



Laboratoire
des usages
en technologies
d'information
numériques

Cité des sciences et de l'industrie

UMS CNRS n° 2809

<http://www.lutin.utc.fr>

Directeur: Dominique Boullier



PARIS 8



AMCSTI

WebCSTI

Culture scientifique et technique sur le Web :

mesures des usages
et dynamique d'un champ

Appel d'offres « usages de l'internet »
Ministère de la Recherche





Objectifs

- Décrire les usages des sites de la culture scientifique et technique en France
- Permettre de repérer les stratégies et d'orienter les politiques éditoriales
- Méthode applicable à tout domaine institutionnel, commercial ou industriel comme **outil de pilotage stratégique et d'intelligence économique**



Présupposés

- Tout domaine
 - possède une existence autonome sur le Web,
 - révèle des acteurs nouveaux,
 - permet des stratégies institutionnelles ou commerciales nouvelles
- Les niveaux d'observation doivent être séparés mais en même temps combinés:
Effets de masse et pratiques individuelles
- Les usages ne peuvent être analysés sans référence aux propriétés de l'offre: toujours une double analyse

• Une stratégie de recherche plurielle SHS-STIC

<i>Plan de travail Web CSTI</i>	<i>Quantitatif Automatisation du recueil de données en masse</i>	<i>Qualitatif (et quanti) Sites et pratiques individuelles</i>
<i>Offre (les sites)</i>	1-Analyse topologique du Web à partir de robots	3- Analyse sémiotique des sites et découpage des écrans
<i>Demande (les pratiques des usagers)</i>	2- Analyse des parcours sur sites à partir des serveurs	4- Analyse oculométrique des parcours individuels

1- Analyse topologique du Web de la CSTI (méthode Tarent- RTGI- UTC)

Couche **la plus visible** du web :
Google, Amazon,
Voila, SNCF, etc...

Couche
intermédiaire :
agrégats,
communautés
en ligne

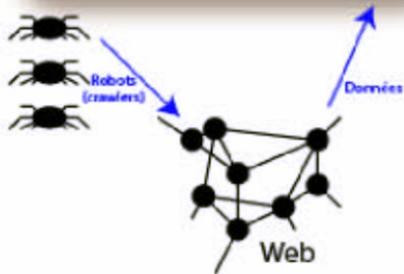
Couche profonde :
bases de données



1- Analyse topologique du Web de la CSTI (méthode Tarent- RTGI- UTC)

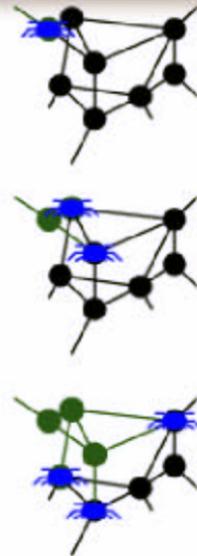
Crawl :

- Envoi de robots sur le web
- Rapatriement des données



Parcours des robots :

- Chaque robot commence à un point d'entrée et suit tous les liens



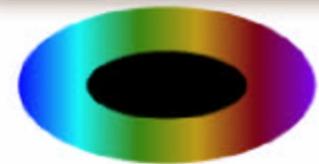
Algorithmes :

- Les sites sont analysés en fonction de leurs liens, pour montrer les regroupements

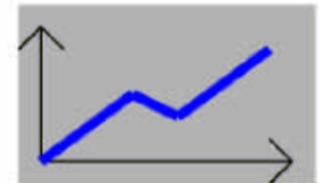


Résultats :

- On dégage la structure
- Les principaux sites sont analysés à la main



Structure du domaine



Données chiffrées



Faits Saillants

1-Analyse topologique du Web de la CSTI

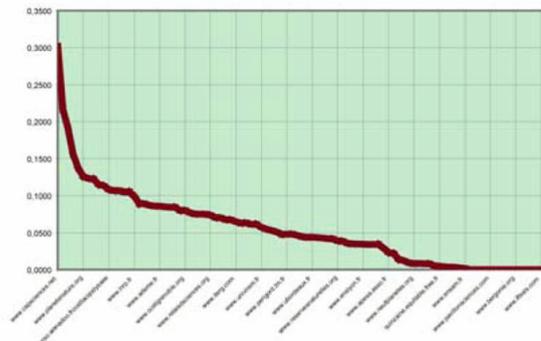
Score d'Autorité par sites,
par ordre décroissant



Liste des sites appartenant au coeur,
par score d'Autorité décroissant :

