



IPTC Photo Metadata Crawler

Brendan Quinn, Managing Director, IPTC

IPTC Photo Metadata Conference 2019

Paris 6 June 2019



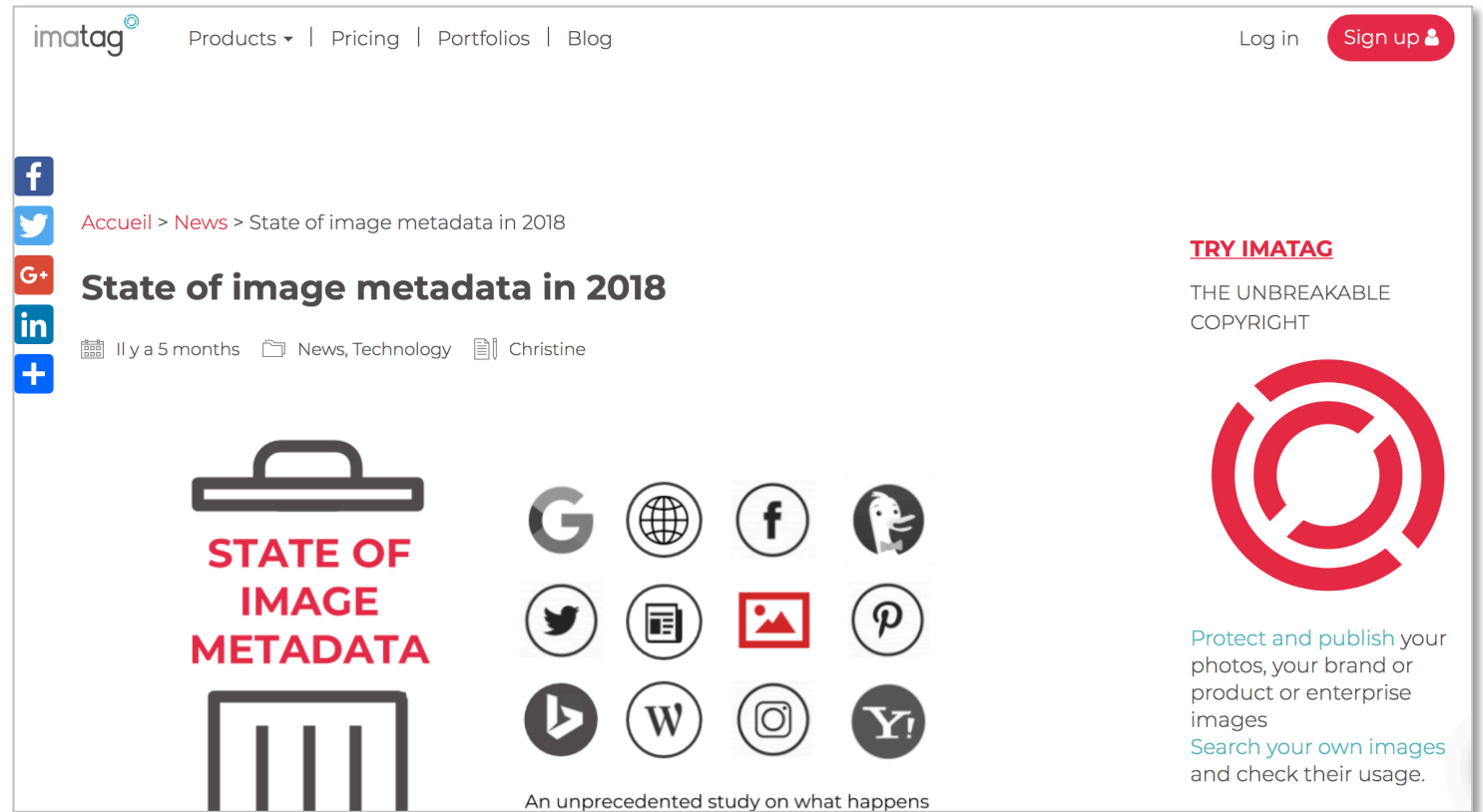
The problem

- We know that some sites are better at implementing IPTC Photo Metadata than others, but...
- Who can we point to as examples of best practice?
- Who should we remind to improve their metadata? (particularly in the news and media space where we are focussed)
- How can we compare sites' usage of IPTC Photo Metadata?
 - who uses which fields?
 - who is putting what text in their fields?
- How can we track sites' usage of IPTC Photo Metadata over time?



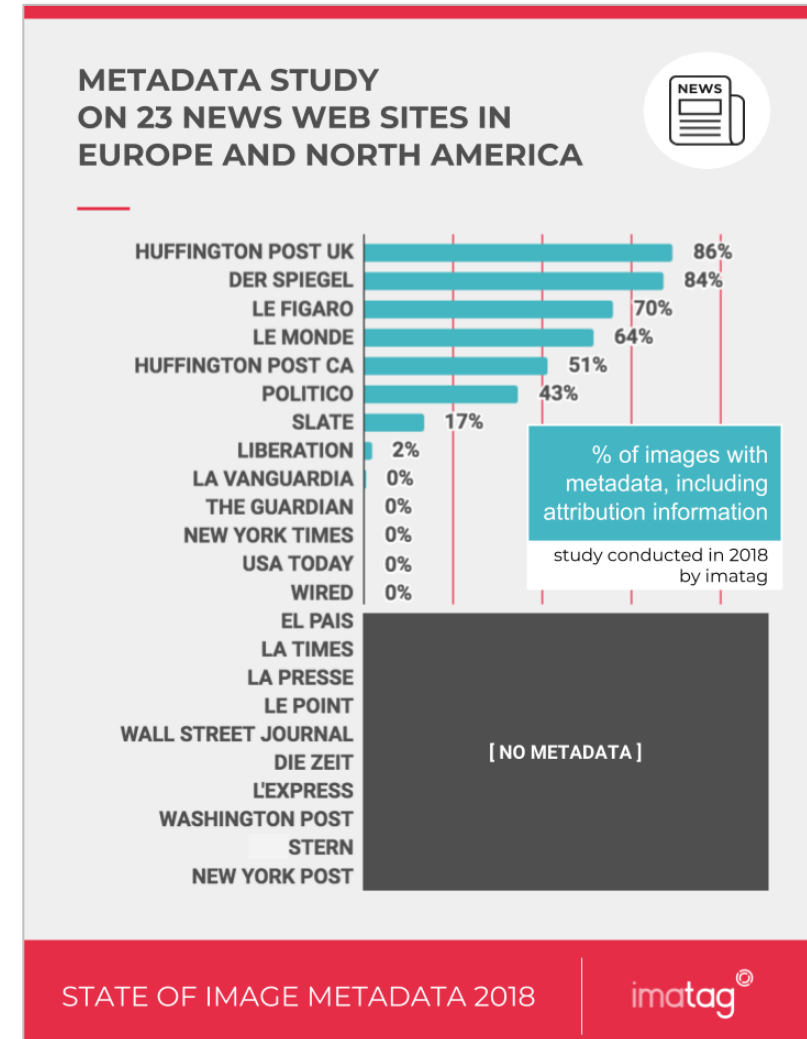
The inspiration

- At the IPTC Photo Metadata Conference in May 2018, Imatag presented their survey of embedded image metadata
- Their crawler identified that 85% of images published on the web have no metadata, and only 3% have copyright information
 - Based on “a sample of over 40 million images, not including social media or image databases”
- The figures become 80% and 8% on “editorial websites”
 - “More than 50,000 images have been analyzed by IMATAG on over 750 different editorial websites worldwide.”



Charts with specific news providers are very useful

- Showing usage of metadata across news providers:
 - Those that use metadata well, including description and rights information
 - Those that use other metadata (EXIF)
 - Those that use no metadata at all
- This might be for a variety of reasons...
 - Not receiving images with metadata from agencies or photographers
 - Deliberate stripping of metadata for privacy reasons, or to protect sources
 - CMS, DAM, or CDN tools may be stripping out metadata inadvertently or deliberately





IMATAG's search can also show embedded metadata


imatag® Products ▾ Pricing | Portfolios | Blog Log in Sign up Et


- <http://www.imatag.com/search/>


getty images

All sizes English All colors

2270703 results found
in 1.908 seconds

 **Marijuana and criminality**
<https://hightimes.com/news/native-american-trib..>
3840 × 2400 - 10/06/2018
Gavel and marijuana.
Getty Images

 **Mature and young couple holding Christmas decorations**
<http://cwwz.net/image/2281400442/>
550 × 362 - 10/04/2018
ab72389.
Getty Images

 **America's Cup - Finals Races 9 & 10**
<http://toursmaps.com/best-summer-destinations-uk/>
617 × 467 - 08/15/2018
jpg?w=300
Getty Images




Image metadata

File properties

Exif

Iptc

Application2

Byline: Emilie Duchesne

BylineTitle: Contributor

Caption: Gavel and marijuana. Concept art

Copyright: Emilie Duchesne

Credit: Getty Images

DateCreated: 2010-06-19

Headline: Marijuana and criminality

Keywords: Legal System,Gavel,Addiction,

ObjectName: 157618833

RecordVersion: 0

Source: E+

SpecialInstructions: Not Released (NR)

