Exploring the French Web of the 90s

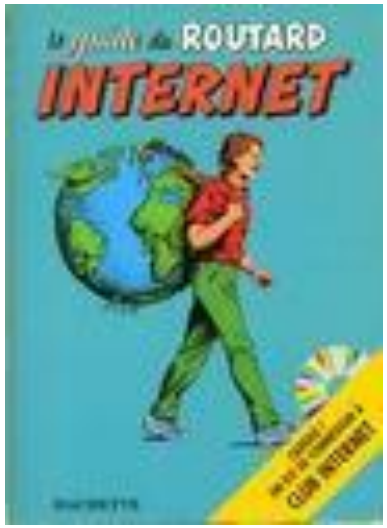
Valérie Schafer

C²DH
University of Luxembourg

Web Archiving Conference
June 2019

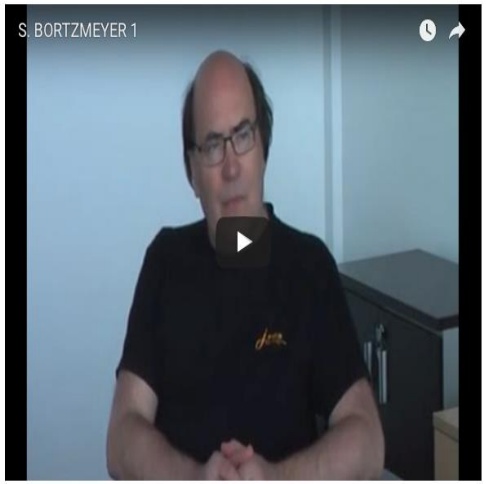
UNIVERSITÉ DU LUXEMBOURG

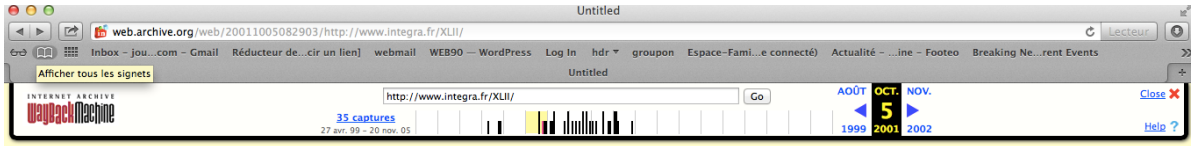
C²DH
LUXEMBOURG CENTRE FOR CONTEMPORARY AND DIGITAL HISTORY



Entretien avec Stéphane Bortzmeyer – 1ère partie

Le 23 avril 2015, Stéphane Botzmeyer était notre premier grand témoin interviewé dans le cadre du projet Web90 (par Camille Paloque-Berges et Valérie Schafer) pour revenir sur sa découverte et son expérience du Web des années 90.





Le site de Quarante-Deux se trouve maintenant à
http://www.quarante-deux.org/

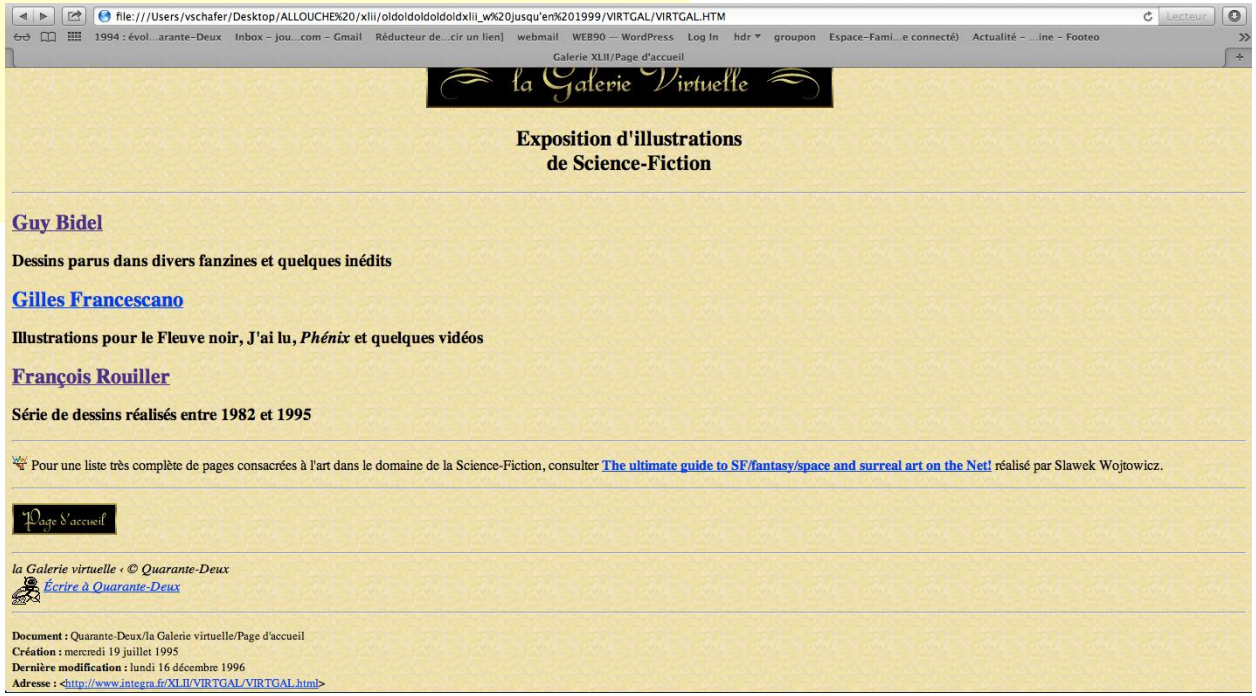
Mettez à jour vos signets/favoris/bookmarks...