0,3037	www.cnrs.fr	0,0523	www.cybersciences.com
0,1851	www.citesciences.fr	0,0523	www.planetenature.org
0,1506	www.mnhn.fr	0,0516	www.francescience.org
0,1485	www.cnam.fr	0,0486	www.physique2005.org
0,1388	www.ubordeaux.fr	0,0485	www.ctba.fr
0,1377	www.ubordeaux2.fr	0,0480	www.latecoere.com
0,1292	www.cnes.fr	0,0479	perso.wanadoo.fr.malnesciences
0,1201	www.palaisdecouverte.fr	0,0469	www.museum.nantes.fr
0,1189	www.ademe.fr	0,0468	www.csticorse.asso.fr
0,1186	www.ubordeaux1.fr	0,0463	www.scienceanimation.org
0,1129	www.inrp.fr	0,0455	www.ocean.asso.fr
0,1109	www.univpoitiers.fr	0,0454	www.mollat.com
0,1043	www.univpau.fr	0,0443	www.quaternet.fr
0,0958	www.pasteur.fr	0,0439	www.agrotecfrance.com
0,0947	www.infosciences.fr	0,0436	www.adera.fr
0,0923	www.citeespace.com	0,0433	www.cctib.org
0,0913	www.aquitaine.fr	0,0429	www.sciencespobordeaux.fr
0,0910	www.onera.fr	0,0429	www.bordeauxunitec.com
0,0894	www.radiofrance.fr	0,0424	www.museumlyon.org
0,0870	www.exploradome.com	0,0423	www.technovest.com
0,0854	www.acordeaux.fr	0,0422	perso.wanadoo.fr.csticlacodysee
0,0841	www.larecherche.fr	0,0419	www.crhonealpes.fr
0,0839	www.capsiences.net	0,0416	www.agropole.com
0,0830	www.anpe.fr	0,0414	www.iterg.com
0,0791	www.ubordeaux3.fr	0,0414	www.iftsils.com
0,0786	www.liberation.fr	0,0405	membres.lycos.fr.cstidelaval
0,0780	www.pourlascience.com	0,0387	www.cyberpapy.com
0,0740	www.nouvelobs.com	0,0382	www.franceinfo.com
0,0736	www.ocim.fr	0,0363	www.univmontp2.fr
0,0683	www.enscpb.fr	0,0362	www.cirasti.org
0,0678	www.illedefrance.fr	0,0359	www.apesa.asso.fr
0,0676	www.mairiebordeaux.fr	0,0350	www.caminteresse.fr
0,0674	www.perigord.tm.fr	0,0349	www.parcsnaturelregionaux.tm.fr
0,0657	www.lescale.net	0,0344	www.edf.fr
0,0653	www.ensam.fr	0,0344	www.talenceinnovation.fr

