│ little attention │
      └─────────────┘           └──────────────┘
```

ArticlePlaceholder

Search
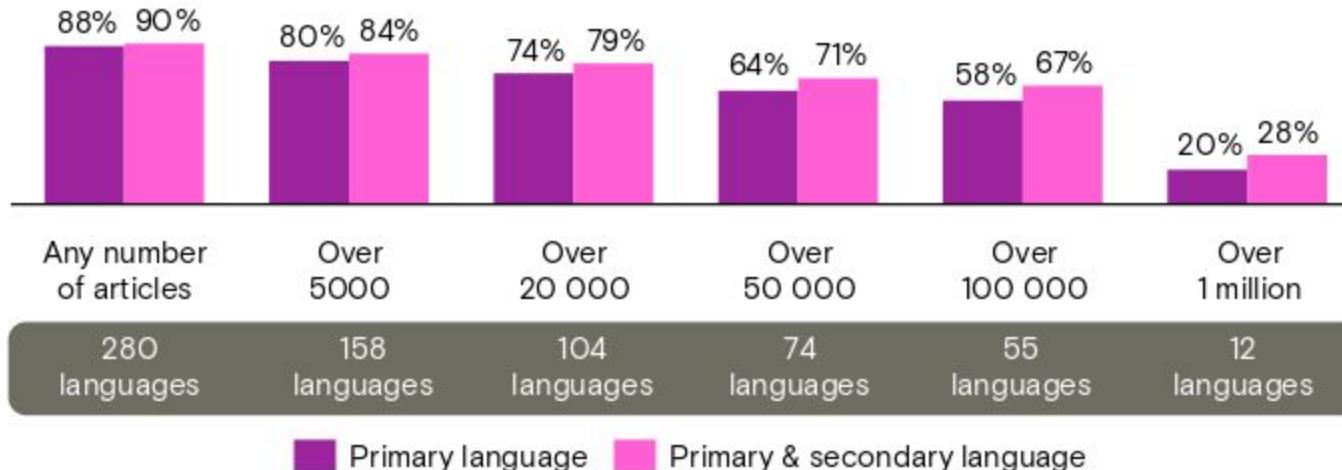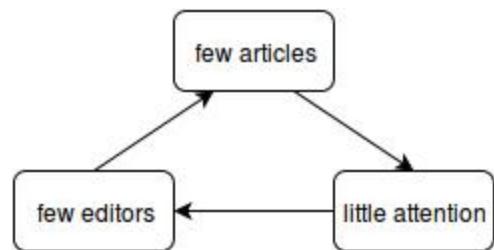
# Ada Lovelace

**Create an article**

English mathematician, considered the first computer programmer

Tools

What links here
Special pages
Printable version
Page information

In other languages
Afrikaans
Alemannisch
العربية
Azərbaycanca
Беларуская
Български
বাংলা
Bosanski
Català
English
Čeština
Cymraeg
Dansk
Deutsch
Ελληνικά
Esperanto
Español
Eesti
Euskara
فارسی
Suomi
Français
Gaeilge
Galego
Gaelg
עברית
Hrvatski
Kreyòl ayisyen
Magyar
Հայերեն
Bahasa Indonesia

| place of death | place of birth | cause of death |
|---|---|---|
| Marylebone[18] | London[19] | uterine cancer[20] |

| occupation | sex or gender | father |
|---|---|---|
| mathematician[21] programmer poet computer scientist engineer | female[22][23][24][25] | Lord Byron |

| birth name | field of work | languages spoken or written |
|---|---|---|
| Augusta Ada Byron[26] | computer science | English[27][28][29] |

## Identifier

National Portrait Gallery (London) person identifier

mp50509

Open Library identifier

OL776398A

## mother

Anne Isabella Byron

## Reference

1. ↑ **imported from**: French Wikipedia
2. ↑ **imported from**: German Wikipedia
3. ↑ **imported from**: French Wikipedia
4. ↑ **imported from**: International Standard Name Identifier
5. ↑ **imported from**: German Wikipedia
6. ↑ **imported from**: German Wikipedia
7. ↑ **imported from**: English Wikisource
8. ↑ **imported from**: Historical Commission of the Bavarian Academy of Sciences
9. ↑ **stated in**: Freebase Data Dumps
10. ↑ **publication date**: 28 October 2013
11. ↑ **stated in**: Virtual International Authority File

# Implementation

## SpecialAboutTopic

- idParser : EntityIdParser
- termLookupFactory : LanguageFallbackLabelDescriptionLookupFactory
- sitelinkLookup : SitelinkLookup