Envelope

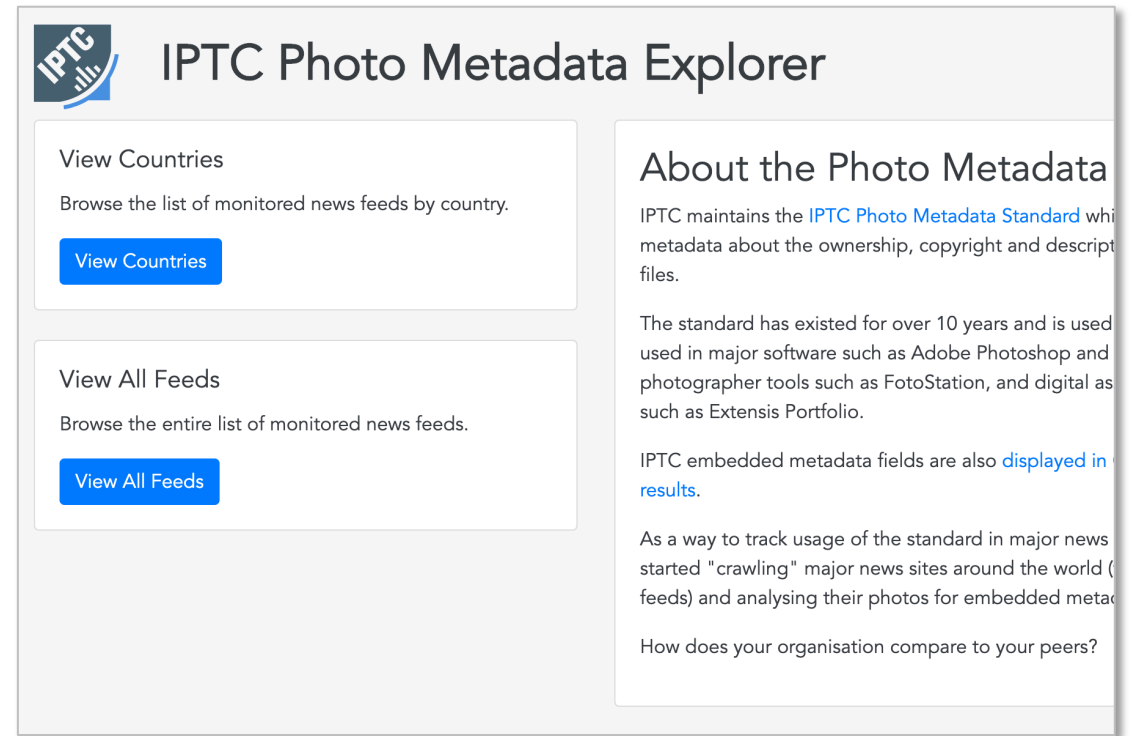
CharacterSet: %G

Xmp



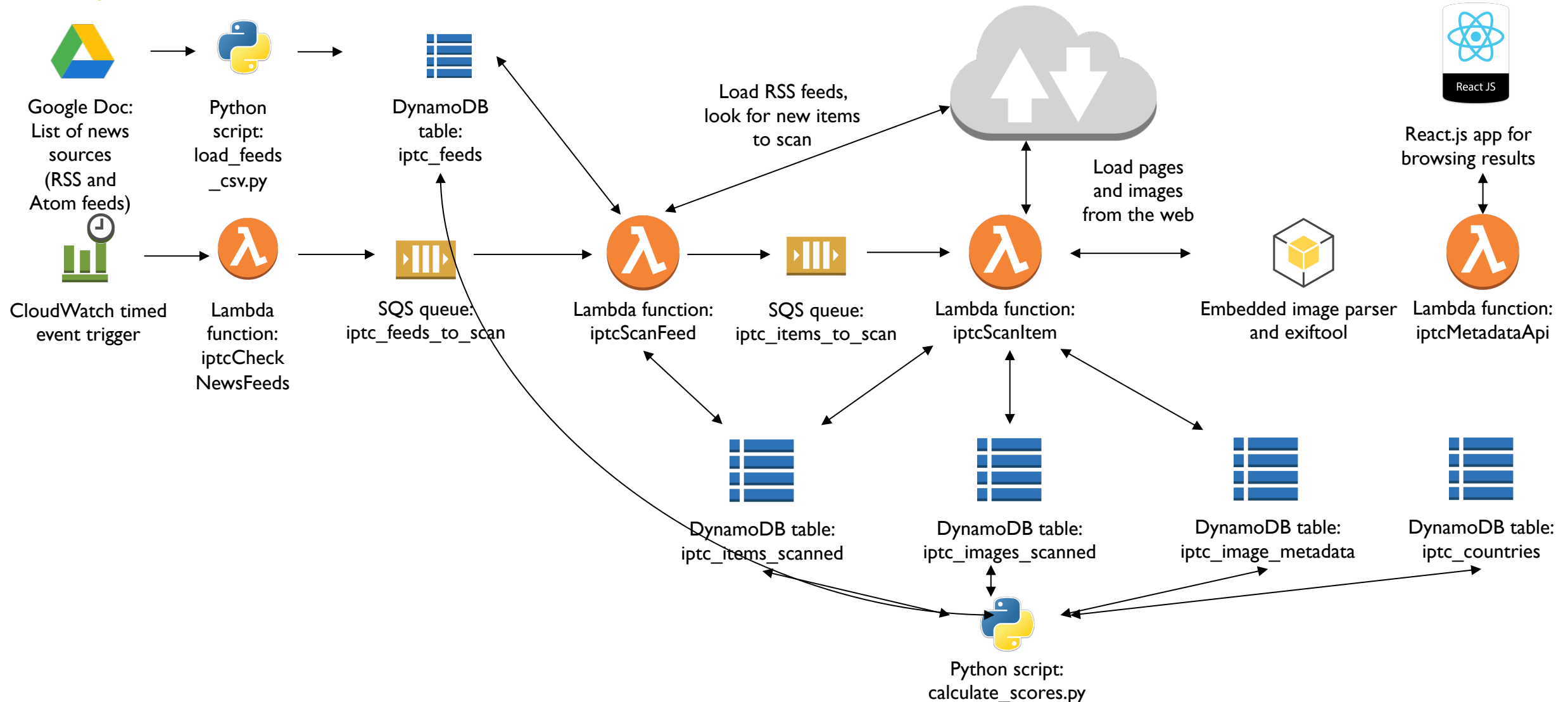
This was inspiration for our own crawler

- We have the photo metadata checker tool (pictured) but that only works on one image or one page at a time
- We thought it would be useful to have something that tracks IPTC Photo Metadata usage on news sites
- So why not build something simple and lightweight that can track image metadata usage?
- We can run it as often as we like to track trends over time
- Lets us analyse usage, find best practices and people who could be encouraged to follow the spec better
- Based on news publishers' public RSS or Atom feeds





IPTC Photo Metadata Crawler Architecture





Checking 200 news feeds

Country	Number of feeds checked	Examples
AT	8	DerStandard, ORF
CA	4	CBC, CTV, Radio Canada, 500px...
CH	2	Neue Zürcher Zeitung, Blick
DE	18	Der Spiegel (ENG/DE), Bild, DW, Die Zeit...
ES	6	El Pais, El Mundo, ABC...
FR	8	Le Monde, Le Figaro, France24...
GB	21	BBC, Times, Daily Mail, Metro, Economist...
IN	5	Times of India, Hindustan Times, The Hindu...
IT	6	ANSA, RAI News
JP	17	Japan Times, Nikkei Shinbun, Yomiuri Shinbun
US	9	CNN, NYT, USA Today, observer.com