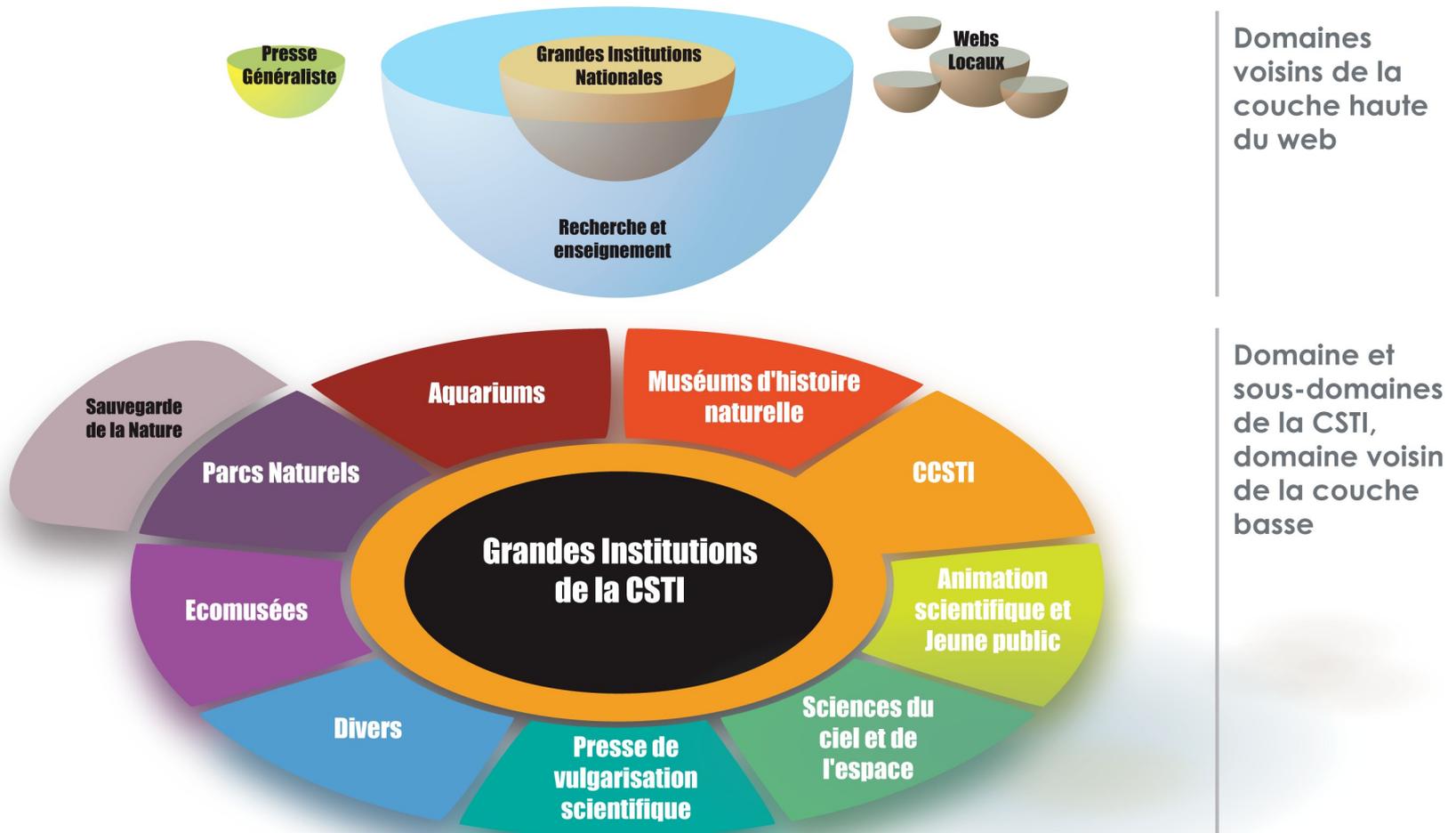
Score de Hub par sites,
par ordre décroissant



Liste des sites appartenant au coeur,
par score de Hub décroissant :

0,3017	www.capsiences.net	0,0435	www.agropole.com
0,2168	www.citesciences.fr	0,0429	www.infosciences.fr
0,1909	www.cnrs.fr	0,0420	www.technovest.com
0,1562	www.ocim.fr	0,0414	www.edf.fr
0,1371	www.ubordeaux1.fr	0,0392	www.reservesnaturelles.org
0,1259	www.planetenature.org	0,0384	www.nouvelobs.com
0,1234	www.bordeauxunitec.com	0,0358	www.agorasciences.org
0,1222	www.ubordeaux3.fr	0,0349	www.crrhonealpes.fr
0,1153	www.scienceanimation.org	0,0347	www.physique2005.org
0,1130	perso.wanadoo.fr.soclorient.csti	0,0345	www.enslyon.fr
0,1083	perso.wanadoo.fr.csticlacodysee	0,0343	www.cctib.org
0,1067	www.onera.fr	0,0342	www.illedefrance.fr
0,1066	www.sciencespobordeaux.fr	0,0342	www.univpau.fr
0,1052	www.labri.fr	0,0292	www.adera.fr
0,1051	www.larecherche.fr	0,0235	www.apesa.asso.fr
0,0990	www.inrp.fr	0,0217	www.enscpb.fr
0,0894	www.citedelamer.com	0,0140	www.cirasti.org
0,0890	www.cnes.fr	0,0119	www.systemic.fr
0,0872	www.estia.fr	0,0094	www.mollat.com
0,0862	www.pasteur.fr	0,0085	www.neufplanetes.org
0,0861	www.ademe.fr	0,0085	www.artsetmetiers.net
0,0850	www.ubordeaux2.fr	0,0082	www.cybersciences.com
0,0847	www.cnam.fr	0,0079	www.cyberpapy.com
0,0844	www.univmontp2.fr	0,0054	www.franceinfo.com
0,0803	membres.lycos.fr.cstidelaval	0,0048	quinzaine.equitable.free.fr
0,0799	www.cstigrenoble.org	0,0038	www.ocean.asso.fr
0,0770	www.francescience.org	0,0033	www.ecolesdesavignac.com
0,0754	www.momes.net	0,0031	www.enseirb.fr
0,0751	www.ouestfrance.fr	0,0024	www.agrotecfrance.com
0,0749	www.aquitaine.fr	0,0016	www.ensam.fr
0,0737	www.relaisciences.org	0,0000	www.talenceinnovation.fr
0,0708	www.acordeaux.fr	0,0000	www.amazon.fr
0,0699	www.pourlascience.com	0,0000	www.cieletespace.fr
0,0679	www.radiofrance.fr	0,0000	www.lescale.net
0,0674	www.parcsnaturelregionaux.tm.fr	0,0000	www.navillonsciences.com

