

Moving to a data-driven Content Discovery strategy

STM Digital Publishing 2017

Mike Roberts
Content Discovery Manager
Emerald Group Publishing

What is 'Content Discovery'?



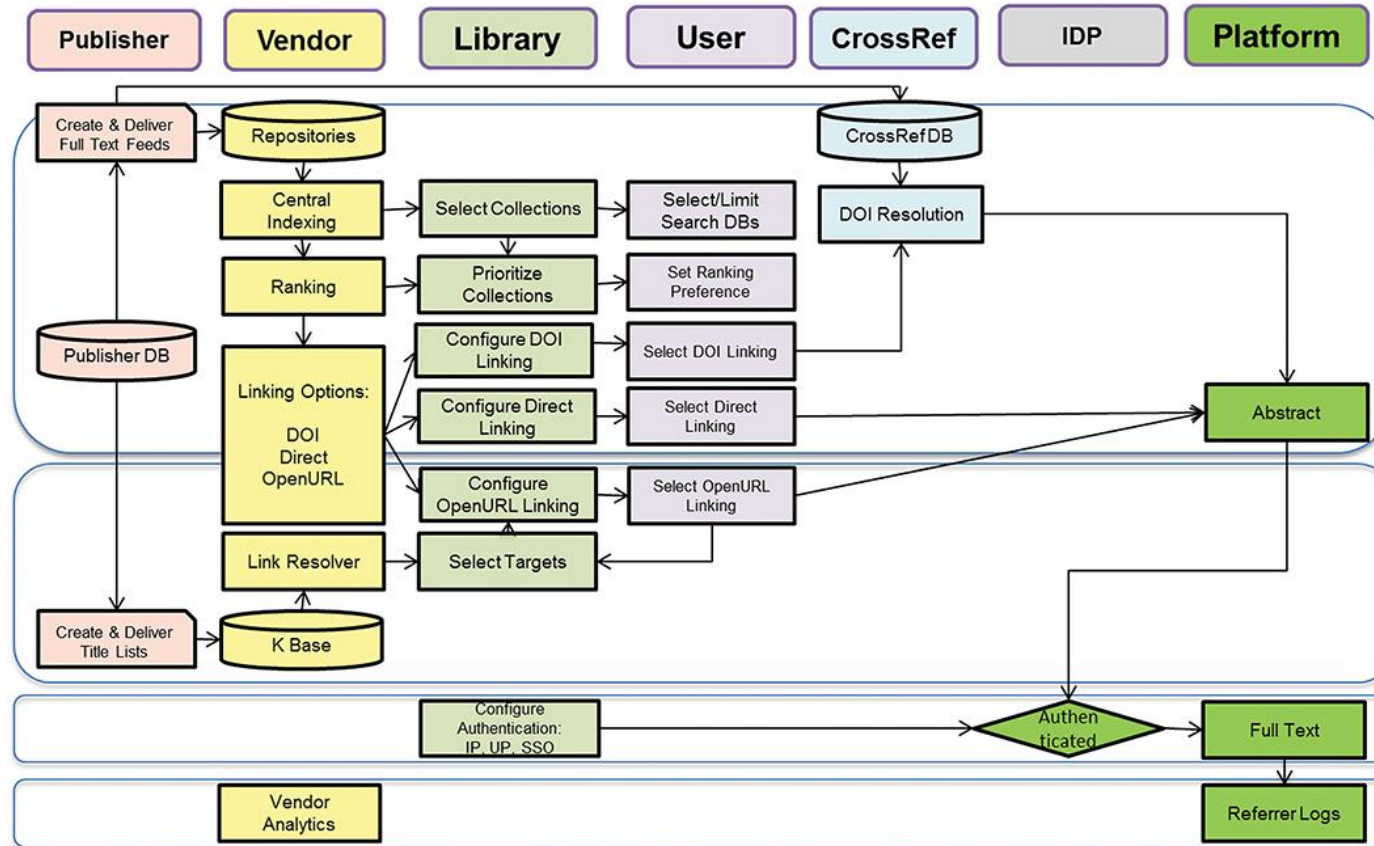
Helping content find its audience



*Drury Lane Theatre,
from the Stage during the performance.
Published Aug. 21. 1804, by Richard Phillips, 71, St. Paul's Church Yard.*