Every EU country has at least three feeds

Feed ID	Last scanner	Country	Site Name	Site URL	Feed URL
73e1af0f-524a-4048-a059-ebc657d14bb2	29/08/2018	CA	500px ISO	https://iso.500px.com/	https://iso.500px.com/feed/
ae46dc15-9e52-4e27-bc8b-42e7df771d8	30/08/2018	DE	Die Welt	http://www.welt.de/	https://www.welt.de/feeds/latest.rss
d1645dec-6c14-4d21-af35-5fee67bc253d	30/08/2018	DE	Süddeutsche Zeitung	http://www.sueddeutsche.de/	http://rss.sueddeutsche.de/rss/Topthemen
4324ac5e-3a37-4a72-bc76-d3fed7512b9c	30/08/2018	DE	Frankfurter Rundschau	http://www.fr.de/	http://www.fr.de/?_XML=rss
f68ed192-6d92-4313-ae5e-e03eb721203b	30/08/2018	DE	Bild	https://www.bild.de/	https://www.bild.de/rss/feeds/vw-home/vw-home-16725562.dzt
99f09b8c-b8d1-4283-9307-a245a81ddabf	30/08/2018	DE	Deutsche Welle	http://www.dw-world.de/	http://rss.dw.com/atom/rss-en-all
20cf1969-86ee-4d7d-9d88-9404ca4cd290	30/08/2018	DE	Die Zeit	https://www.zeit.de/	http://newsfeed.zeit.de/all
6005073c-62f6-4dc4-8a77-6c03e195b39a	30/08/2018	DE	Spiegel Online International	https://www.spiegel.de/	http://www.spiegel.de/international/index.rss
9aabb1586-abfa-4a0c-8c2e-bb88be50d7a1	30/08/2018	DE	taz	http://www.taz.de/	http://www.taz.de/Themen-des-Tages/p15/rss/
3cd4885b-d6a9-48a5-915d-b6f51d239f1a	30/08/2018	DE	ZDF Heute	https://www.zdf.de/nachrichten/	https://www.zdf.de/rss/zdf/nachrichten
7da512e4-144e-49c7-80ea-17049390a17b	30/08/2018	DE	FOCUS Online	http://www.focus.de/	https://rss.focus.de/
cdb67b58-7350-445f-aed1-df4df2a7883e	30/08/2018	DE	The Local	http://www.thelocal.de/	https://www.thelocal.de/feeds/rss.php
1503c688-bb0c-463e-9042-d90d2f15c4c6	30/08/2018	DE	stern.de	http://www.stern.de/	https://www.stern.de/feed/standard/all/
584c1964-efce-4667-aae6-0d1c48f73716	30/08/2018	DE	ITB Berlin	https://itb-berlin-news.com/	https://itb-berlin-news.com/feed/
15717713-0ab9-4d82-a488-72937349595	30/08/2018	DE	Munich NOW	http://munichnow.com/	http://munichnow.com/feed/
27689d94-aeef-4ed1-b547-f4ee277657c6	30/08/2018	DE	Frankfurter Allgemeine	http://www.faz.net/	http://www.faz.net/rss/aktuell/
954d6212-e37c-424d-a0b0-728186f7c247	30/08/2018	DE	taqesschau.de	https://www.taqesschau.de/	https://www.taqesschau.de/vml/rss2
569207fe-420b-4f18-9ef0-2262dd70951d	30/08/2018	FR	20minutes	https://www.20minutes.fr/	https://www.20minutes.fr/rss/
9a09325a-7002-4106-b90d-903c7ce09049	30/08/2018	FR	L'Internaute	https://www.linternaute.com/	https://www.linternaute.com/rss/
c3645b29-384f-4b81-abf0-5ccc2c486951	30/08/2018	FR	Euronews	http://www.euronews.com/	http://feeds.feedburner.com/euronews/en/home/
e27451cc-48cb-4fa3-a679-ac682c945143	30/08/2018	FR	Le Monde	https://www.lemonde.fr/	https://www.lemonde.fr/rss/une.xml
1e58dba0-89b5-4ccf-9dc5-77c56c09c19d	30/08/2018	FR	Le Figaro	http://www.lefigaro.fr/	http://www.lefigaro.fr/rss/figaro_actuallites.xml
72c4aee9-577f-4391-9e66-085edd0ab535	30/08/2018	FR	France24	https://www.france24.com/fr/	https://www.france24.com/fr/france/rss
a28e4c8f-d40a-422e-b5f3-b6cc7b3aa691	29/08/2018	GB	The Daily Mirror / Sunday Mirror	https://www.mirror.co.uk/	https://www.mirror.co.uk/?service=rss
c0f333b8-4eae-4490-a4fc-d5ed7913e3fd	29/08/2018	GB	Spectator	https://www.spectator.co.uk/	https://www.spectator.co.uk/feed/
738ee54a-e26a-4a51-a980-e416cc1fc20d	29/08/2018	GB	Daily Mail / Mail on Sunday	http://www.dailymail.co.uk/	http://www.dailymail.co.uk/home/index.rss
f42a0c0f-a6a4-4058-9d14-b0b0ff0427f1	29/08/2018	GB	Evening Standard	https://www.standard.co.uk/	https://www.standard.co.uk/rss
9c8f1a2f-7207-417f-8ea4-b0fa91d0b611	29/08/2018	GB	The Independent	http://www.independent.co.uk/	https://www.independent.co.uk/rss
f17eb337-737c-449c-8f1d-9595c03ba9c5	29/08/2018	GB	The Sun	https://wwwthesun.co.uk/	https://wwwthesun.co.uk/feed/
ca201833-7aaa-4ac2-ab1e-14162030f24d	29/08/2018	GB	Channel 4 News	https://www.channel4.com/news/	https://www.channel4.com/news/feed



A RESTful API to the data lets us build an interactive Photo Metadata Explorer tool

```
https://2kpne6m75m.execute-api.us-east-1.amazonaws.com/dev/feeds/e4fd5c37-22a4-4944-b59...

[
  {
    item_rights_licensing_score: 0,
    item_category_score: 0,
    item_score: 0,
    item_iim_score: 0,
    image_count: "7",
    feed_id: "e4fd5c37-22a4-4944-b59",
    item_descriptive_score: 0,
    item_things_shown_score: 0,
    item_xmp_score: 0,
    feed_url: "http://rss.gazeta.pl/",
    item_title: "Skala strajku może",
    scanned_timestamp: 1554666495,
    id: "28184de9-82f0-45f8-8e26-f71",
    item_url: "http://wyborcza.pl/7",
    scanned_date: "2019-04-07"
  },
  {
    item_rights_licensing_score: 0,
    item_category_score: 0,
    item_score: 0,
    item_iim_score: 0,
    image_count: "8",
    feed_id: "e4fd5c37-22a4-4944-b59",
    item_descriptive_score: 0,
    item_things_shown_score: 0,
    item_xmp_score: 0,
    feed_url: "http://rss.gazeta.pl/",
    item_title: "Eurowybory. Kaczyński",
    scanned_timestamp: 1554660451,
    id: "7c0d7850-ba5e-4a01-b7db-0bd",
    item_url: "http://wyborcza.pl/7",
    scanned_date: "2019-04-07"
  },
  {
    item_rights_licensing_score: 0,
    item_category_score: 0,
    item_score: 0,
    item_iim_score: 0
  }
]
```



IPTC Photo Metadata Explorer

View Countries

Browse the list of monitored news feeds by country.

[View Countries](#)

View All Feeds

Browse the entire list of monitored news feeds.

[View All Feeds](#)

About the Photo Metadata Explorer

IPTC maintains metadata files.

The standard used in many photographs such as EXIF.

IPTC embeds results.

As a way to started "cr" feeds) and

How does



IPTC Photo Metadata Explorer

[Home](#) / [Countries](#)

All Countries

Country name	Country code	Number of feeds	IPTC Descriptive	IPTC Things Shown	IPTC Rights & Lice
Australia	AU	9	4.2%	1.9%	3.9%
Austria	AT	8	1.6%	0.1%	1.5%
Belgium	BE	3	0.4%	0.0%	0.0%
Bulgaria	BG	3	0.0%	0.0%	0.0%
Canada	CA	4	4.7%	0.1%	3.8%
Croatia	HR	4	0.0%	0.0%	0.0%
Cyprus	CY	3	4.4%	2.8%	3.4%
Czech Republic	CZ	6	0.0%	0.0%	0.0%
Denmark	DK	4	0.6%	0.5%	0.5%
Estonia	EE	6	0.3%	0.0%	0.3%
Finland	FI	3	0.0%	0.0%	0.0%



Feeds by country (sortable)



IPTC Photo Metadata Explorer

[Home](#) / [Countries](#) / DE

Feeds in DE

Feed name	country	Scanned date	Item Count	IPTC Descriptive	IPTC Things Shown	IPTC Rights & Licensing
taz	DE	24 May 2019	100	87.9%	75.4%	87.9%
Spiegel Online (German)	DE	24 May 2019	100	82.1%	53.7%	80.0%
Spiegel Online International	DE	24 May 2019	86	56.6%	21.2%	50.6%
The Local	DE	24 May 2019	100	11.4%	11.3%	11.3%
Munich NOW	DE	24 May 2019	13	3.1%	1.7%	3.1%
Frankfurter Allgemeine	DE	24 May 2019	100	1.5%	1.0%	1.5%
tagesschau.de	DE	24 May 2019	100	1.0%	0.8%	1.0%
Frankfurter Rundschau	DE	24 May 2019	0	0.0%	0.0%	0.0%
Bild	DE	24 May 2019	100	0.0%	0.0%	0.0%
stern.de	DE	24 May 2019	100	0.0%	0.0%	0.0%
ZDF Heute	DE	24 May 2019	95	0.0%	0.0%	0.0%



Browsing to individual images



IPTC Photo Metadata Explorer

[Home](#) / [Feeds](#) / taz

Feed: taz

(Feed percentage: 87.9%)

Articles in this feed:

Article title	Scanned date	Image count	IPTC Descriptive	IPTC Things Shown	IPTC Rights & Licensing	IPTC Metadata
Finanzen der extremen Rechten: Die Anti-Europäer	2019-05-09	1	100.0%	0.0%	100.0%	100.0%
Sudan nach der Revolution: Mit Geduld in die Freiheit	2019-05-09	1	100.0%	100.0%	100.0%	100.0%
Atomabkommen mit Iran: Teheran platzt der Kragen	2019-05-09	2	50.0%	50.0%	50.0%	50.0%
Die Westdeutschen und die Systemfrage: Kampfbegriff Sozialismus	2019-05-09	1	100.0%	100.0%	100.0%	100.0%
Datenschutz bei Gentests: Gefährliches Wissen	2019-05-	1	100.0%	100.0%	100.0%	100.0%



Browsing to individual images



IPTC Photo Metadata Explorer

[Home](#) / [Feeds](#) / [taz](#)

Feed: taz

(Feed percentage: 87.9%)

Articles in this feed:

Article title

[Finanzen der extremen Rechten: Die Anti-](#)

[Sudan nach der Revolution: Mit Geduld in](#)

[Atomabkommen mit Iran: Teheran platzt d](#)

[Die Westdeutschen und die Systemfrage: Sozialismus](#)

[Datenschutz bei Gentests: Gefährliches W](#)



IPTC Photo Metadata Explorer



[Feeds](#) / [taz](#) / [Atomabkommen mit Iran: Teheran platzt der Kragen](#)

Article: Atomabkommen mit Iran: Teheran platzt der Kragen

Scanned on 9 May 2019

Original URL: <http://www.taz.de/Atomabkommen-mit-Iran/!5589921/>

List of images in this article:

Image (click for details)	Scanned date	IPTC Descriptive Count	IPTC Things Shown Count	IPTC Rights & Licensing	IPTC Metadata?
	2019-05-09	10	6	4	true
	2019-05-09	0	0	0	false



Browsing to individual images



IPTC Photo Metadata Explorer

[Home](#) / [Feeds](#) / [taz](#)

Feed: taz

(Feed percentage: 87.9%)

Articles in this feed:

Article title

[Finanzen der extremen Rechten: Die Anti-](#)

[Sudan nach der Revolution: Mit Geduld in](#)

[Atomabkommen mit Iran: Teheran platzt d](#)

[Die Westdeutschen und die Systemfrage: Sozialismus](#)

[Datenschutz bei Gentests: Gefährliches W](#)



IPTC Photo Metadata Explorer


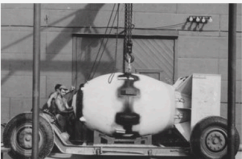
[Feeds](#) / [taz](#) / [Atomabkommen mit Iran: Teheran platzt der Kragen](#)