1-Analyse topologique du Web de la CSTI



2- Analyse des parcours en masse (Méthode LIP6)

Traitement des logs de consultations de chaque page

Common Log Format :

```
127.0.0.1 - Wallace [10/Oct/2000:13:55:36 -0700] "GET /image.gif  
HTTP/1.0" 200 2326'
```

Combined Log Format :

```
127.0.0.1 - Wallace [10/Oct/2000:13:55:36 -0700] "GET /image.gif  
HTTP/1.0" 200 2326 "http://lofti.lip6.fr" "Mozilla/4.08 [en] (Win98; I  
;Nav) "
```

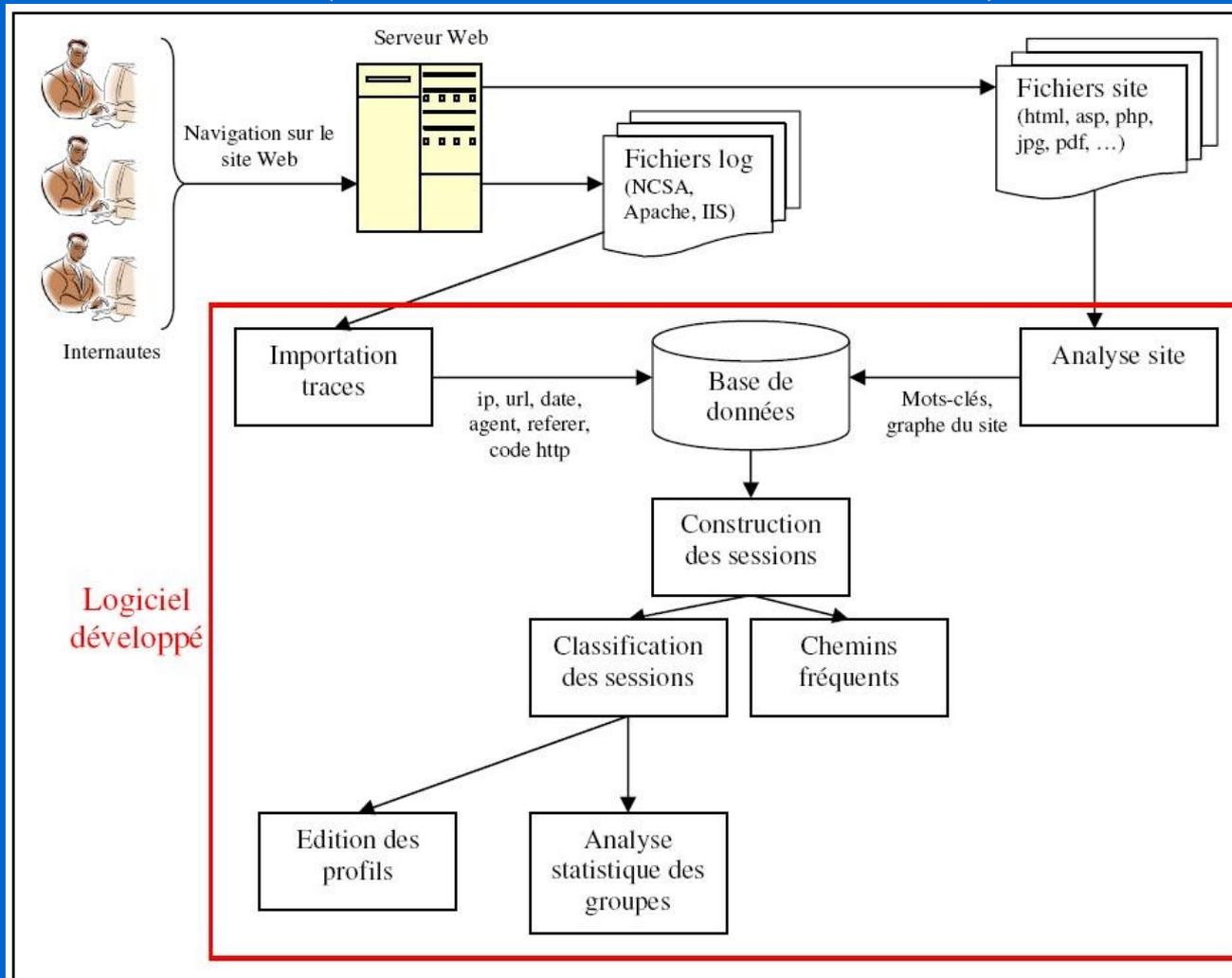
MicrosoftIISLogFormat :

```
127.0.0.1, Wallace,03/20/98, 7:55:20, SERVER, IPSERVER, 192.168.114.201,  
4502, 163, 3223, 200, 0, GET, image.gif
```