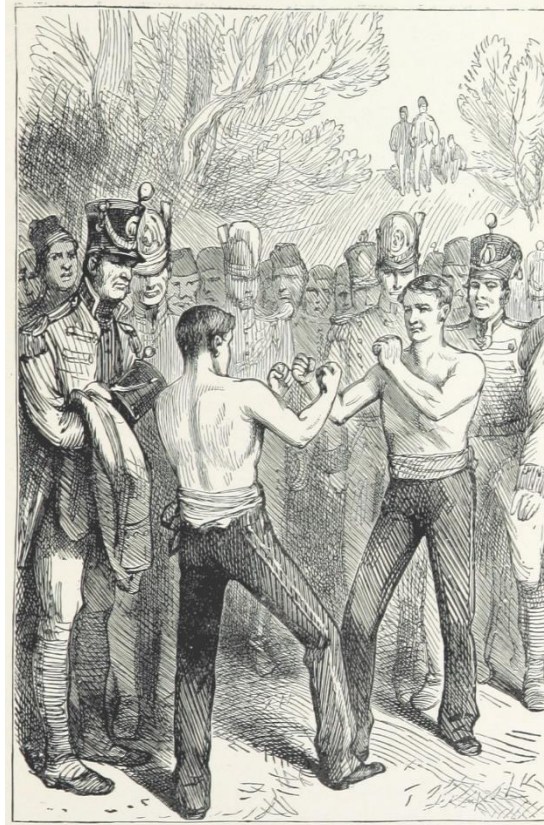
Discovery Service Data Flows



Google

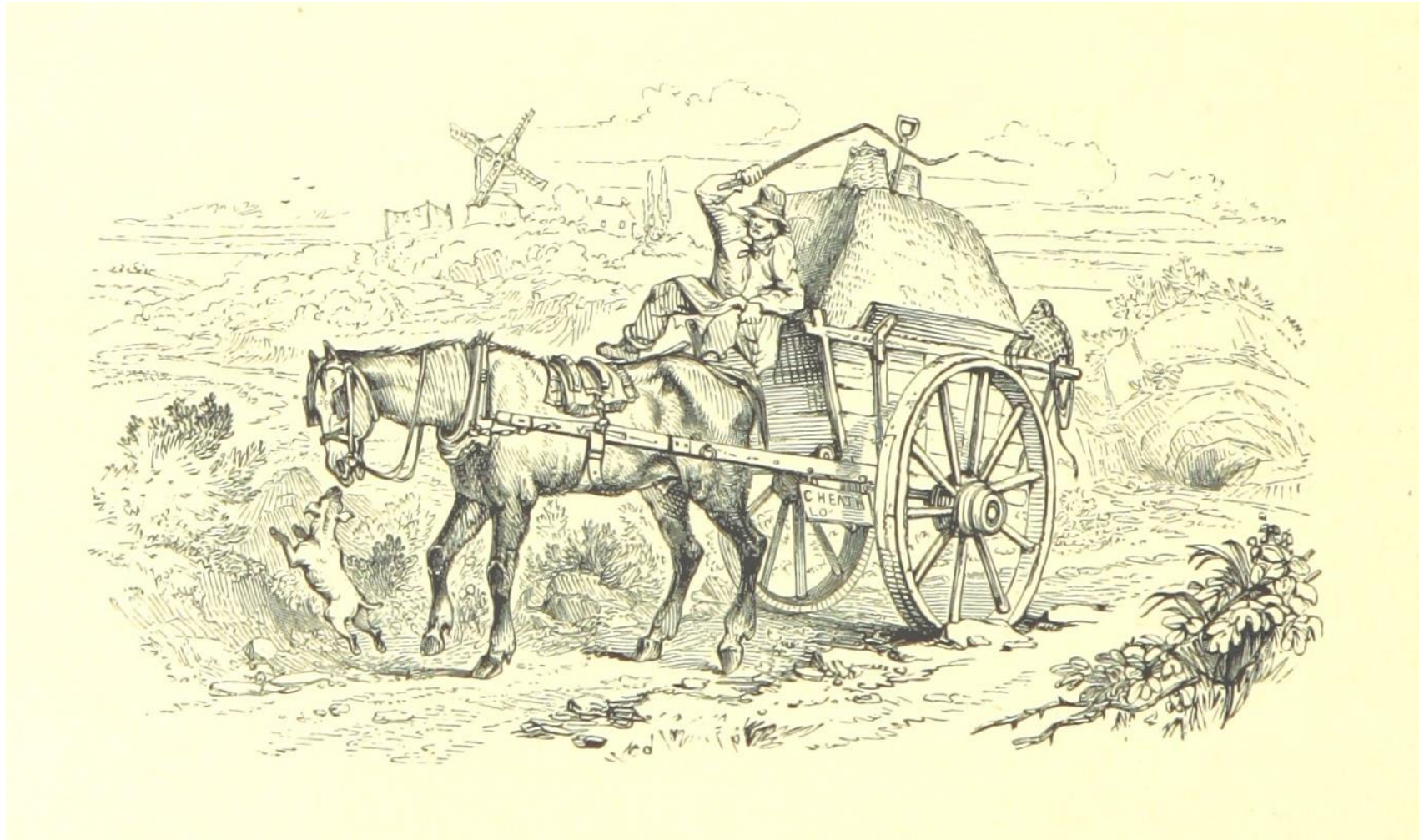


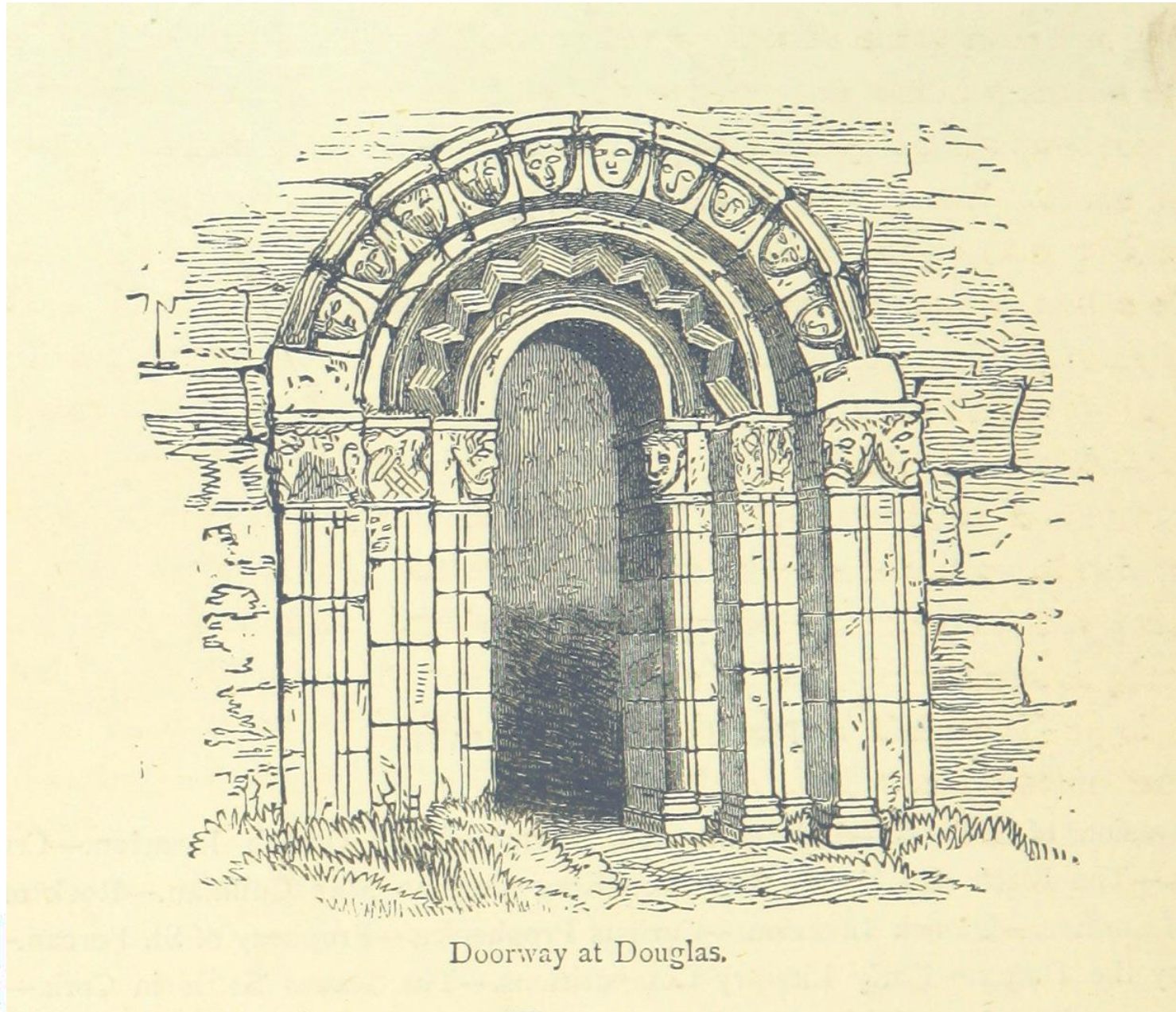
Metadata vs Full-text





Metadata deliveries





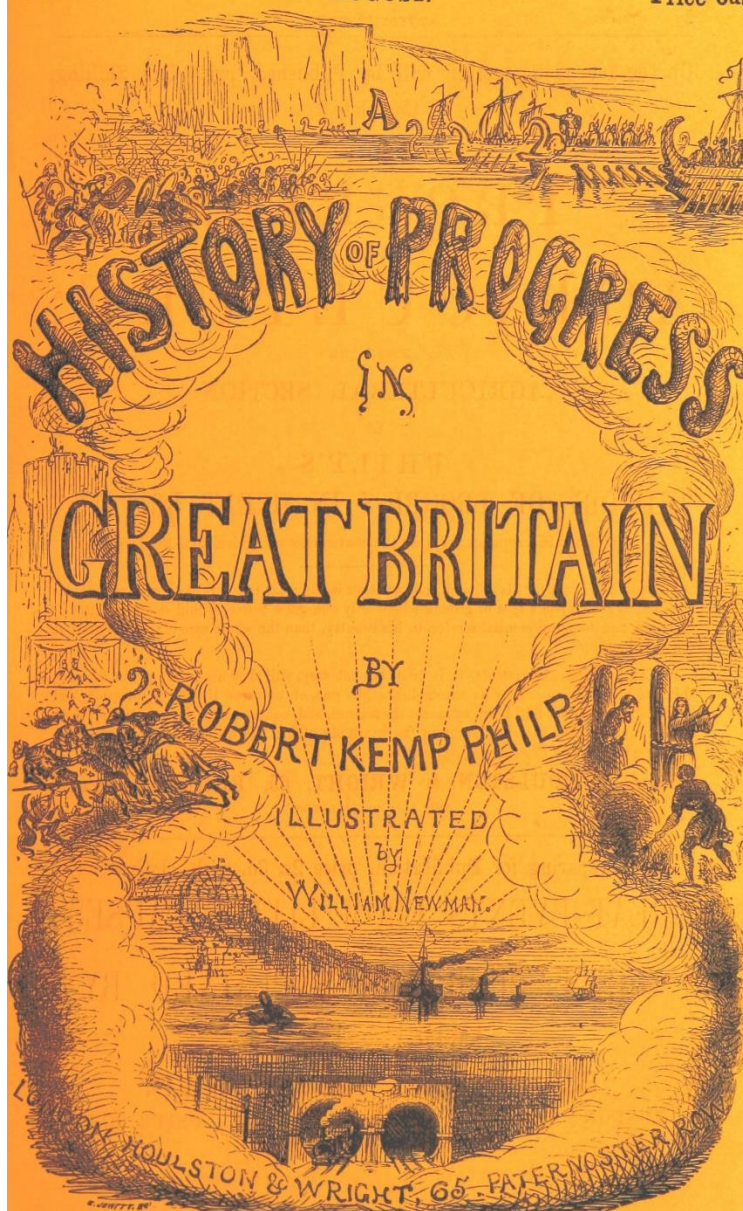
Doorway at Douglas.



No. III.

AUGUST.

Price 6d.