Article: Atomabkomm

Scanned on 9 May 2019

Original URL: <http://www.taz.de/Atomabkommen-mit>

List of images in this article:

Image (click for details)	Scanned date	IP
	2019-05-09	10
	2019-05-09	0



IPTC Photo Metadata Explorer

[Feeds](#) / [taz](#) / [Atomabkommen mit Iran: Teheran platzt der Kragen](#) / [Image](#)

Image detail



Scanned on 9 May 2019

Original URL: http://www.taz.de/picture/3427794/624/Iran_AKW_Busshehr_110893492.jpeg

Descriptive Metadata

Field name	Field value
Byline (IIM 2:80)	Abedin Taherkenareh
Caption/Abstract (IIM 2:120)	ARCHIV - 21.08.2010, Iran, Buschehr: Zwei weibliche Sicherheitsbeamte stehen vor dem Atomkraftwerk und beobachten Medienvertreter.Sanktionen;USA;Iran Foto: Abedin Taherkenareh/EPA/dpa +++ dpa-Bildfunk +++
Categories (IIM	



IPTC Photo Metadata is split into groups: Descriptive, Things Shown, Rights & Licensing



IPTC Photo Metadata Explorer

Feeds / taz / Atomabkommen mit Iran: Teheran platz der Kragen / Image

Image detail



Scanned on 9 May 2019

Original URL: http://www.taz.de/picture/3427794/624/Iran_AKW_Buschehr_110893492.jpeg

Descriptive Metadata

Field name	Field value
Byline (IIM 2:80)	Abedin Taherkenareh
Caption/Abstract (IIM 2:120)	ARCHIV - 21.08.2010, Iran, Buschehr: Zwei weibliche Sicherheitsbeamte stehen vor dem Atomkraftwerk und beobachten Medienvertreter.Sanktionen;USA;Iran Foto: Abedin Taherkenareh/EPA/dpa +++ dpa-Bildfunk +++ I
Category (IIM 2:15)	pl
Date Created (IIM 2:55)	2010:08:21
Headline (IIM 2:105)	Atomkraftwerk in Buschehr
Keywords (IIM 2:25)	['Iran', 'USA', 'KKK', 'Atom', 'Buschehr', 'Konflikte', 'kernkraftwerk', 'Krisen', 'Krieg', 'Konflikte', 'atomanlage', 'Politik', 'Sanktionen']
Object Name (IIM 2:05)	110893492
Description (XMP IPTC-Core 6.4)	ARCHIV - 21.08.2010, Iran, Buschehr: Zwei weibliche Sicherheitsbeamte stehen vor dem Atomkraftwerk und beobachten Medienvertreter.Sanktionen;USA;Iran Foto: Abedin Taherkenareh/EPA/dpa +++ dpa-Bildfunk +++ I

Headline (IIM 2:105)	Atomkraftwerk in Buschehr
Keywords (IIM 2:25)	['Iran', 'USA', 'KKK', 'Atom', 'Buschehr', 'Konflikte', 'kernkraftwerk', 'Krisen', 'Krieg', 'Konflikte', 'atomanlage', 'Politik', 'Sanktionen']
Object Name (IIM 2:05)	110893492
Description (XMP IPTC-Core 6.4)	ARCHIV - 21.08.2010, Iran, Buschehr: Zwei weibliche Sicherheitsbeamte stehen vor dem Atomkraftwerk und beobachten Medienvertreter.Sanktionen;USA;Iran Foto: Abedin Taherkenareh/EPA/dpa +++ dpa-Bildfunk +++ I
Headline (XMP IPTC-Core 6.5)	Atomkraftwerk in Buschehr
Title (XMP IPTC-Core 6.16)	110893492

Things Shown Metadata

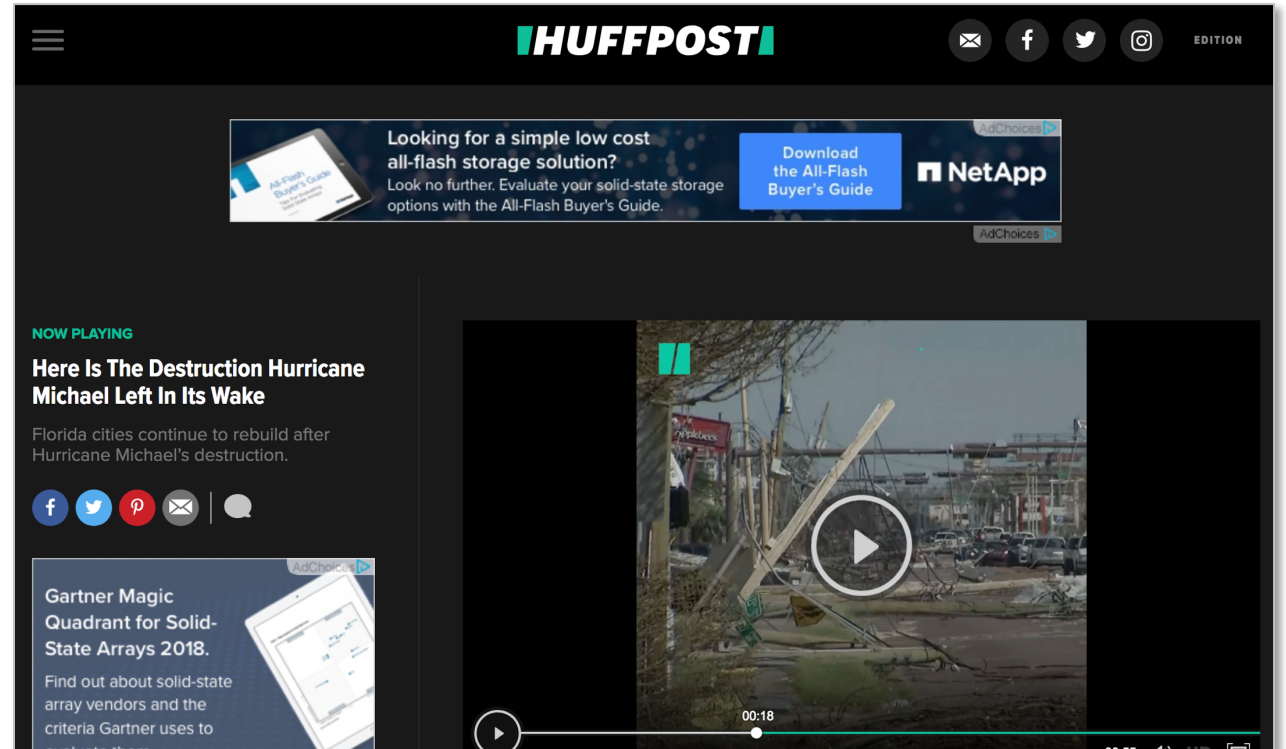
Field name	Field value
City (IIM 2:90)	Buschehr
Country Primary Location Code (IIM 2:100)	IRN
Country Primary Location Name (IIM 2:101)	Iran
City (XMP IPTC-Core 6.1) (legacy)	Buschehr
Country (XMP IPTC-Core 6.2) (legacy)	Iran
Country Code (XMP)	IRN

Rights & Licensing Metadata

Field name	Field value
Copyright Notice (IIM 2:116)	Verfügbar für Kunden mit Rechnungsadresse in Deutschland.
Credit (IIM 2:110)	picture alliance/dpa
Credit Line (XMP IPTC-Core 6.21)	picture alliance/dpa
Source (XMP IPTC-Core 6.23)	EPA

Limitations and exceptions

- Only works for sites with public RSS or Atom feeds
- Won't work if the main image is smaller than 150x150 pixels
- Won't work on embedded images (e.g. posts with only Twitter, Instagram or YouTube embeds)
- Editorial visualisations and infographics rarely have metadata attached (but they should!)
- More and more articles use video rather than images which won't be scanned
- Won't re-read metadata if it is changed after we have scanned a page






How do we calculate scores?

At image level, we count occurrences of metadata tags from both XMP and IIM embedded metadata in 3 categories:

Descriptive	Things shown	Rights & Licensing
<ul style="list-style-type: none">Title (XMP IPTC-Core 6.16), Object Name (IIM 2:05)Headline (XMP IPTC-Core 6.5 / IIM 2:105)Description (XMP IPTC-Core 6.4)Keywords (XMP IPTC-Core 6.7 / IIM 2:25)Caption/Abstract (IIM 2:120)Date Created (XMP IPTC-Core 6.12 / IIM 2:55)Time Created (IIM 2:60)Category (IIM 2:15)Supplemental Categories (IIM 2:20)Creator (IIM 2:80 By-line),Byline Title (IIM 2:85)	<ul style="list-style-type: none">Location (XMP and IIM Location code, city/province/state)Person (XMP PersonInImage, PersonInImageDescription, PersonInImageCVTerm)Organisation (XMP OrganisationInImageCode, OrganisationInImageName)Artwork (XMP ArtworkOrObject, Artwork Circa Date Created, ArtworkCreator,, ArtworkStylePeriod, ArtworkTitle, ...)Product (ProductInImage, ProductInImageGTIN)	<ul style="list-style-type: none">Credit (IIM 2:110) / Credit Line (XMP IPTC-Core 6.21)Source (XMP IPTC-Core 6.23)Copyright Notice (IIM 2:116)CopyrightYearArtworkCopyrightNotice, ArtworkCopyrightOwnerID, ArtworkCopyrightOwnerNameModel release (Model Age, Model Additional Information)Artwork LicensorEmbedded Encoded Rights Expression / Embedded Linked Rights ExpressionRegistry ID / Registry Item ID

Calculating “scores” for pages, feeds and countries


- Individual images:
 - Simply count all occurrences in each category



IPTC Photo Metadata Explorer

[Feeds](#) / [Le Figaro](#) / [L'empereur Naruhito du Japon accède officiellement au trône](#) / Image

Image detail



Scanned on 1 May 2019

Original URL: <https://i.f1g.fr/media/eidos/680x/2019/05/01/XVM45892aac-6bd6-11e9-8e94-ff895166400e-805x453.jpg>

Descriptive Metadata

Field name	Field value
Creator (IIM 2:80 Byline)	POOL
Caption/Abstract (IIM 2:120)	Japan's Emperor Naruhito, flanked by Empress Masako, delivers a speech during a ceremony called Sokui-go-Choken-no-gi, his first audience after the accession to the throne , at the Imperial Palace in Tokyo, May 1, 2019. Japan Pool/Pool via REUTERS JAPAN OUT. NO RESALES. NO ARCHIVES.