NCSA Log Format :

```
127.0.0.1 Wallace [08/Apr/1997:17:39:04 -0800] "GET /image.gif HTTP/1.0"  
200 3401'
```

2- Analyse des parcours en masse (Méthode LIP6)

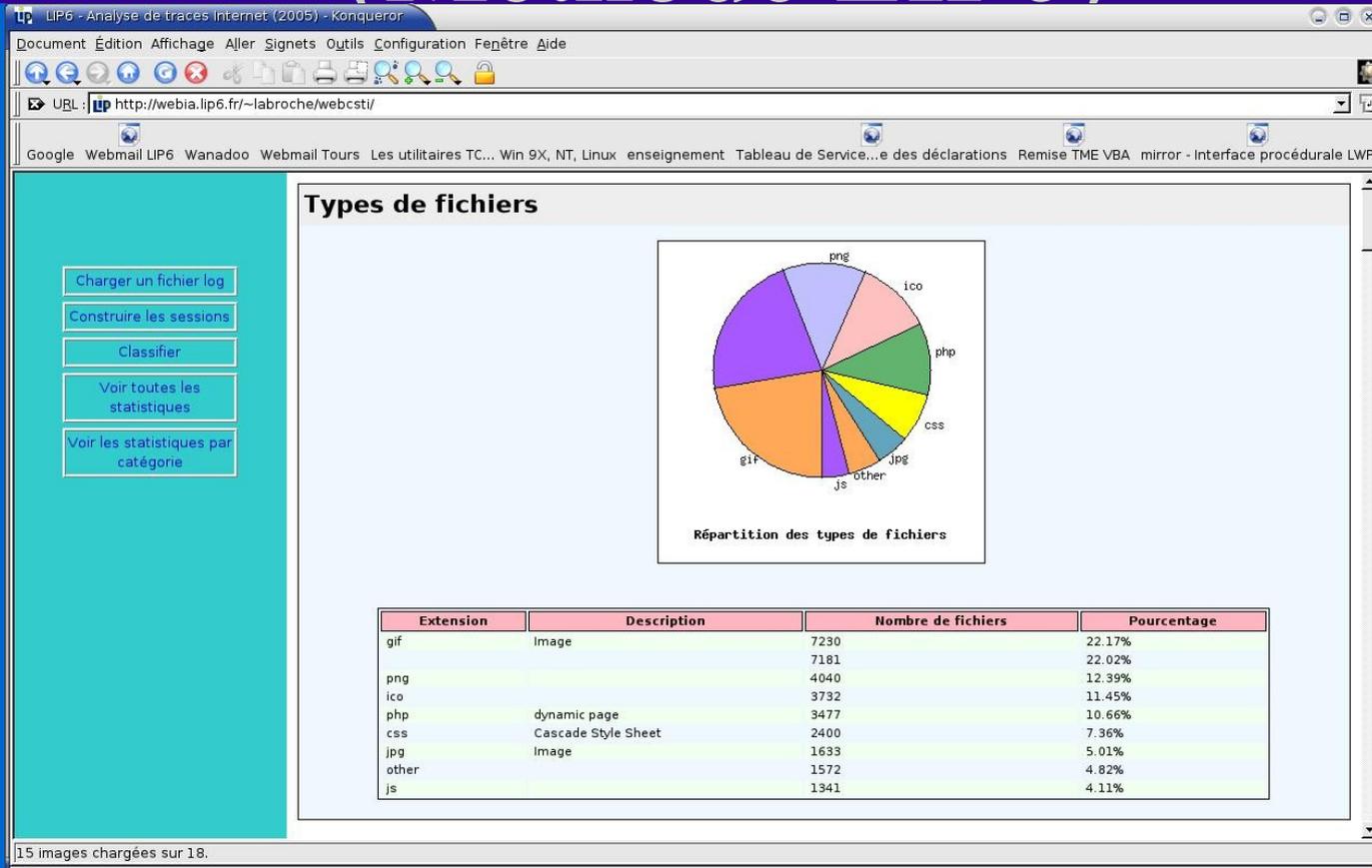


2- Analyse des parcours en masse (Méthode LIP6)