HISTORY OF PROGRESS
IN
GREAT BRITAIN

BY
ROBERT KEMP PHILP

ILLUSTRATED

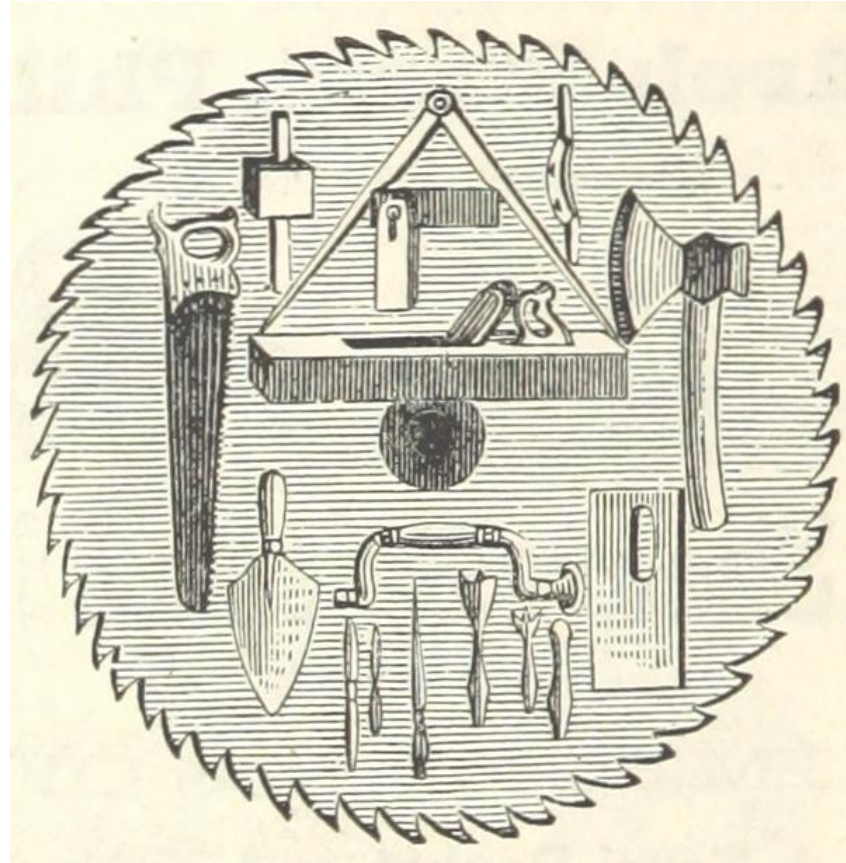
BY
WILLIAM NEWMAN

HOLLSTON & WRIGHT, 65, PATERNOSTER ROW

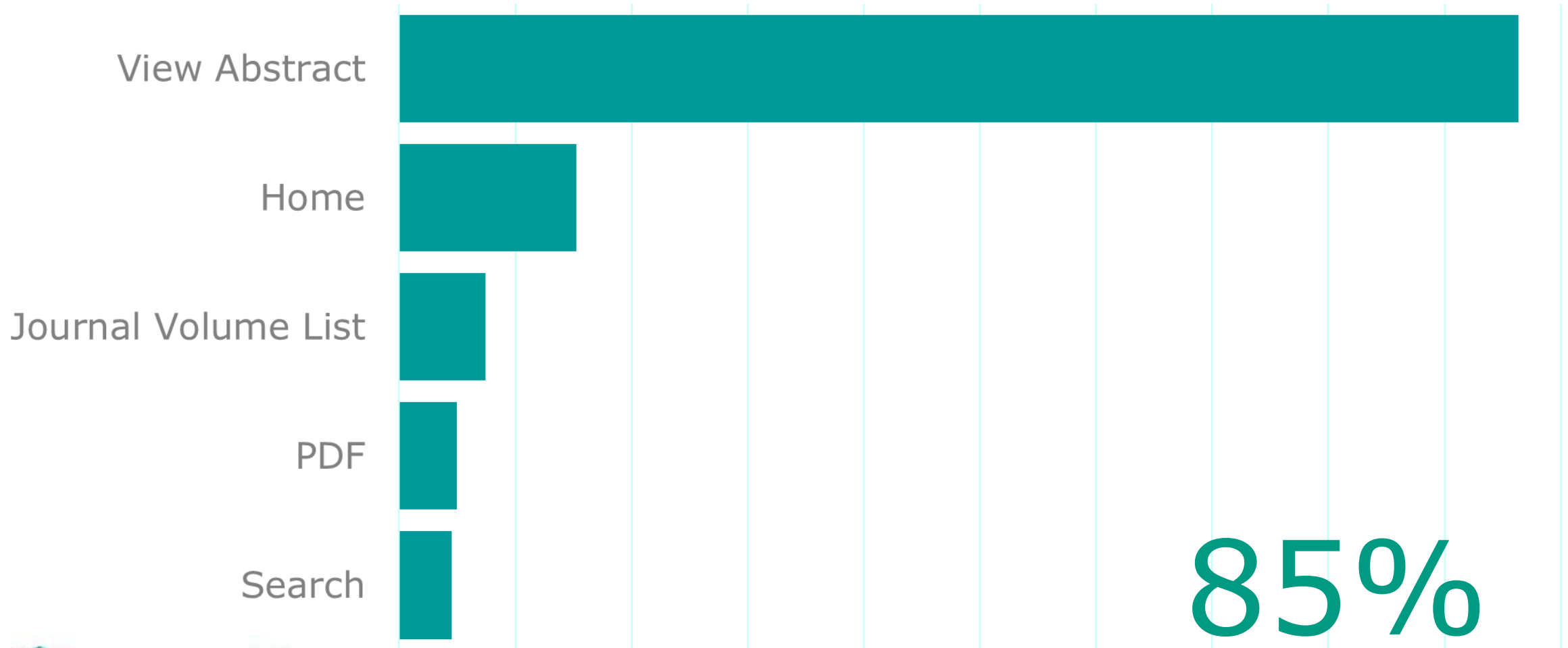
On- vs off-site discovery



Tools



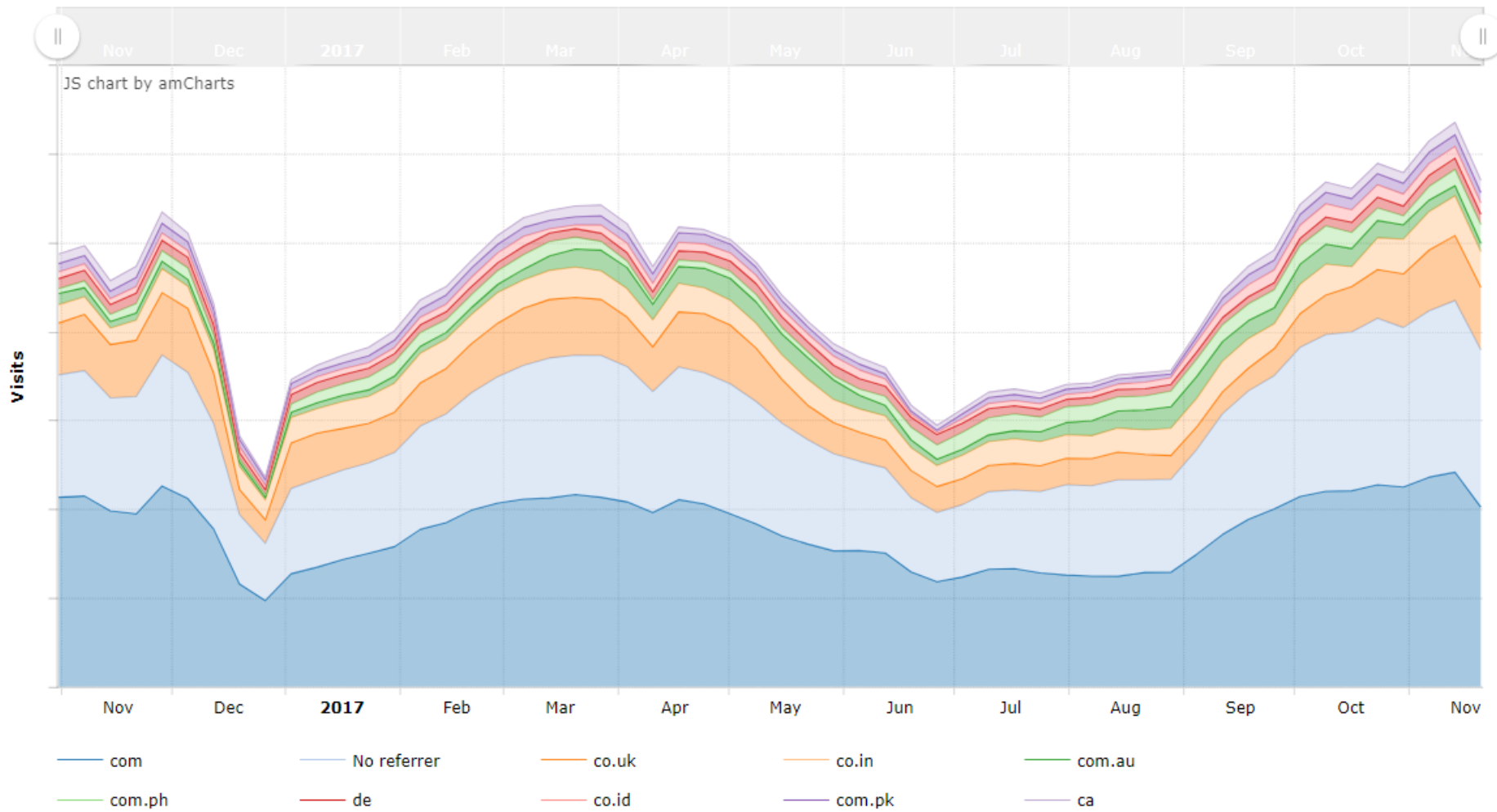