- siteStore : SiteStore
- titleFactory : TitleFactory
- siteGlobalId : String
- entityLookup : EntityLookup

---

+ construct( EntityIdParser, LanguageFallbackLabelDescriptionLookupFactory, SiteLinkLookup, SiteStore, TitleFactory, siteGlobalId, EntityLookup )
+ execute(  String)
- showContent( String )
+ getTextIdParam( String, String )
- showPlaceholder( ItemId )
- showCreateArticle( String )
- getLabel( IdemId )
- showTitle( String )
- showLanguageLinks( ItemId )
- getArticleOnWiki( ItemId )
# geteGroupName()
# createForm
# getFormElements

## SpecialPage

## Hooks

+ onScribuntoExternalLibraries( String, Array )
+ registerScribuntoExternalLibraryPaths( String, Array)
+ onUnitTestsList( Array )

## SearchHookHandler

- termIndex : TermIndex
- termSearchInteractor : TermIndexSearchInteractor
- languageCode : String

---

+ __construct( TermIndex, TermIndexSearchInteractor, String )
+ onSpecialSearchResultsAppend( SpecialSearch, OutputPage, String|null )
+ addToSearch( SpecialSearch, OutputPage, String )
- newFromGlobalState( String )
-  getSearchHeader()
- getSearchResult( String )
- createResult( TermSearchResult, String )
- searchEntities( String )

## Scribunto_LuaArticlePlaceholderLibrary

+ getImageProperty()
+ register()

## Scribunto_LuaLibraryBase

Marie Curie

# Search results

Marie Curie    **Search**

Content pages    Multimedia    Everything    Advanced

**Create the page "Marie Curie" on this wiki!**

There were no results matching the query.

marie curie

# Search results

marie curie    **Search**

Result **1** of **1**

Content pages   Multimedia   Everything   Advanced

**Create the page "Marie curie" on this wiki!**

## Discover data on the topic

**Marie Curie:** French-Polish physicist and chemist
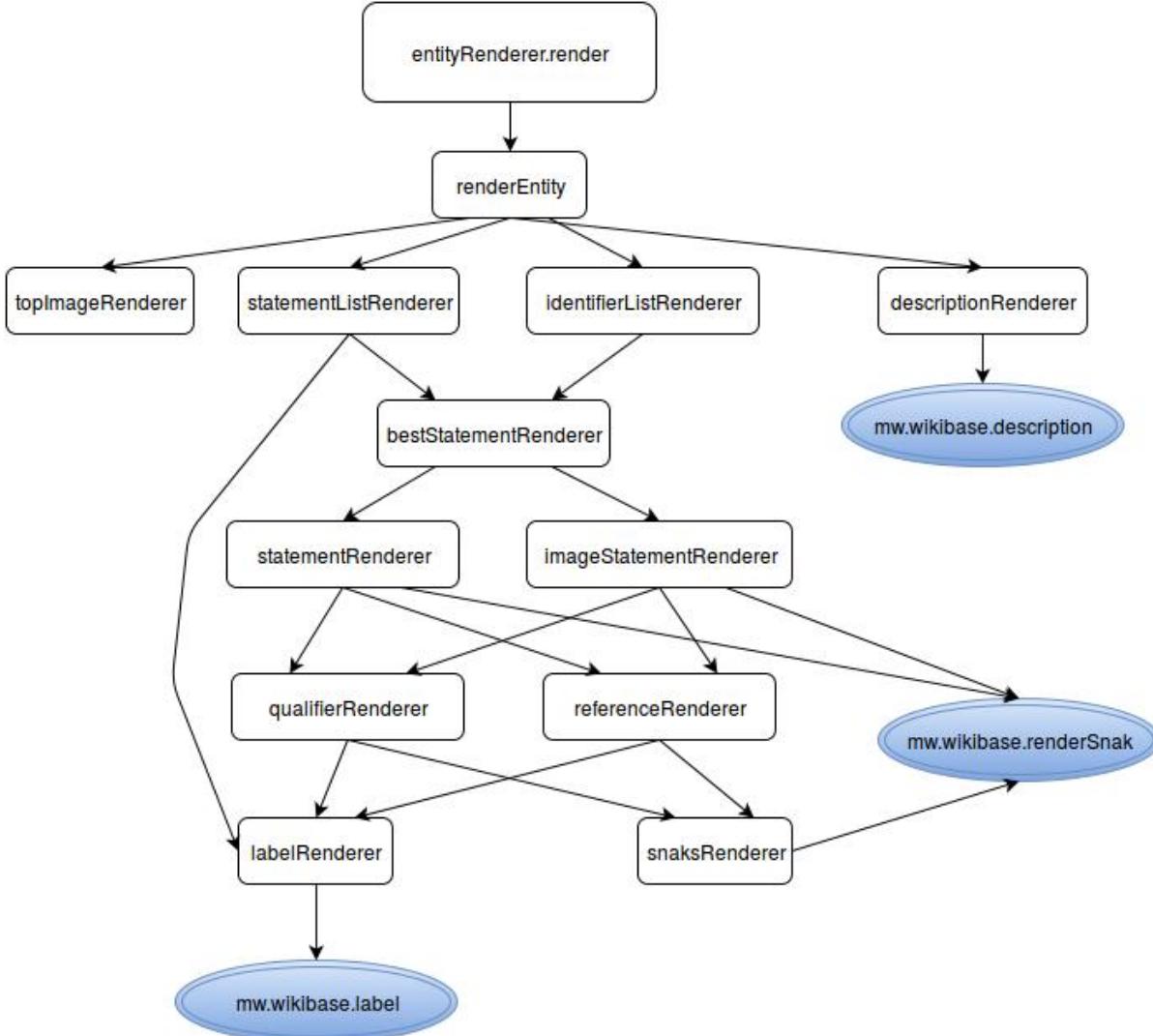
# Deployment

Deployed to the Beta Cluster

Deployed to Test Wikipedia

Upcoming:

Beta Feature

Enable by default on a Wikipedia

More Wikipedias