Calculating “scores” for pages, feeds and countries

- Individual images:
 - Simply count all occurrences in each category
- Pages (“Items”):
 - The percentage of images containing *any* metadata in that category

IPTC Photo Metadata Explorer

Feeds / Le Figaro / L'empereur Naruhito du Japon accède officiellement au trône

Article: L'empereur Naruhito du Japon accède officiellement au trône

Scanned on 1 May 2019

Original URL: <http://www.lefigaro.fr/international/l-empereur-naruhito-du-japon-accede-officiellement-au-trone-20190501>

List of images in this article:

Image (click for details)	Scanned date	IPTC Descriptive Count	IPTC Things Shown Count	IPTC Rights & Licensing	IPTC Metadata?
	2019-05-01	9	3	1	true
	2019-05-01	8	3	1	true



Calculating “scores” for pages, feeds and countries

- Individual images:
 - Simply count all occurrences in each category
- Pages (“Items”):
 - The percentage of images containing *any* metadata in that category
 - “IPTC Metadata” score is the percentage of images that contain any IPTC metadata fields at all

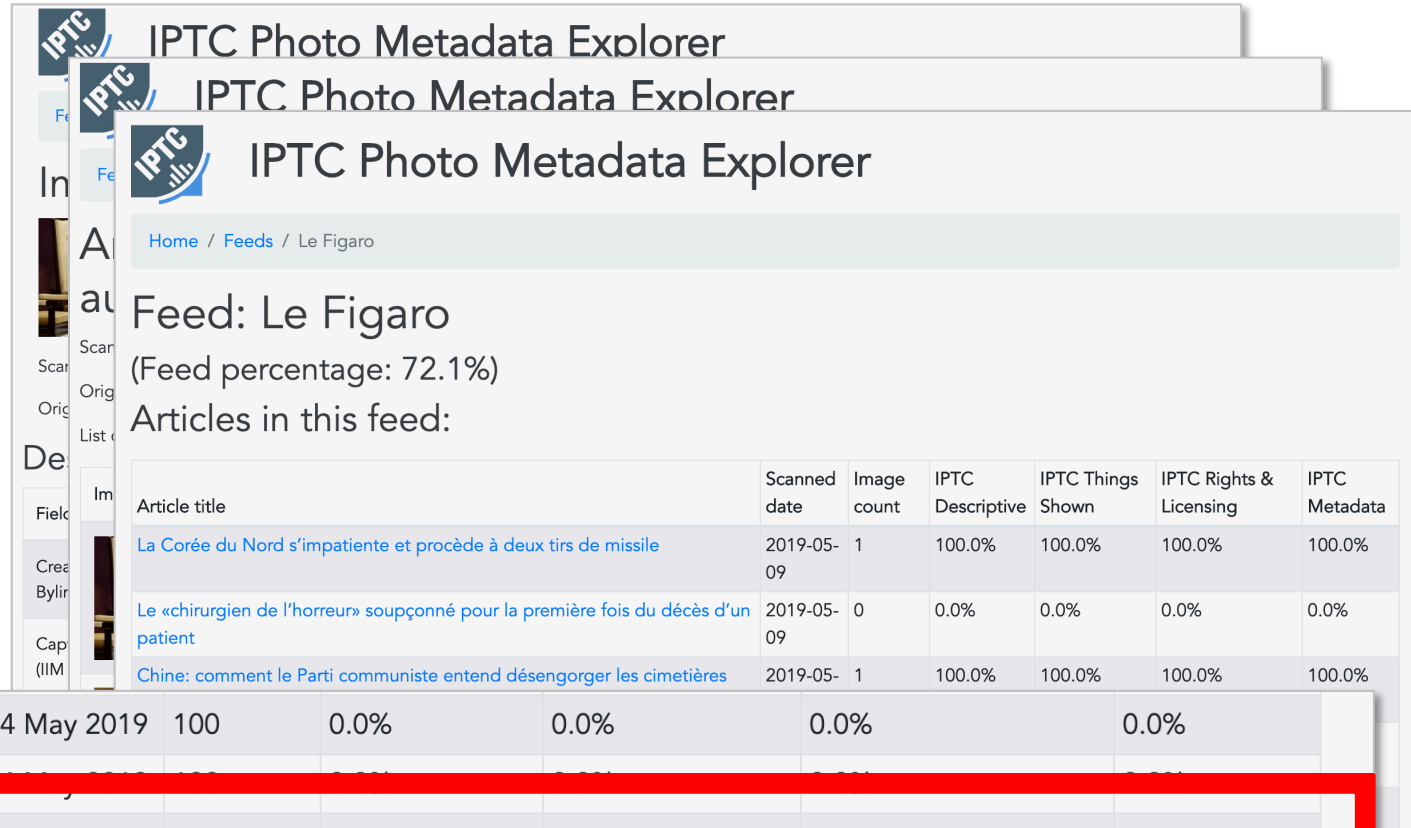
The screenshots show the IPTC Photo Metadata Explorer interface for the 'Feed: Le Figaro'. The breadcrumb trail is 'Home / Feeds / Le Figaro'. The page title is 'Feed: Le Figaro' with a subtitle '(Feed percentage: 72.1%)'. Below this, it says 'Articles in this feed:'. A table lists five articles with their respective IPTC metadata scores.

Article title	Scanned date	Image count	IPTC Descriptive	IPTC Things Shown	IPTC Rights & Licensing	IPTC Metadata
La Corée du Nord s'impatiente et procède à deux tirs de missile	2019-05-09	1	100.0%	100.0%	100.0%	100.0%
Le «chirurgien de l'horreur» soupçonné pour la première fois du décès d'un patient	2019-05-09	0	0.0%	0.0%	0.0%	0.0%
Chine: comment le Parti communiste entend désengorger les cimetières	2019-05-09	1	100.0%	100.0%	100.0%	100.0%
«Unis, nous sommes plus forts»: les 27 s'engagent pour l'avenir de l'Union européenne	2019-05-09	0	0.0%	0.0%	0.0%	0.0%
Entre les fonctionnaires et Macron, le désamour ne date pas de mai 2017	2019-05-09	1	100.0%	100.0%	100.0%	100.0%



Calculating “scores” for pages, feeds and countries

- Individual images:
 - Simply count all occurrences in each category
- Pages (“Items”):
 - The percentage of images containing *any* metadata in that category
 - “IPTC Metadata” score is the percentage of images that contain any IPTC metadata fields at all
- Feeds:
 - Average of percentages on all items in that feed (ignoring items with no images)



The screenshot shows the IPTC Photo Metadata Explorer interface. It displays a feed for 'Le Figaro' with a feed percentage of 72.1%. Below this, a table lists articles in the feed with their respective metadata scores. The table has columns for Article title, Scanned date, Image count, IPTC Descriptive, IPTC Things Shown, IPTC Rights & Licensing, and IPTC Metadata.

Article title	Scanned date	Image count	IPTC Descriptive	IPTC Things Shown	IPTC Rights & Licensing	IPTC Metadata
La Corée du Nord s'impatiente et procède à deux tirs de missile	2019-05-09	1	100.0%	100.0%	100.0%	100.0%
Le «chirurgien de l'horreur» soupçonné pour la première fois du décès d'un patient	2019-05-09	0	0.0%	0.0%	0.0%	0.0%
Chine: comment le Parti communiste entend désengorger les cimetières	2019-05-	1	100.0%	100.0%	100.0%	100.0%

Page	Country	Date	Total Images	IPTC Descriptive	IPTC Things Shown	IPTC Rights & Licensing	IPTC Metadata
La Stampa	IT	24 May 2019	100	0.0%	0.0%	0.0%	0.0%
Le Figaro	FR	24 May 2019	100	70.1%	63.1%	62.1%	72.1%
Le Monde	FR	24 May 2019	100	3.6%	3.3%	3.7%	3.6%
Le Quotidien	LU	24 May 2019	100	24.1%	17.5%	24.0%	24.1%
Le Soir	BE	24 May 2019	100	0.0%	0.0%	0.0%	0.0%



Some comparisons: IPTC metadata usage by country



IPTC Photo Metadata Explorer

[Home](#) / [Countries](#)