Landing page analysis



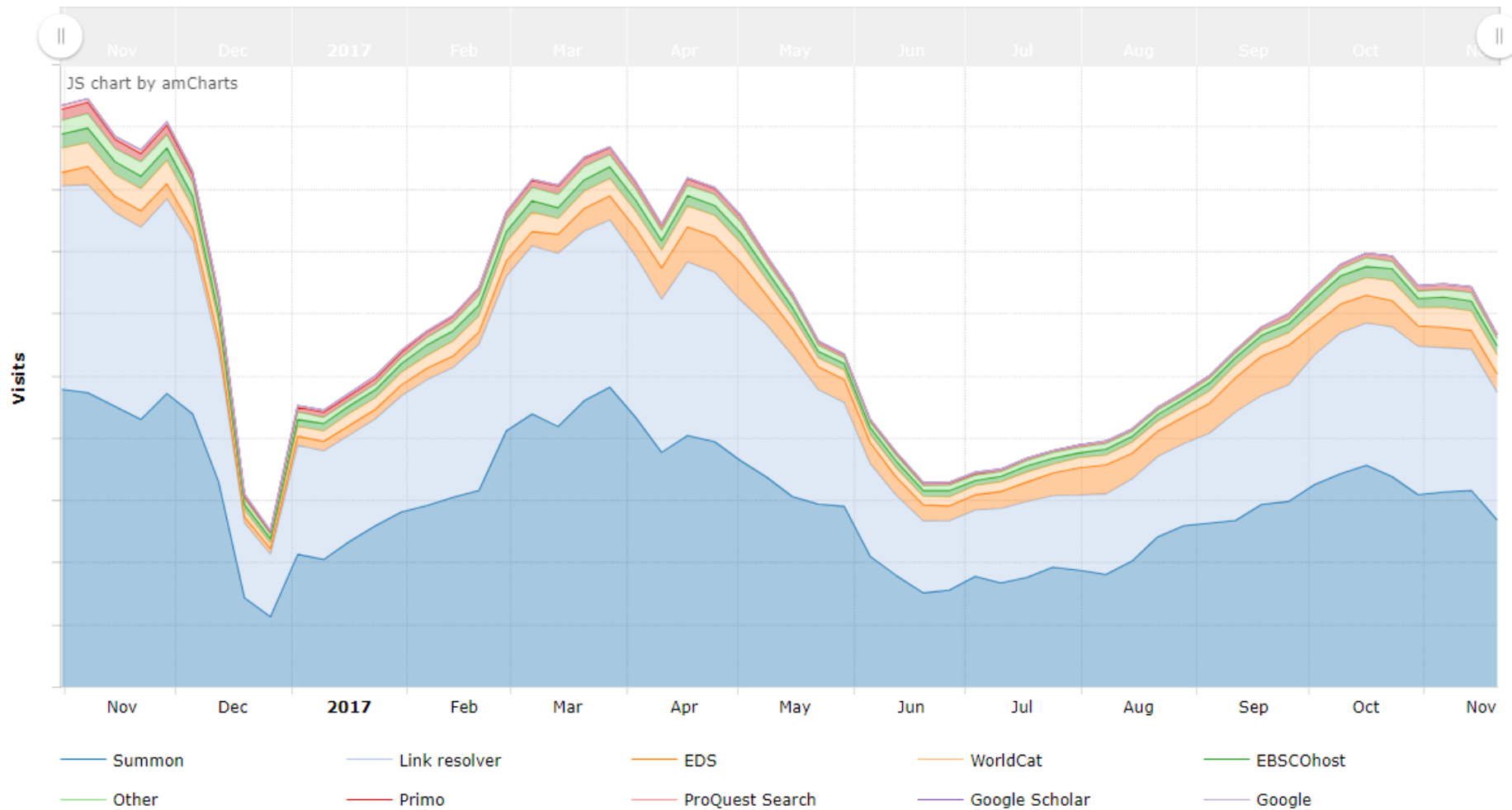
Reporting



Web referral analysis



Source aggregation



SFX

EZ Proxy

EBSCO

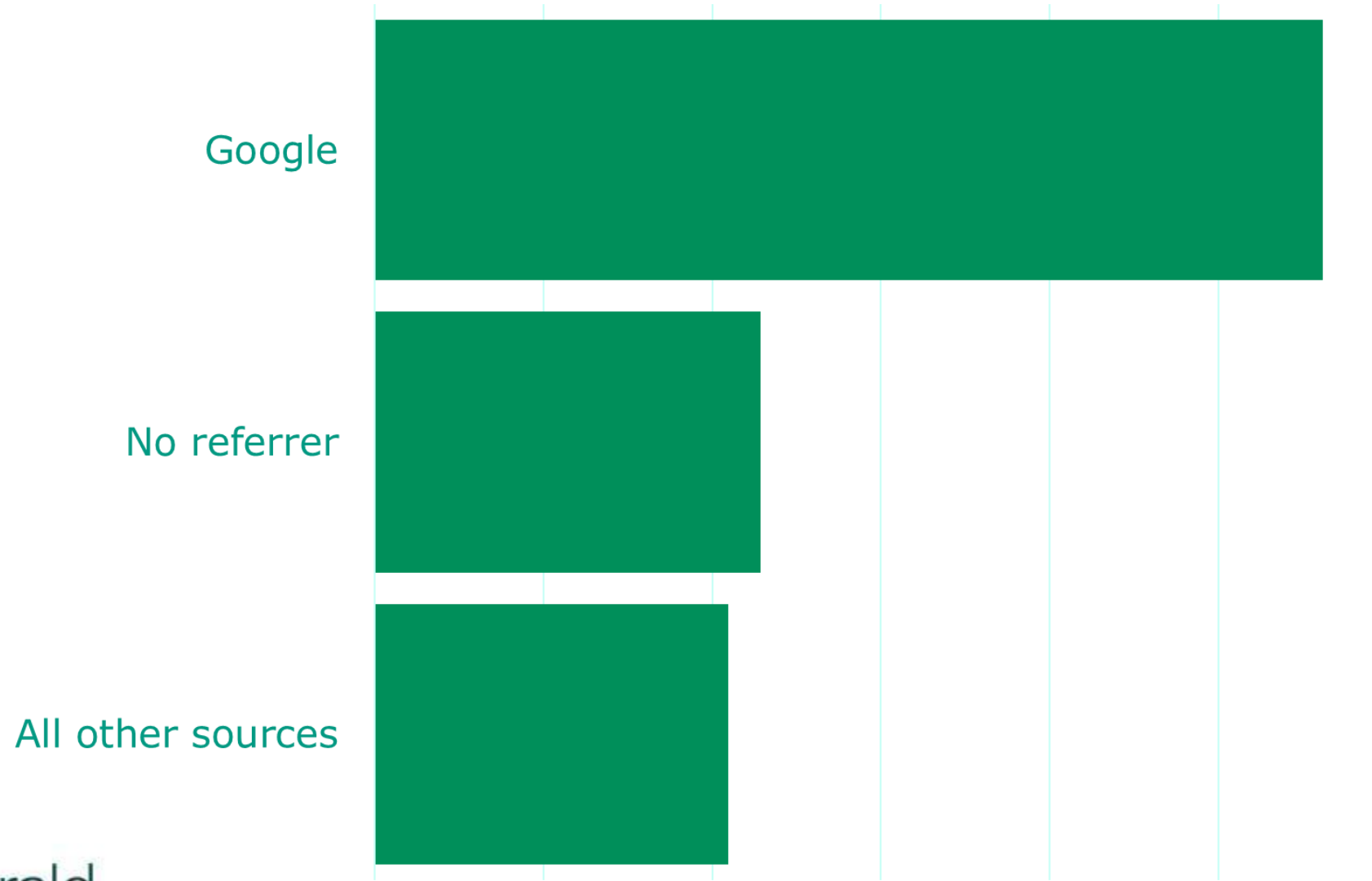
Alma



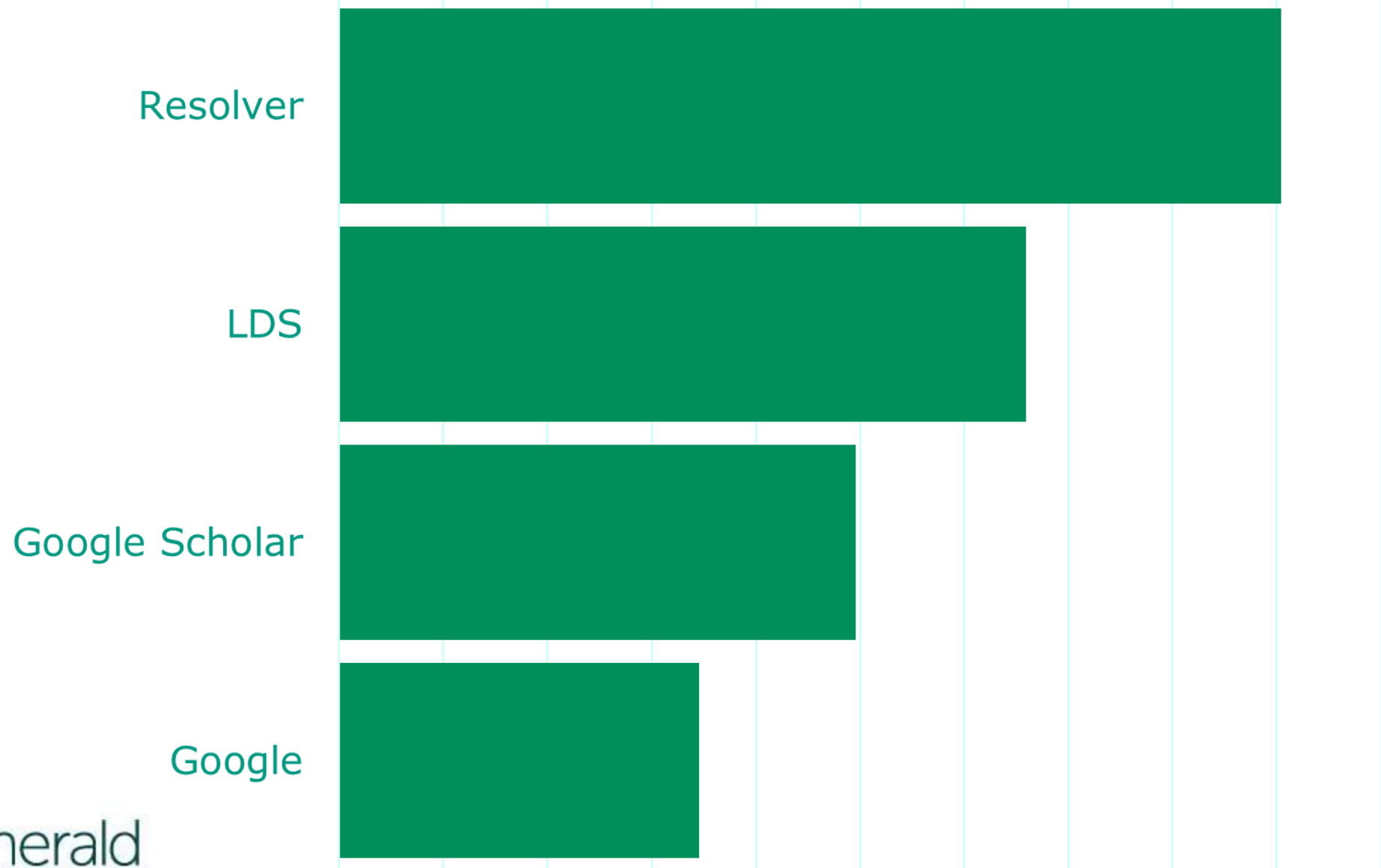
Where are users coming from?