The website of Quarante-Deux is now at http://www.quarante-deux.org/

Please update your bookmarks.

---

quarante-deux.org
First archive in the Wayback Machine in 2000



Quarante-Deux
Historique de l'évolution du site

plan du site

Recherche

tout Quarante-Deux

Années

2014
2013
2012
2011
2010
2009
2008
2007

Année 1994

Ce qui se passe ensuite en 1995

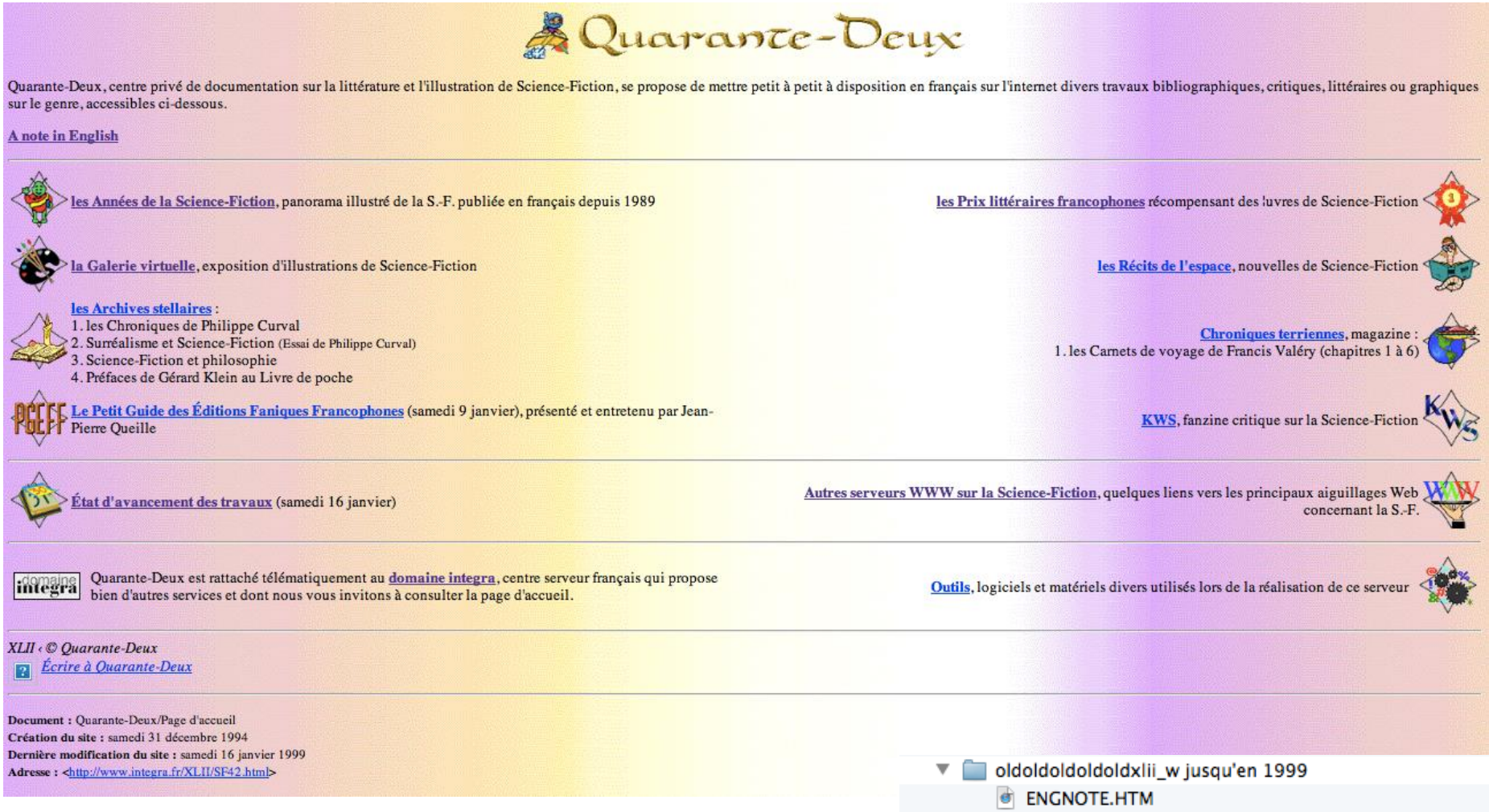
Étapes regroupées par année depuis la création du site, fin 1994, et sa date de première mise en ligne, le 11 février 1995 ; à lire de manière chronologique à partir de la fin de chaque page…