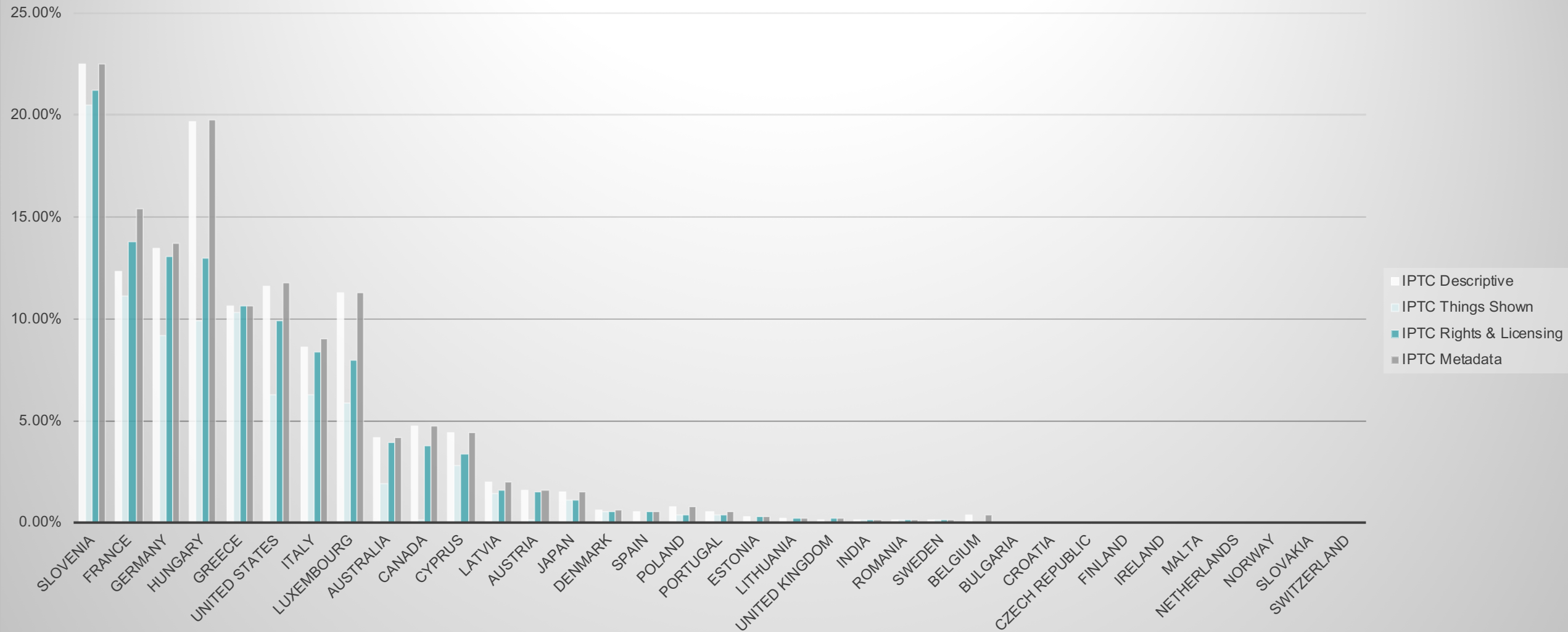
All Countries

Country name	Country code	Number of feeds	IPTC Descriptive	IPTC Things Shown	IPTC Rights & Licensing	IPTC Metadata
Slovenia	SI	3	22.5%	20.5%	21.2%	22.5%
Hungary	HU	4	19.7%	9.9%	13.0%	19.8%
France	FR	6	12.3%	11.1%	13.8%	15.4%
Germany	DE	18	13.5%	9.2%	13.1%	13.7%
United States	US	9	11.6%	6.3%	9.9%	11.8%
Luxembourg	LU	3	11.3%	5.9%	8.0%	11.3%
Greece	GR	3	10.6%	10.3%	10.6%	10.6%
Italy	IT	6	8.6%	6.3%	8.4%	9.0%
Canada	CA	4	4.7%	0.1%	3.8%	4.7%
Cyprus	CY	3	4.4%	2.8%	3.4%	4.4%
Australia	AU	9	4.2%	1.9%	3.9%	4.2%
Latvia	LV	3	2.0%	1.4%	1.6%	2.0%



IPTC Photo Metadata usage in 35 countries studied

IPTC Photo Metadata usage by country





IPTC Photo Metadata - Top 10 countries

Rank	Country name	Number of feeds	IPTC Descriptive	IPTC Things Shown	IPTC Rights & Licensing	IPTC Metadata
1	Slovenia	3	22.50%	20.50%	21.20%	22.50%
2	Hungary	4	19.70%	9.90%	13.00%	19.80%
3	France	6	12.30%	11.10%	13.80%	15.40%
4	Germany	18	13.50%	9.20%	13.10%	13.70%
5	United States	9	11.60%	6.30%	9.90%	11.80%
6	Luxembourg	3	11.30%	5.90%	8.00%	11.30%
7	Greece	3	10.60%	10.30%	10.60%	10.60%
8	Italy	6	8.60%	6.30%	8.40%	9.00%
9	Canada	4	4.70%	0.10%	3.80%	4.70%
10	Cyprus	3	4.40%	2.80%	3.40%	4.40%

Notes:

- Slovenia's and Hungary's high scores are skewed by the fact that one of the few feeds studied has high metadata usage (can be fixed by analysing more feeds)
- France and Germany have relatively high metadata usage, especially given how many feeds we're analysing
- Number 10 in the list has

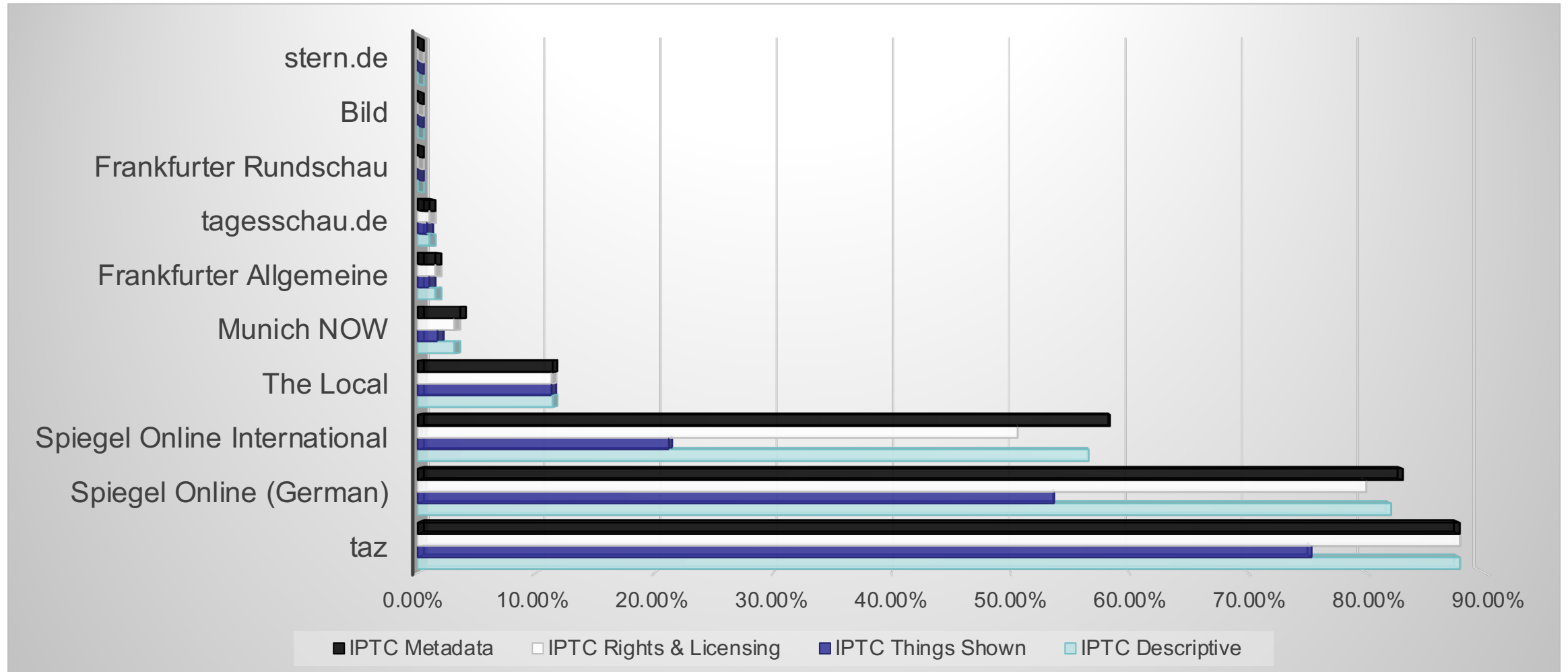


Lots of room for improvement...