Traffic sources



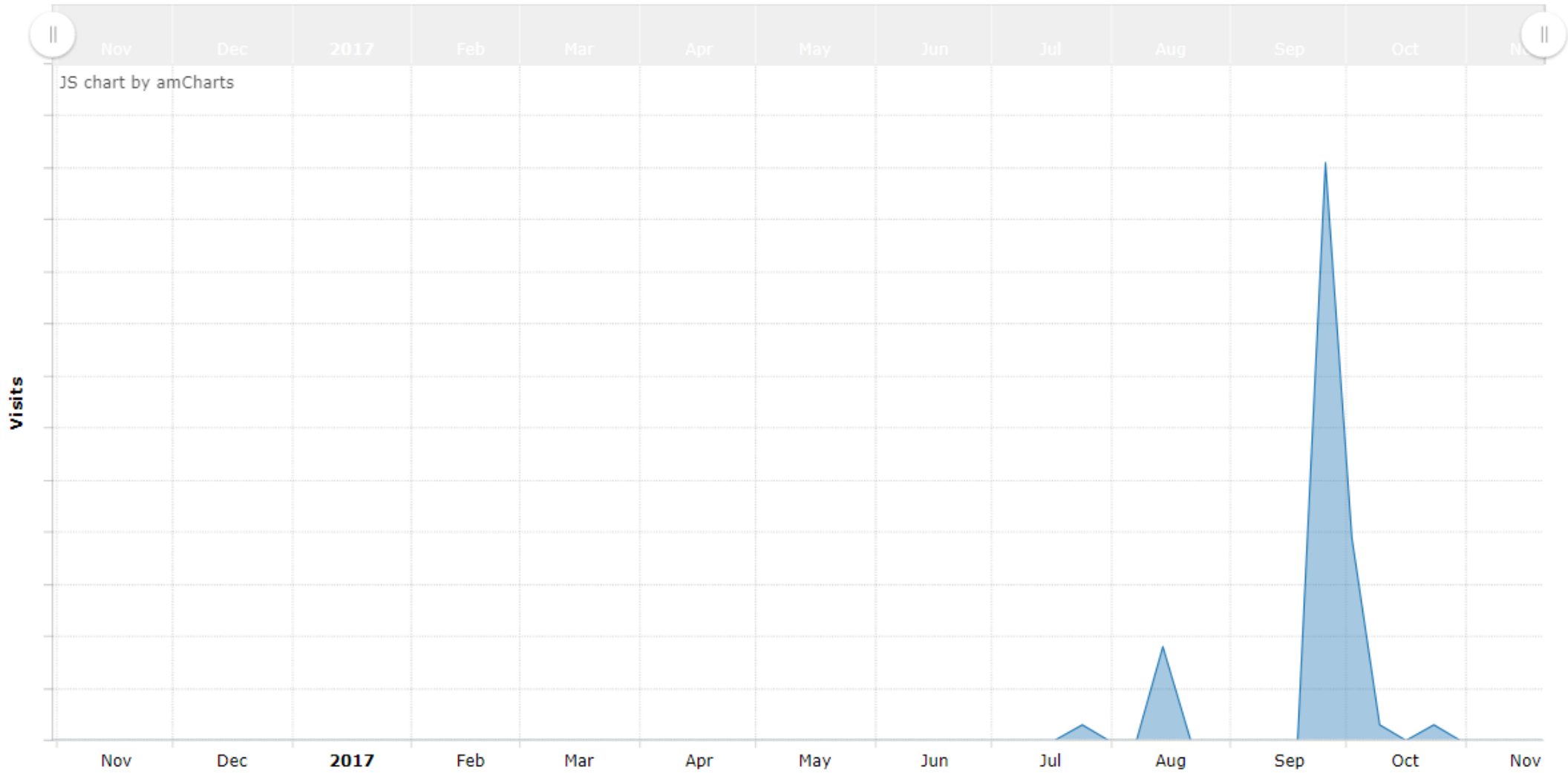
FTD Success Rate (%) by Referrer



Trends in referrals data



Referrals from bbc.com



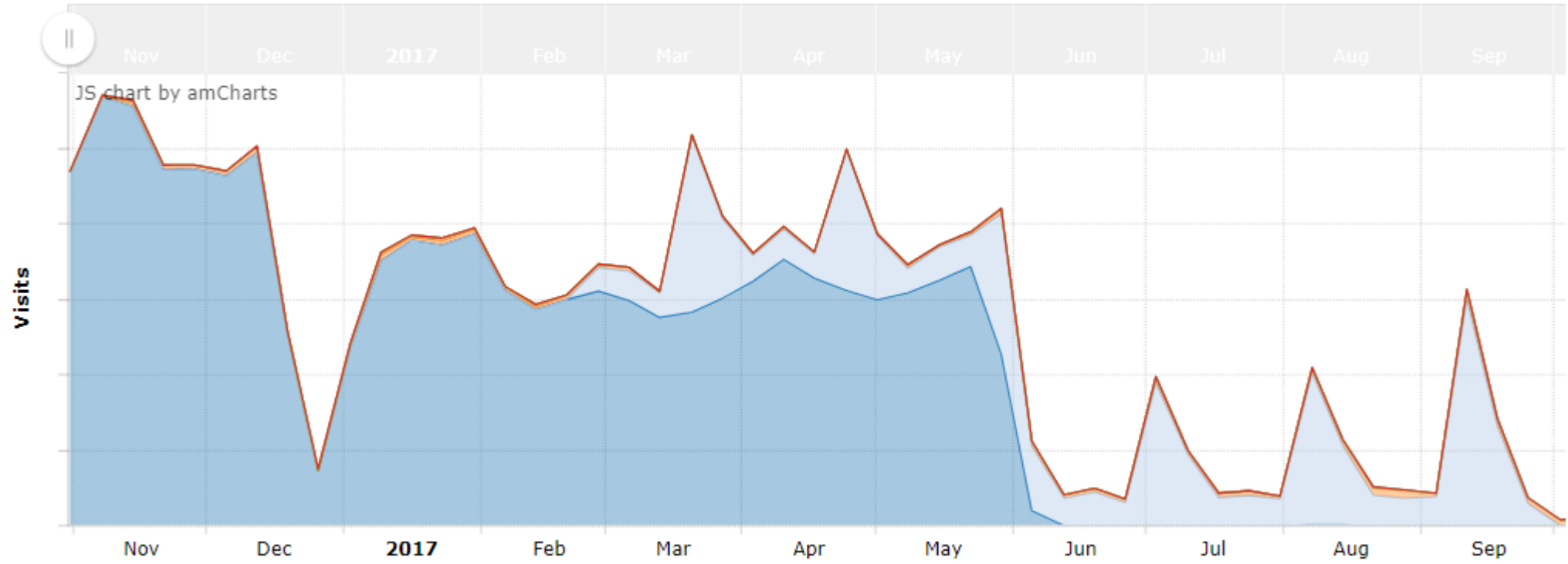


During the Cold War, Vozrozhdeniya Island was a top-secret testing ground for deadly Soviet super-pathogens. Despite over two decades of abandonment, their legacy lives on.