1994, octobre à décembre

Dialogue : « De la doc sur la DTD HTML, on en trouve sur le net mais c'est assez parcellaire et ça se contredit souvent…
— Il y en a pourtant qui arrivent à mettre des pages en ligne, non…? ».

1994, printemps

Annonce sur CalvaCom : « Un forum a été créé en cité RCI où vous pourrez mentionner les URL que vous avez jugées intéressantes ; n'hésitez pas à y contribuer… »



The first quarante-deux webpage

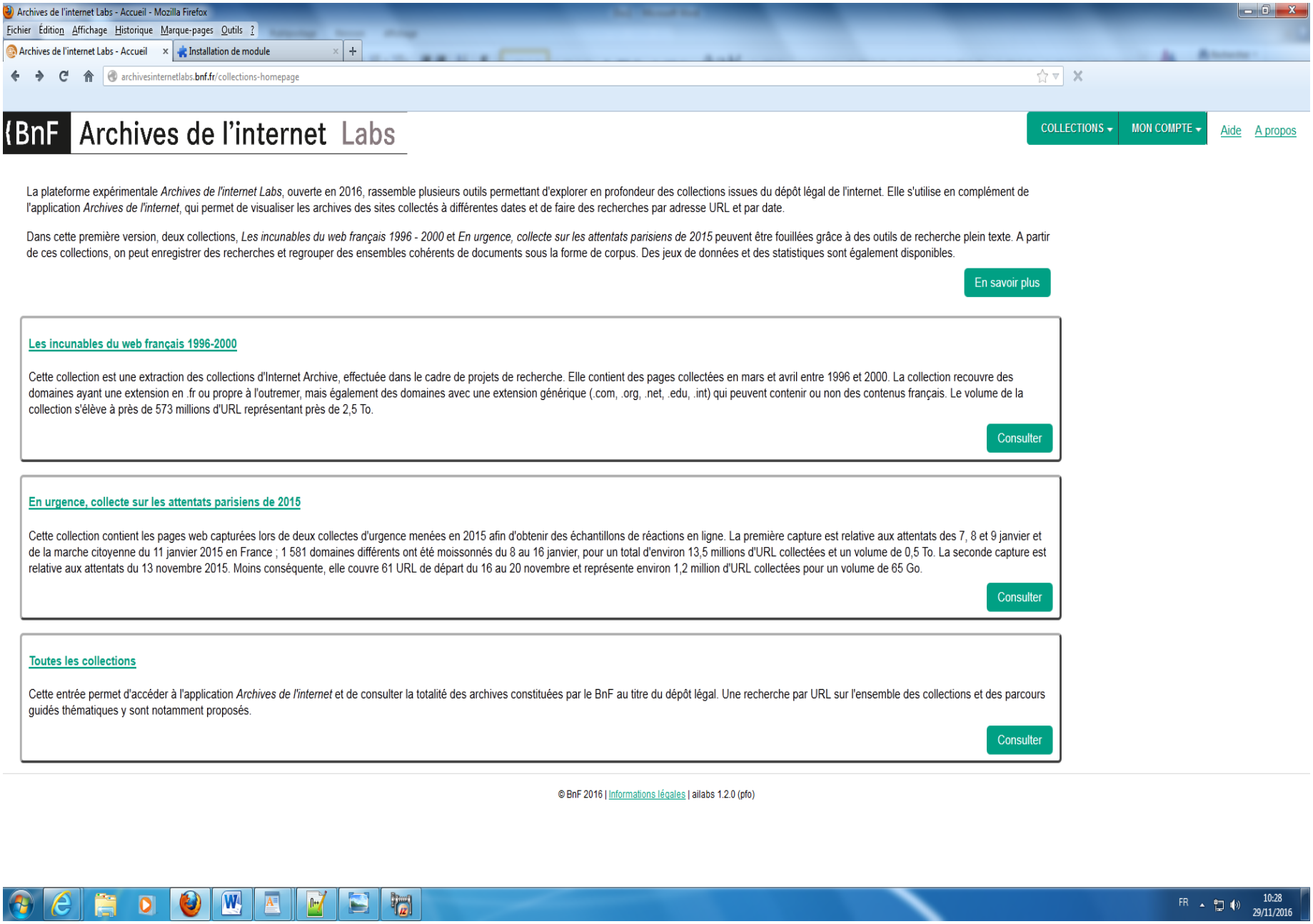
The Historian of the Web : Crawler, Browser or Lurker?