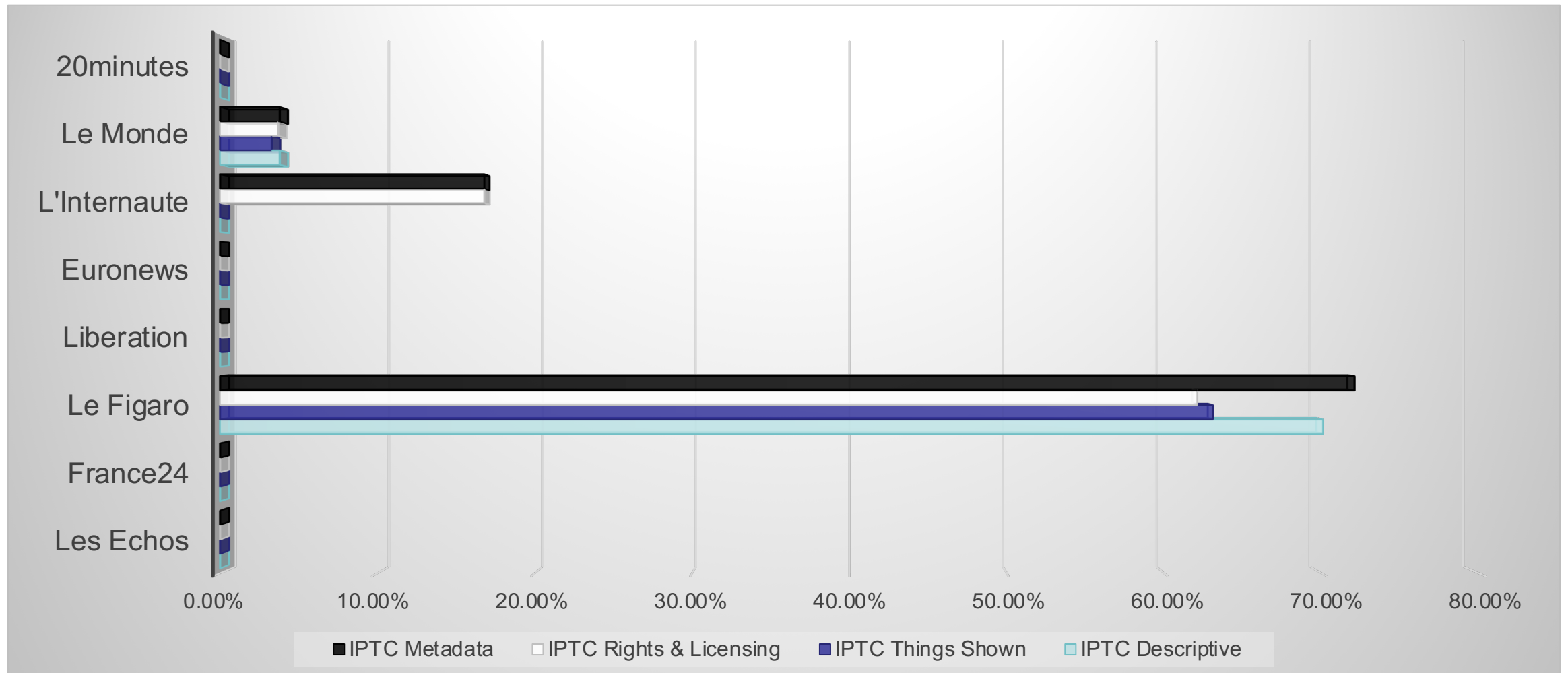
Rank	Country name	Number of feeds	IPTC Descriptive	IPTC Things Shown	IPTC Rights & Licensing	IPTC Metadata
16	Spain	6	0.50%	0.00%	0.50%	0.50%
17	Poland	3	0.80%	0.40%	0.40%	0.80%
18	Portugal	5	0.50%	0.40%	0.40%	0.50%
19	Estonia	6	0.30%	0.00%	0.30%	0.30%
20	Lithuania	4	0.20%	0.00%	0.20%	0.20%
21	United Kingdom	21	0.10%	0.00%	0.20%	0.20%
22	India	5	0.10%	0.10%	0.10%	0.10%
23	Romania	3	0.10%	0.10%	0.10%	0.10%
24	Sweden	5	0.10%	0.10%	0.10%	0.10%
25	Belgium	3	0.40%	0.00%	0.00%	0.40%
26	Bulgaria	3	0.00%	0.00%	0.00%	0.00%
27	Croatia	4	0.00%	0.00%	0.00%	0.00%
28	Czech Republic	6	0.00%	0.00%	0.00%	0.00%
29	Finland	3	0.00%	0.00%	0.00%	0.00%
30	Ireland	4	0.00%	0.00%	0.00%	0.00%
31	Malta	3	0.00%	0.00%	0.00%	0.00%
32	Netherlands	4	0.00%	0.00%	0.00%	0.00%
33	Norway	6	0.00%	0.00%	0.00%	0.00%
34	Slovakia	3	0.00%	0.00%	0.00%	0.00%
35	Switzerland	2	0.00%	0.00%	0.00%	0.00%

- 20 countries (of 35 studied) have little or no embedded metadata
- We have our work cut out for us!