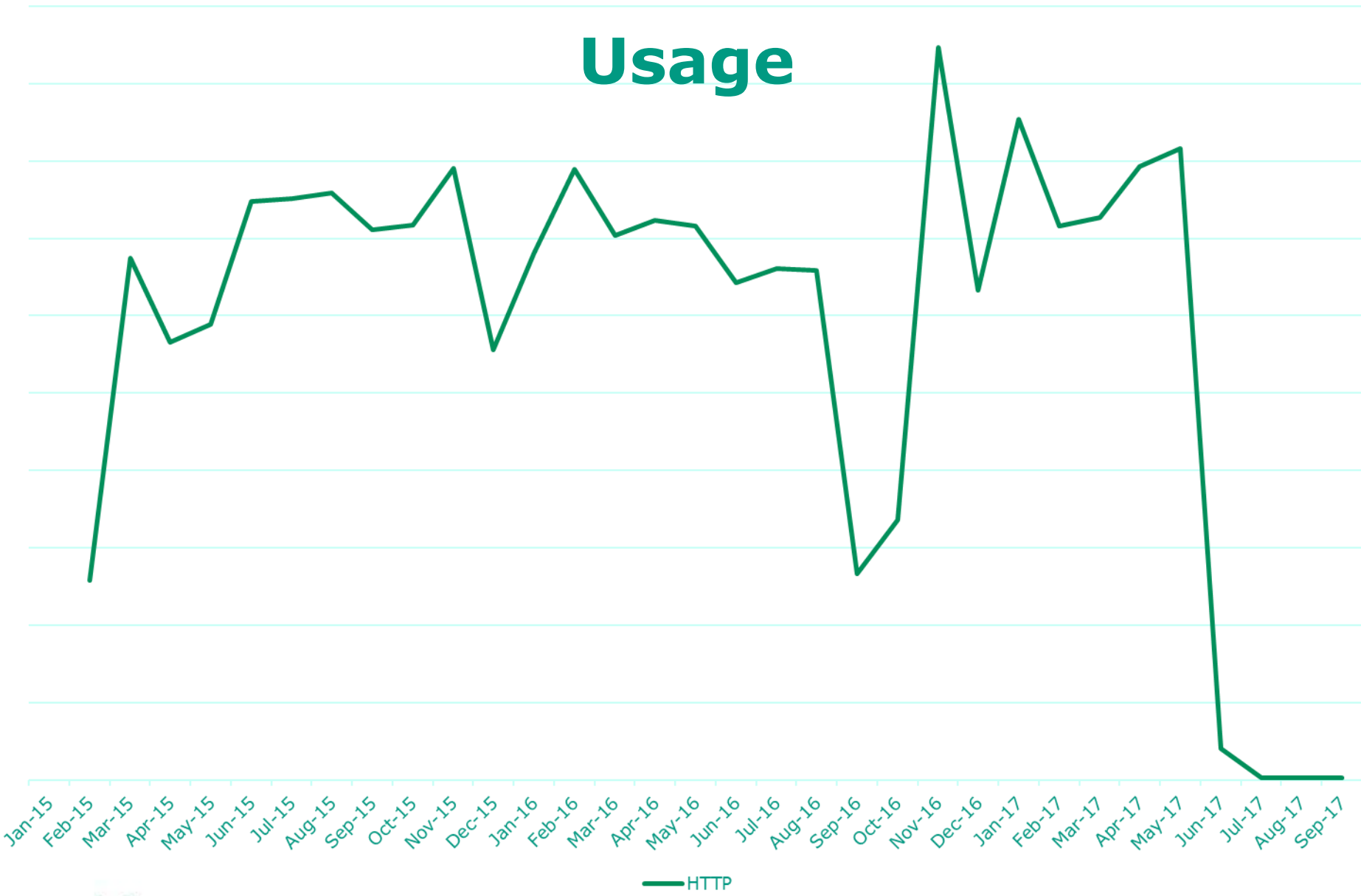




Web referrals



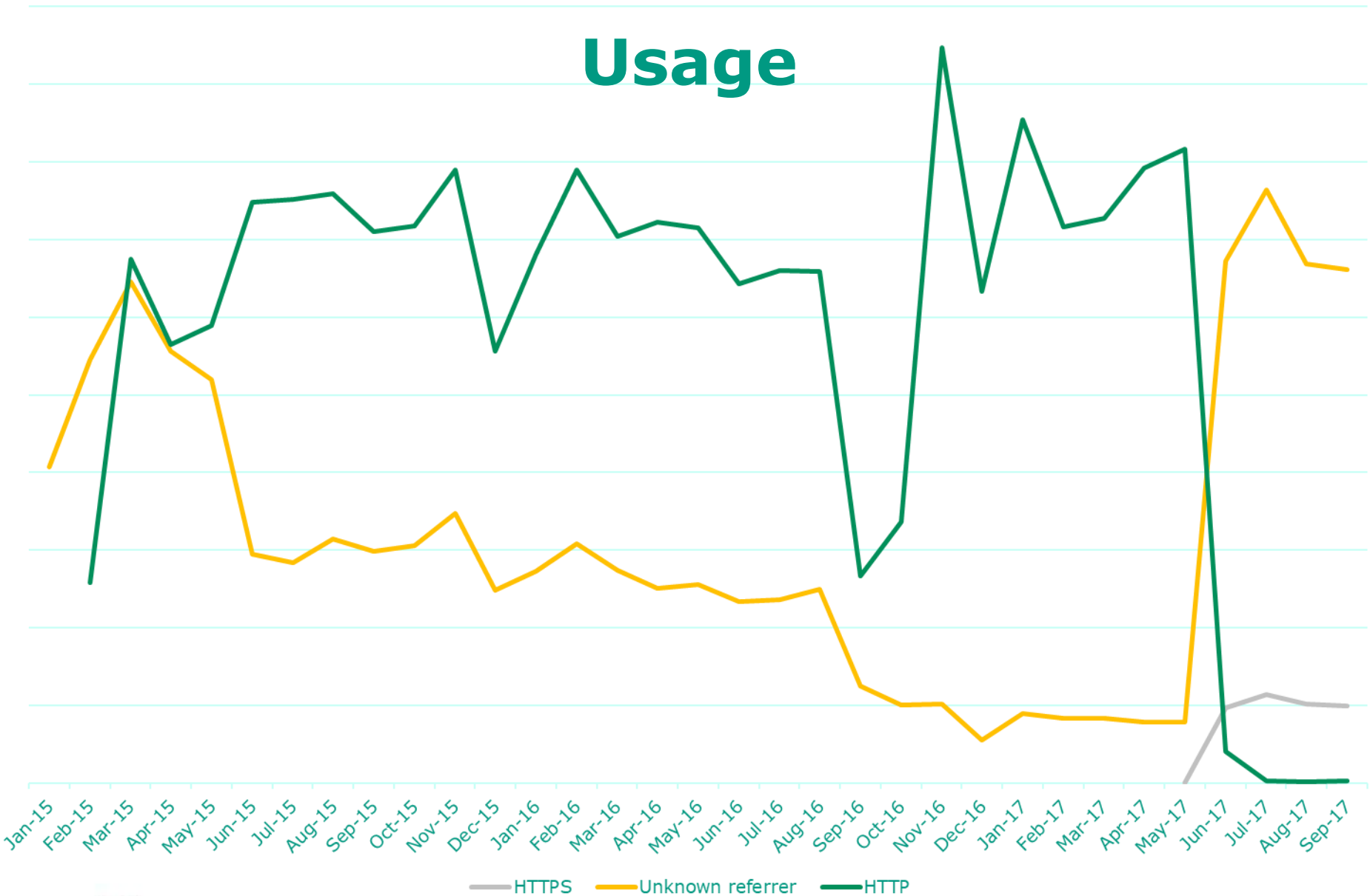
Usage



— HTTP

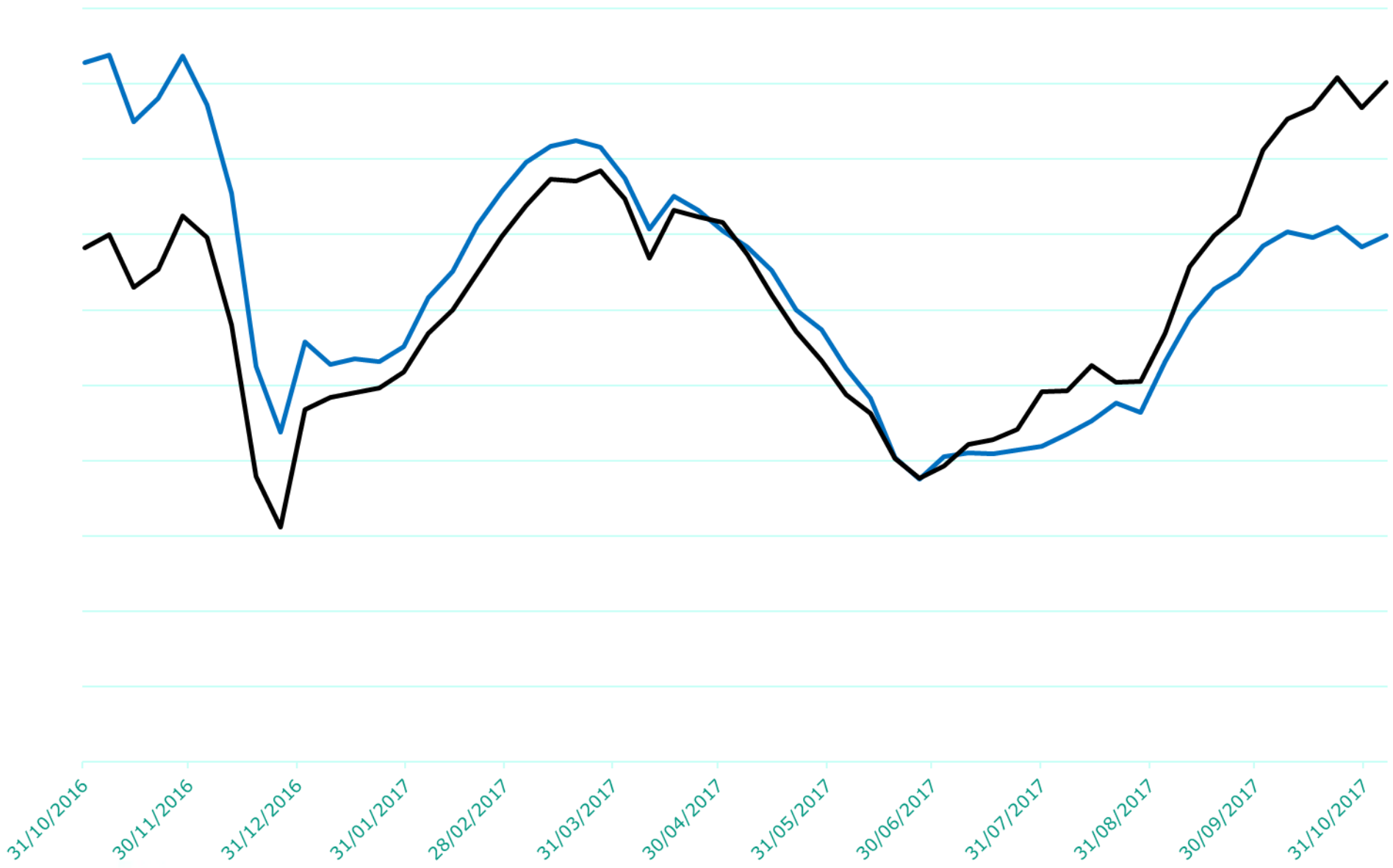