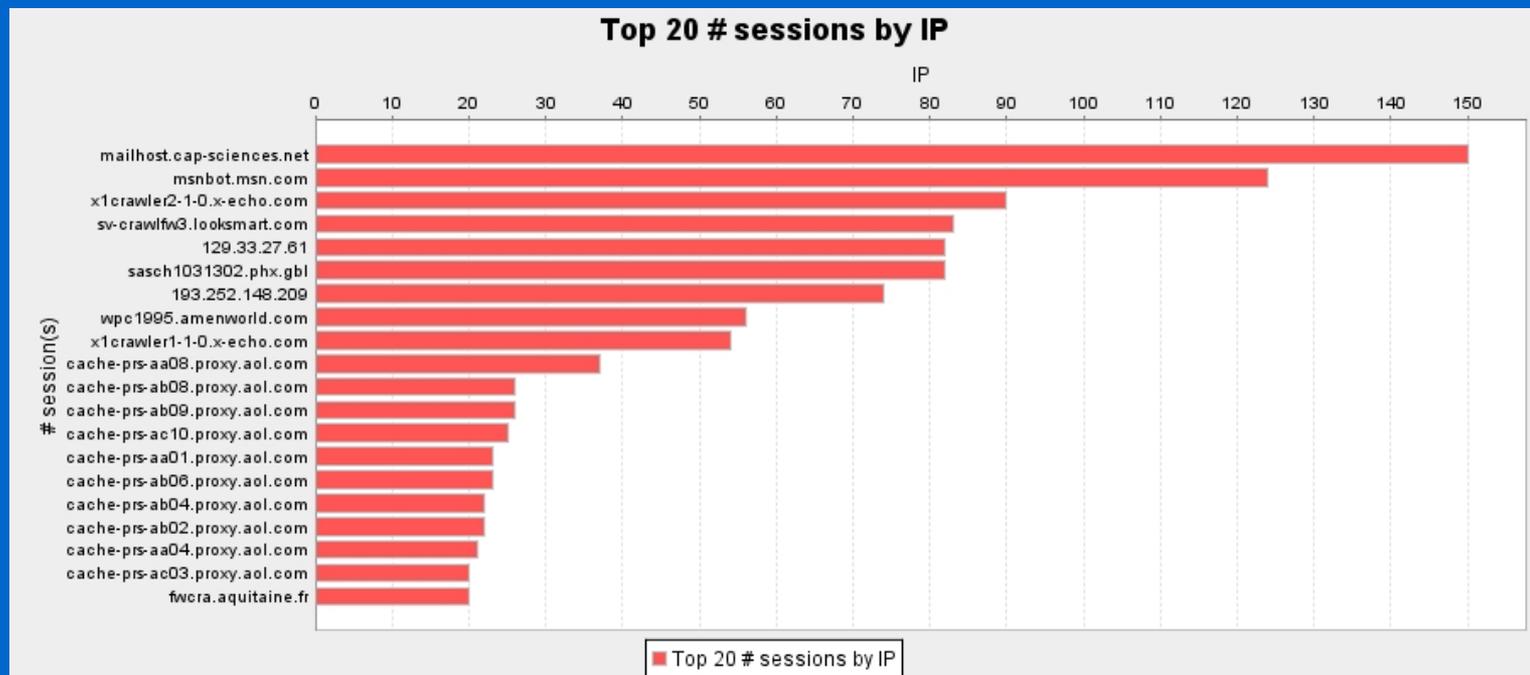


Graphe d'un parcours

2- Analyse des parcours en masse (Méthode LIP6)



2- Analyse des parcours en masse (Méthode LIP6)



4- Analyse des parcours oculaires des usagers

- Comprendre et mesurer exactement ce qui est perçu et faire le rapprochement avec les intentions de l'offre
- Scénarios de découverte clics sur les 10 sites sélectionnés