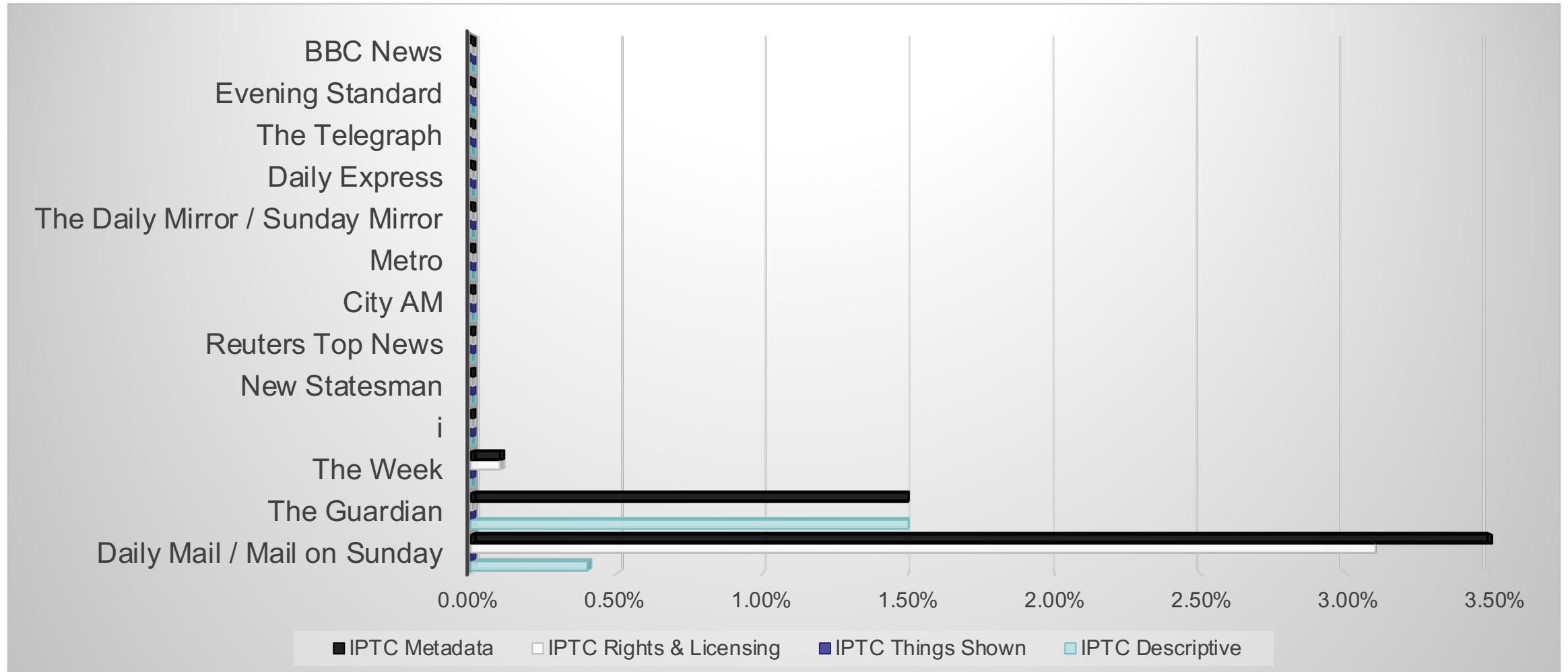
Feeds in Germany



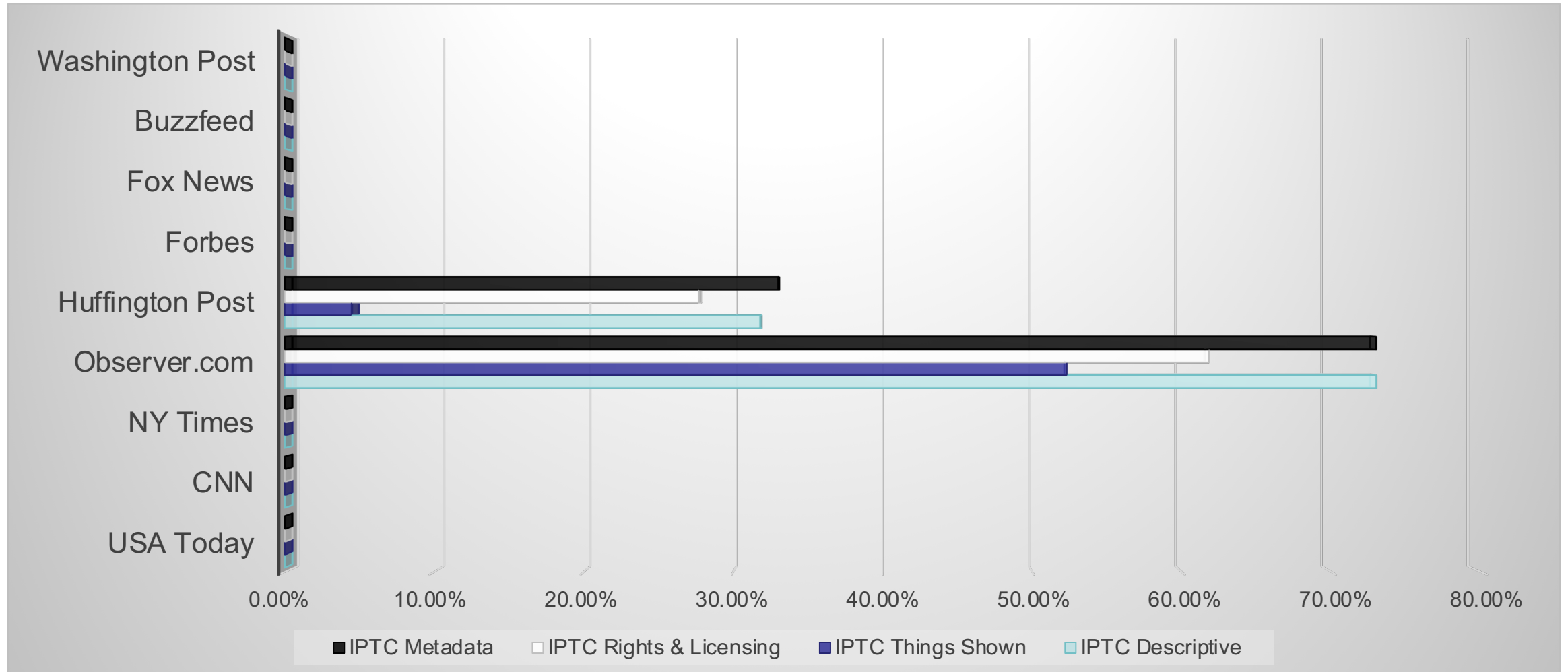
Feeds in France



Feeds in United Kingdom

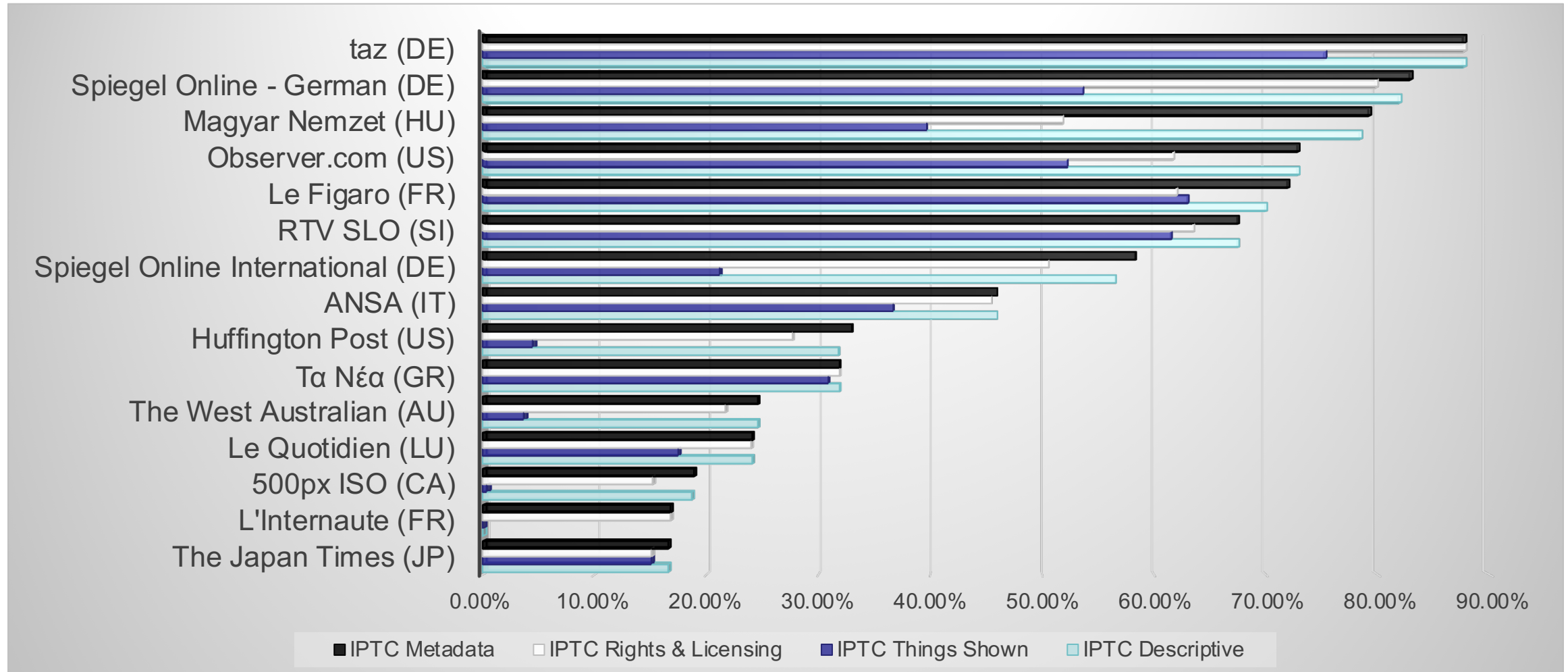


Feeds in United States



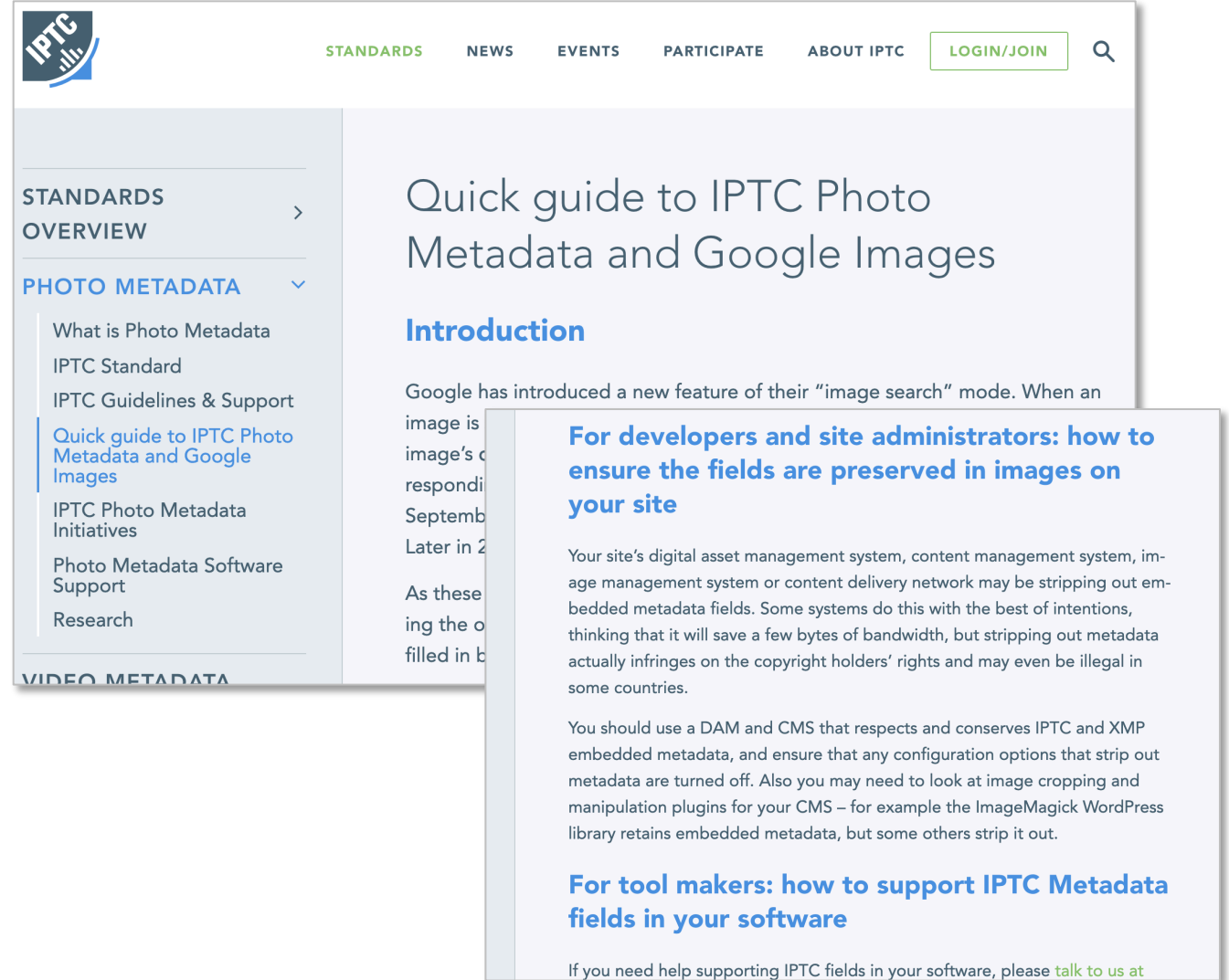
And now for the final results...

Overall top feeds globally



Conclusions

- A lot of education is needed
- Hopefully publishing some of this information will encourage publishers to treat embedded metadata more kindly
- Better tools are still needed
- Awareness will drive better support in DAMs, CMSs, CDNs
- More documents like the [“Quick Guide to IPTC Photo Metadata and Google Images”](#) could help



The screenshot shows the IPTC website with a navigation bar at the top containing links for STANDARDS, NEWS, EVENTS, PARTICIPATE, ABOUT IPTC, and a LOGIN/JOIN button. A search icon is also present. On the left, a sidebar menu lists various resources under 'STANDARDS OVERVIEW' and 'PHOTO METADATA', with the 'Quick guide to IPTC Photo Metadata and Google Images' article highlighted. The main content area displays the title of the article, an introduction paragraph, and two callout boxes. The first callout box, titled 'For developers and site administrators: how to ensure the fields are preserved in images on your site', discusses the importance of preserving metadata in digital asset management systems. The second callout box, titled 'For tool makers: how to support IPTC Metadata fields in your software', provides advice on configuring DAM and CMS systems to retain metadata. Both callouts include links to further resources.

For developers and site administrators: how to ensure the fields are preserved in images on your site

Your site's digital asset management system, content management system, image management system or content delivery network may be stripping out embedded metadata fields. Some systems do this with the best of intentions, thinking that it will save a few bytes of bandwidth, but stripping out metadata actually infringes on the copyright holders' rights and may even be illegal in some countries.

You should use a DAM and CMS that respects and conserves IPTC and XMP embedded metadata, and ensure that any configuration options that strip out metadata are turned off. Also you may need to look at image cropping and manipulation plugins for your CMS – for example the ImageMagick WordPress library retains embedded metadata, but some others strip it out.

For tool makers: how to support IPTC Metadata fields in your software

If you need help supporting IPTC fields in your software, please [talk to us at](#)



Next steps

- Publish initial results on iptc.org blog
- Make the browser available to IPTC member companies
- Add some features:
 - Month-by-month comparisons over time
 - Get better at only extracting editorial images (we're currently grabbing some incidental images e.g. from Outbrain panels)
 - Compare usage of particular fields across publishers and countries
- Add more feeds, more countries
- All suggestions greatly received!



Please get in touch!

Twitter: @brendanquinn / @IPTC

Email: mdirector@iptc.org