March 13, 2015 | Uncategorized | 1990s, digital heritage, Louise Merzeau, Nineties, Web90
peterwebster

[A special guest post by **Valérie Schafer** and **Francesca Musiani**.]

Web explorers… "Small is beautiful"!

In the Web90 project, so as to account for the set of conditions that have shaped the Web experience of Internet users in the Nineties – especially their browsing habits – we have chosen to study on one hand the general framework and the power of 'massification'. But on the other hand, and in parallel, we wish to 'stay close to the archive', and eventually to follow paths previously traced by others (e.g. directories or closed spaces such as Infonie), open doors such as ISP portals or the Yahoo! Directory, or those sites that were recommended by the press or by guidebooks, such as the 1998 *Guide du Routard de l'Internet*.) Of course, in this domain, Web archives are precious assets, but the word "browser" takes on here its full original etymological sense: to encompass our mobilization of printed sources (press archives, State-driven reports, guidebooks for the general public) but also audio-visual archives, oral testimonies, or newsgroups. These sources invite historians to become lurkers around exchanges past. But very often, they soon need to emerge from this "passive" status…


Archives de l'internet Labs - Accueil - Mozilla Firefox
Fichier Édition Affichage Historique Marque-pages Outils ?
Archives de l'internet Labs - Accueil
Installation de module
archivesinternetlabs.bnf.fr/collections-homepage
BnF Archives de l'internet Labs
COLLECTIONS
MON COMPTE
Aide
A propos
La plateforme expérimentale Archives de l'internet Labs, ouverte en 2016, rassemble plusieurs outils permettant d'explorer en profondeur des collections issues du dépôt légal de l'internet. Elle s'utilise en complément de l'application Archives de l'internet, qui permet de visualiser les archives des sites collectés à différentes dates et de faire des recherches par adresse URL et par date.
Dans cette première version, deux collections, Les incunables du web français 1996 - 2000 et En urgence, collecte sur les attentats parisiens de 2015 peuvent être fouillées grâce à des outils de recherche plein texte. A partir de ces collections, on peut enregistrer des recherches et regrouper des ensembles cohérents de documents sous la forme de corpus. Des jeux de données et des statistiques sont également disponibles.
En savoir plus
Les incunables du web français 1996-2000
Cette collection est une extraction des collections d'Internet Archive, effectuée dans le cadre de projets de recherche. Elle contient des pages collectées en mars et avril entre 1996 et 2000. La collection recouvre des domaines ayant une extension en .fr ou propre à l'outremer, mais également des domaines avec une extension générique (.com, .org, .net, .edu, .int) qui peuvent contenir ou non des contenus français. Le volume de la collection s'élève à près de 573 millions d'URL représentant près de 2,5 To.
Consulter
En urgence, collecte sur les attentats parisiens de 2015
Cette collection contient les pages web capturées lors de deux collectes d'urgence menées en 2015 afin d'obtenir des échantillons de réactions en ligne. La première capture est relative aux attentats des 7, 8 et 9 janvier et de la marche citoyenne du 11 janvier 2015 en France ; 1 581 domaines différents ont été moissonnés du 8 au 16 janvier, pour un total d'environ 13,5 millions d'URL collectées et un volume de 0,5 To. La seconde capture est relative aux attentats du 13 novembre 2015. Moins conséquente, elle couvre 61 URL de départ du 16 au 20 novembre et représente environ 1,2 million d'URL collectées pour un volume de 65 Go.
Consulter
Toutes les collections
Cette entrée permet d'accéder à l'application Archives de l'internet et de consulter la totalité des archives constituées par le BnF au titre du dépôt légal. Une recherche par URL sur l'ensemble des collections et des parcours guidés thématiques y sont notamment proposés.
Consulter
© BnF 2016 | Informations légales | ailabs 1.2.0 (pfo)
FR
10:28
29/11/2016