Usage

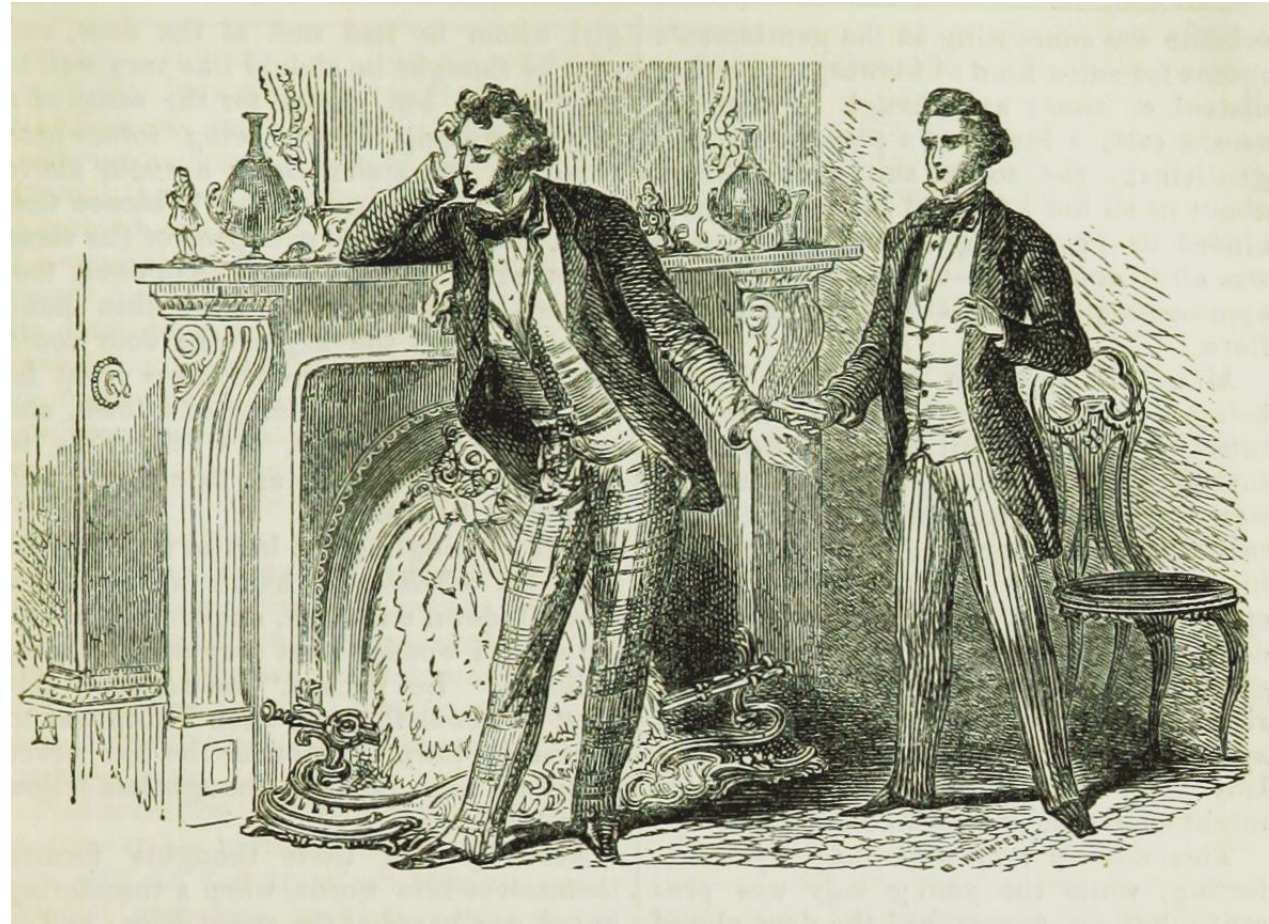


— HTTPS — Unknown referrer — HTTP

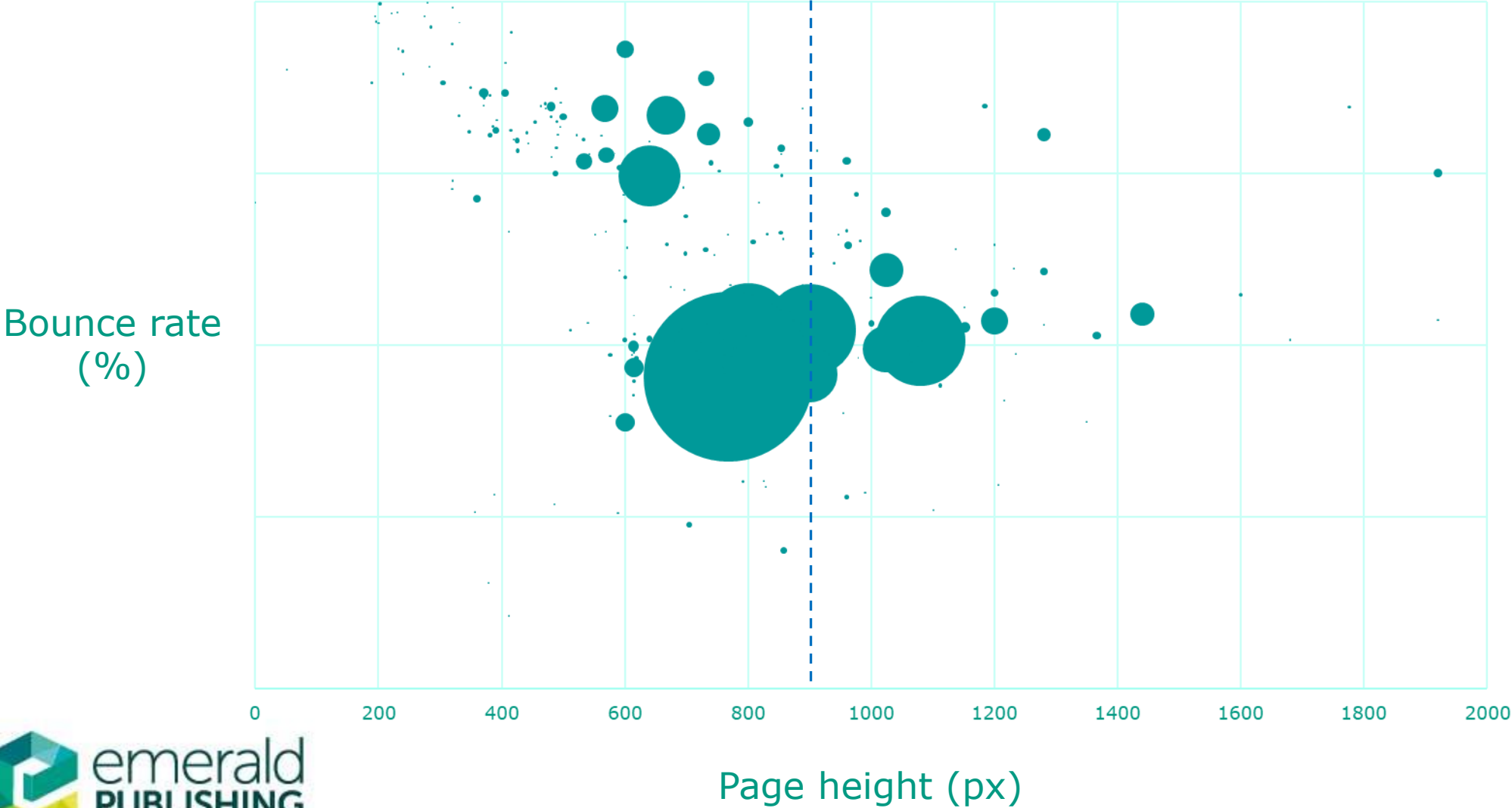




Where is it going wrong?