Sister projects

____

# Community interest

Among the first Wikipedias with interest in the extension:

Odia Wikipedia

Esperanto Wikipedia

Haitian Creole Wikipedia

——

# How can you contribute?

Besides editing and translating data on wikidata.org

➢ Translate the extension itself on translatewiki.net
➢ Translate the most common labels https://tools.wmflabs.org/wikidata-terminator/
➢ Contribute to the code (PHP, Lua, JavaScript, CSS)
➢ Test the Lua modules on-wiki
➢ Help out with documentation/tests
➢ Pick a task on Phabricator https://phabricator.wikimedia.org/project/view/1416/

# Contact

## Wikidata

#wikidata on Freenode

@wikidata (Twitter)

wikidata@lists.wikimedia.org

www.wikidata.org/wiki/Wikidata:Project_chat

## Lucie

frimelle on #wikidata (freenode)

@frimelle (Twitter)

lucie.kaffee@wikimedia.de

[[User:Frimelle]]

Thanks to GFOSS

# References

Slide 3: Wikidata logo, [[User:Planemad]], public domain.

Slide 4: WikidataTermDescriptorDiagram, Charlie Kritschmar [[User:Incabell]], CC-BY-SA.

Slide 5: MediaWiki logo, [[User:Isarra]], [[User:Anthere]], [[User:Eloquence]] and [[User:Aka]], CC-BY-SA.

Slide 7: Usage of content languages for websites, http://w3techs.com/technologies/overview/content_language/all (last visited Jan. 29, 2016).

Slide 8:  Languages with at least 50 million first-language speakers, https://www.ethnologue.com/statistics/size (last visited Jan. 29, 2016).

Slide 11: English version of Wikipedia gets more than six times as many page views as next largest language version, http://www.pewresearch.org/fact-tank/2016/01/14/wikipedia-at-15/ft_16-01-13_wikipedia_bubble/ (last visited April 15, 2016).

Slide 12: State of connectivity 2015, https://fbinternetorg.files.wordpress.com/2016/02/state-of-connectivity-2015-2016-02-21-final.pdf (last visited April 15, 2016).

Slide 15: MagicUnicorn, Logo of the ArticlePlaceholder extension, Charlie Kritschmar [[User:Incabell]], CC-BY-SA.