©BnF

BnF Archives de l'internet Labs

COLLECTIONS MON COMPTE Aide A propos

Les incunables du web français 1996-2000

Statistiques

Des statistiques détaillées décrivent l'ensemble de la collection. Elles sont fournies sous forme de chiffres (nombre d'URL, de domaines, de formats, poids...) et de listes (répartition des URL par extension, 200 plus gros domaines...) et permettent d'appréhender la collection dans sa totalité.

Nom du fichier	Description	Date de mise à jour	Poids en octets	Télécharger
indicateurs-generaux-arc.txt	Principales statistiques sur le nombre d'URL ou de domaines, sur le poids, sur les formats...	16/06/2016	565	télécharger
tld-domain.txt	Répartition des 2 092 398 domaines collectés dans les 19 extensions (TLD)	16/06/2016	157	télécharger
tld-url.txt	Répartition des 573 202 794 URL collectées dans les 19 extensions (TLD)	16/06/2016	188	télécharger
domaine-url-top200.txt	Nombre d'URL collectées pour les 200 plus grands domaines	16/06/2016	3 851	télécharger
domaine-url_fr-top200.txt	Nombre d'URL collectées pour les 200 plus grands domaines en .fr	16/06/2016	3 304	télécharger
hote-url-top200.txt	Nombre d'URL collectées pour les 200 plus grands hôtes	16/06/2016	4 688	télécharger
annee-url.txt	Répartition des 573 202 794 URL collectées par année (de 1996 à 2000)	16/06/2016	201	télécharger
codehttp-url.txt	Répartition des 573 202 794 URL collectées dans les 52 codes HTTP	16/06/2016	458	télécharger
typemime-url.txt	Répartition des 573 202 794 URL collectées dans les 4 498 formats (type MIME)	16/06/2016	225 678	télécharger

© BnF 2016 | informations légales | ailabs 1.2.0 (pfs)

Archives de l'Internet Labs
©BnF

BnF Archives de l'internet Labs

COLLECTIONS MON COMPTE Aide A propos

En urgence, collecte sur les attentats parisiens de 2015

Modifier la recherche | Nouvelle recherche

36 882 résultats

Mot(s) recherché(s) : "charlie hebdo"
Hôte, Domain ou extension : twitter.com
Enregistrer la recherche
Exporter les résultats

Trier par : Date de collecte | Croissant | << < Page 1 sur 3 689 > >> | 10 résultats par page

Enregistrer la sélection (0)

Trier les valeurs des facettes : 0-9 | A-Z

Année (1) inclure | exclure
+ 2015 (36 882)

Nom de domaine (1) inclure | exclure
+ twitter.com (36 882)

Extension (1) inclure | exclure
+ com (36 882)

Langue (10+) inclure | exclure
+ fr (14 905)
+ en (8 230)
+ bn (1 281)
+ es (833)
+ tl (603)
+ it (568)
+ cs (553)
+ nl (518)
+ vi (452)

1 "Conseil général 90 (@cg90_fr) | Twitter"

Archive du 08 janvier 2015
Format : html - Pertinence : 0.17881839
https://twitter.com/cg90_fr