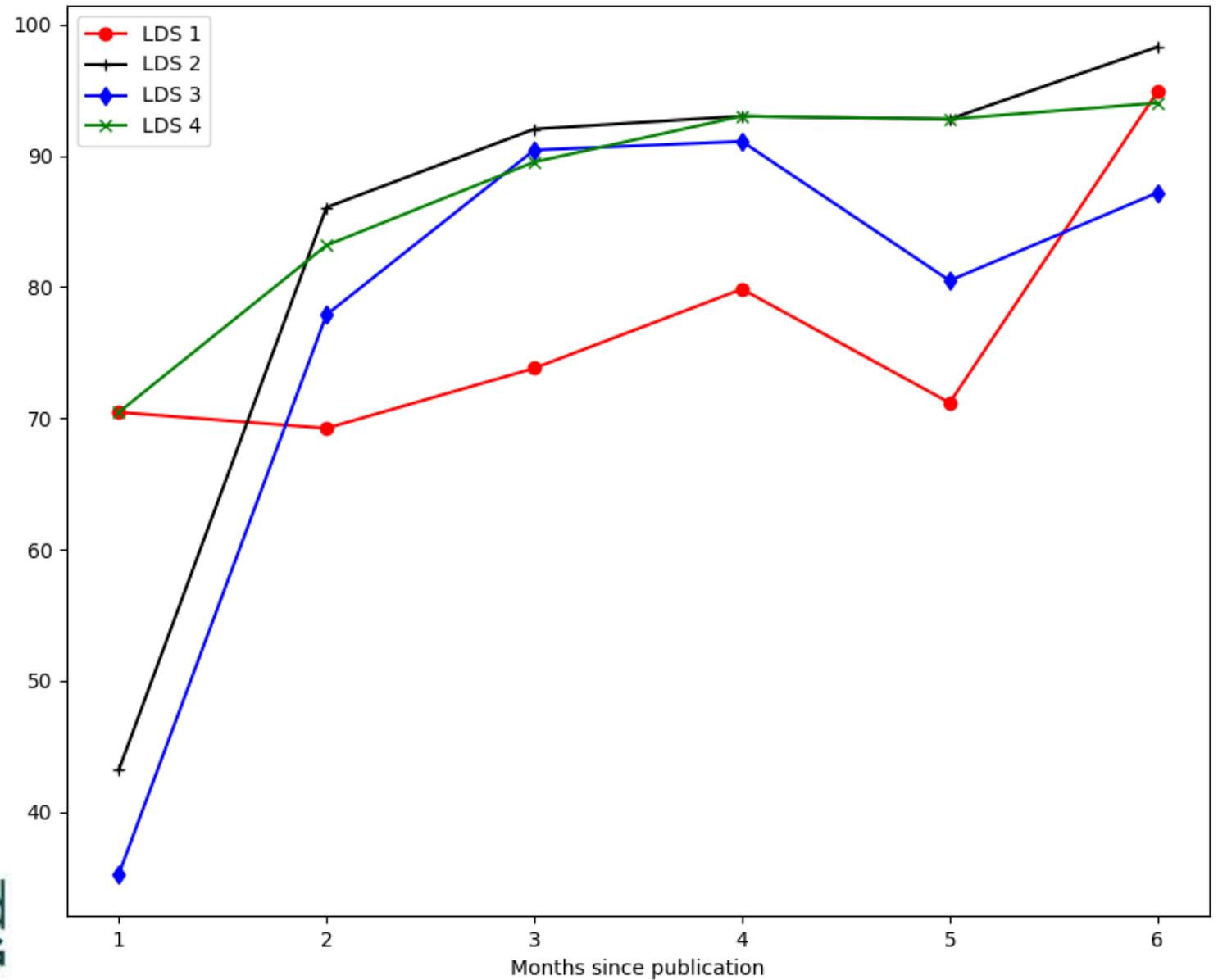
Page height vs bounce rate



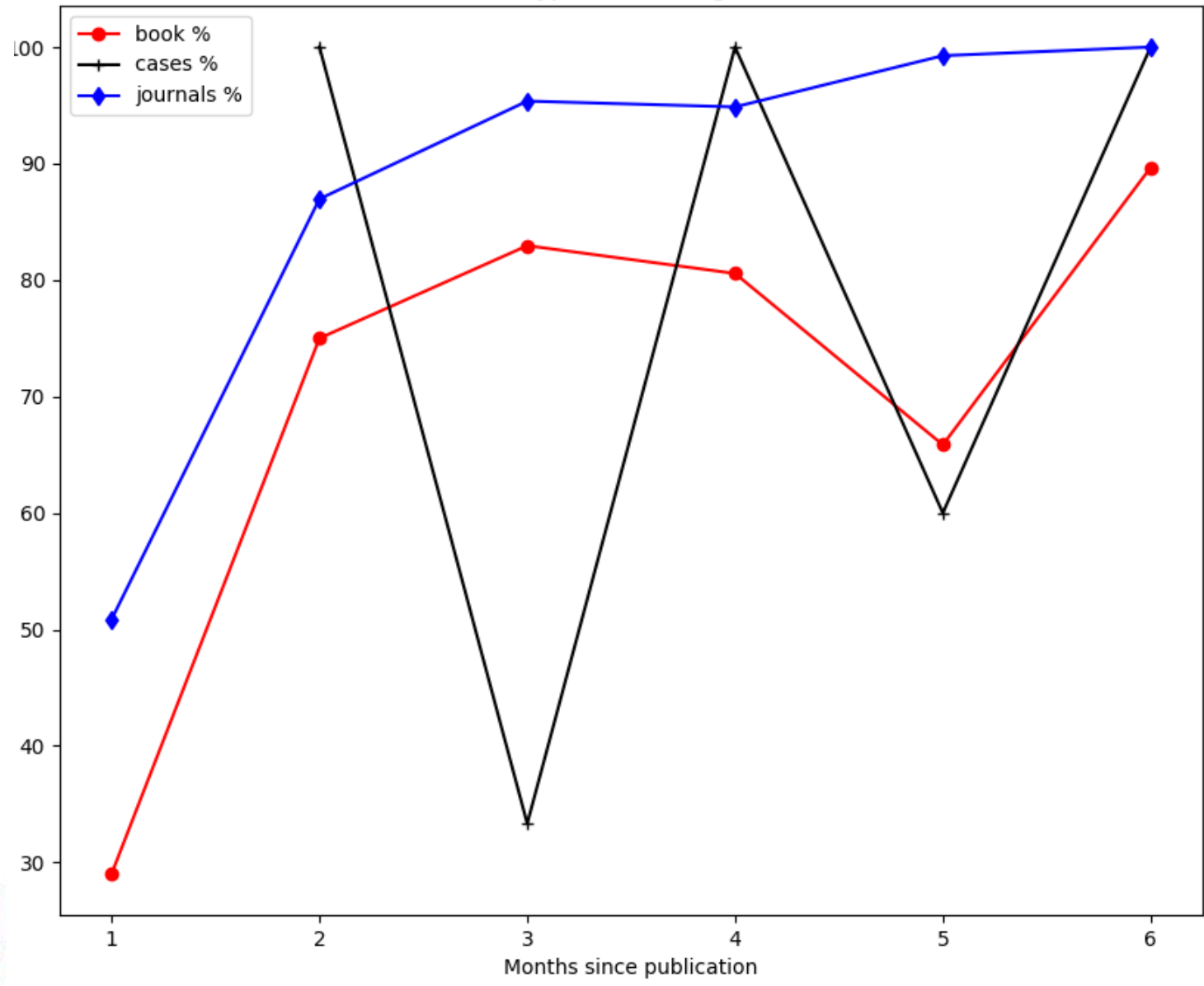
LDS audits



Total: % coverage



Type: % coverage



Customer site visits



Customer site visits

- ▶ SSO not working
- ▶ CSS being lost
- ▶ Platform sign-in bug
- ▶ Metadata issues



What next?



Immediate actions

- ▶ Change article page layout
- ▶ HTTPS
- ▶ Diversify traffic sources
 - ▶ TrendMD & ScienceOpen pilots
 - ▶ Yewno & Semantic Scholar
- ▶ Collaboration with LDS vendors



Further analysis & data sources

- ▶ Understand bounce rates
 - ▶ Possible connection with turnaway rates
- ▶ How well is authentication working?
- ▶ On-site discovery
- ▶ Crossref metadata audits
- ▶ Crossref event data API – real-time insight