Belfort @bleubelfort · 4 h il y a 4 heures #Belfort : la place d'Armes noire de monde.Hommage aux 12 morts de **Charlie Hebdo** #jesuischarlie pic.twitter.com/Miir2tdNiz Voir plus de photos et vidéos 0

2 "AMF (@AMF_actu) | Twitter"

Archive du 08 janvier 2015
Format : html - Pertinence : 0.17881839
https://twitter.com/AMF_actu

y a 6 heures Les salariés de l'#AMF se réuniront à 12h pour 1 min de silence en mémoire des victimes de l'attentat contre **Charlie Hebdo** #JeSuisCharlie 0 réponse 0 Retweet 3 favoris Répondre Retweeter

3 "Pays de la Loire (@paysdelaloire) | Twitter"

Archive du 08 janvier 2015
Format : html - Pertinence : 0.25288737
https://twitter.com/paysdelaloire

©BnF

```
WARC-Target-URI: http://www.univ-
angers.fr:80/universite/partenariat.html
WARC-Date: 1998-05-26T12:14:38Z
WARC-Record-ID: <urn:uuid:013e90e8-10ab-44d9-a0f6-
4db2fc436209>
WARC-Refers-To: <urn:arc:BNF-FS-478076.arc.gz:154>
Content-Type: application/json
Content-Length: 1716

{"Envelope":{"Format":"ARC","ARC-Header-Metadata":
{"Date":"19980526121438","Content-Length":"2361","Content-
Type":"text/html","Target-URI":"http://www.univ-
angers.fr:80/universite/partenariat.html","IP-
Address":"193.49.144.1"},"ARC-Header-Length":"100","Payload-
Metadata":{"Trailing-Slop-Length":"1","Actual-Content-
Type":"application/http; msgtype=response","HTTP-Response-
Metadata":{"Headers":{"ETag":"\"220a6-85c-31c02d7b
\"","Date":"Tue, 26 May 1998 12:26:01 GMT","Content-
Length":"2140","Last-Modified":"Thu, 13 Jun 1996 15:02:19
GMT","Content-Type":"text/html","Accept-
Ranges":"bytes","Server":"Apache/1.2.5"},"Headers-
Length":"221","Entity-Trailing-Slop-Length":"0","Entity-
Length":"2140","Response-Message":
{"Status":"200","Version":"HTTP/1.1","Reason":"OK"},"HTML-
Metadata":{"Links":
[{"path":"IMG@/src","url":"/sImages/up.gif"},
{"path":"A@/href","url":"index.html"},
{"path":"IMG@/src","url":"/linebgy.gif"},{"text":"Jacques
Allo","path":"A@/href","url":"/~allo/allo.html"},
{"text":"Jacques Allo","path":"A@/href","url":"/
~allo/allo.html"},{"text":"webadmin@univ-
angers.fr","path":"A@/href","url":"mailto:webadmin@univ-
angers.fr"},{"path":"IMG@/src","url":"/Images/logoicon.gif"},
{"text":"retour &agrave; la page d'accueil de
l'Universit&eacute;","path":"A@/href","url":"index.html"}],"He
ad":{"Title":"Universit&eacute; d'Angers -
PARTENARIAT"}},"Entity-
Digest":"sha1:W3E1JMCYCGK4BNTJA6F5KKIV63QAB675"},"Block-
Digest":"sha1:YUHLPNEPFUARCPW2KKY25Y2ZHBYKYGNK","Actual-
Content-Length":"2361"}},"Container":{"Compressed":true,"Gzip-
Metadata":{"Footer-Length":"8","Deflate-
Length":"1321","Header-Length":"10","Inflated-CRC":"-
1576703488","Inflated-
Length":"2462"},"Offset":"154","Filename":"BNF-FS-
478076.arc.gz"}}

WARC/1.0
WARC-Type: metadata
WARC-Target-URI: http://www.univ-tln.fr:80/enseignement/aceco/
```

MEDIUM READING

- As noted by Jane Winters (‘Web Archives for Humanities Research, in Brügger N., Schroeder R. (ed.) *The Web as History*, UCL Press, 2017: 239)

‘For most humanities scholars it will be a very long time before they transition to using solely digital sources, let alone solely born- digital sources, and for many this will never be the case. They will continue to mix and match, to compare and contrast, and to work with overlapping sets of material’.