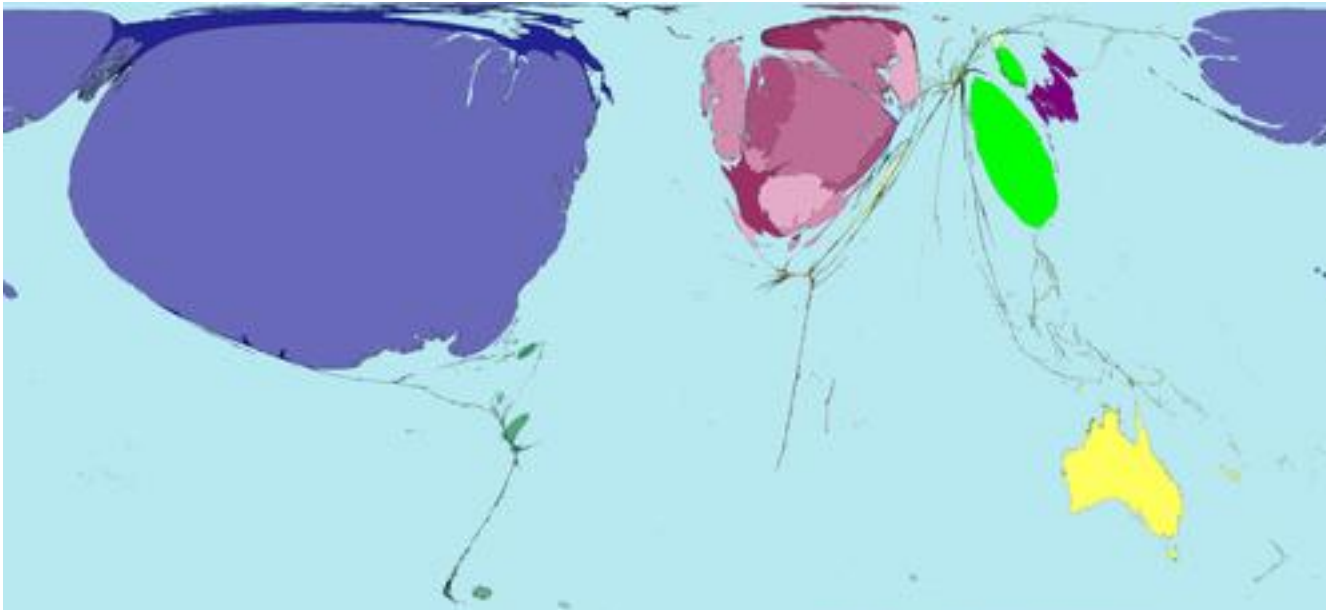


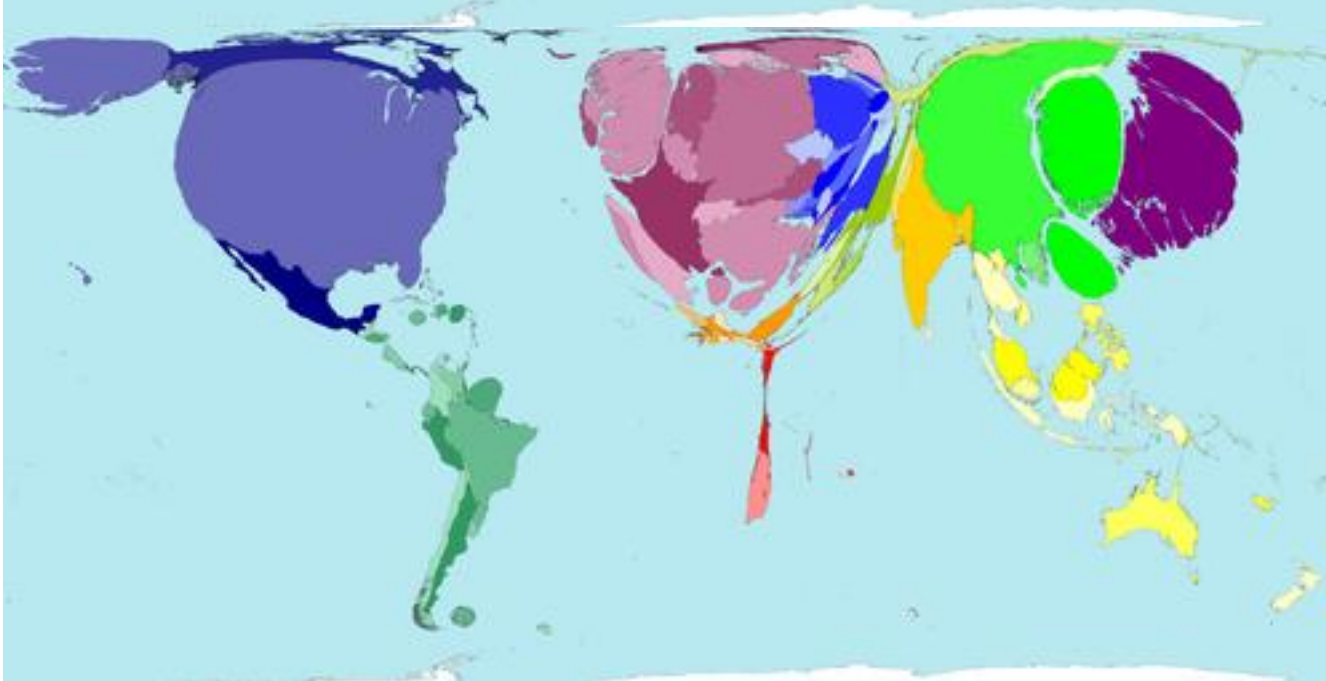
Ressources et réflexion sur les moteurs de recherche

Les utilisateurs
d'internet dans le
monde. Source
Worldmapper.

1990



2012



Il existe plusieurs types d'outils de recherche

- **Des moteurs de recherche** comme Google, Yahoo, Bing, Qwant (français), Baidu (chinois, 2^{ème} rang mondial) Yandex (russe, 4^{ème} rang), Exalead (groupe Dassault, français), DuckDuckGo.
Un moteur de recherche est une interface qui indexe de façon automatique les pages web par mots clés. Il est conçu pour scanner à intervalle régulier le web visible. Il n'accède pas au « deep web » protégé par des mots de passe ou des chemins très spécifiques.
- **Des métamoteurs** comme PolyMeta, Ixquick.
Un métamoteur est un moteur qui interroge simultanément plusieurs moteurs de recherche ainsi que certains grands sites. Il permet ainsi d'accéder à des réponses de plusieurs moteurs de recherche différents. Les résultats, plus variés, font apparaître des sites qui seraient restés invisibles par le biais de certains moteurs.
- **Des bases documentaires** de type BCDI.
Elles permettent d'accéder à une information triée par un logiciel de traitement des données. On accède à des fiches sur les articles et ouvrages, donc à des sources secondaires avant d'accéder à des sources primaires. Cela permet une utilisation raisonnée du CDI. Cet outil permet de présenter les mots clefs proches du sujet pour affiner ou étendre la recherche.

Une grande variété des moteurs de recherche

v · d · m	Moteurs de recherche	[masquer]
Généraux	A9.com · ASPseek · AlltheWeb · AltaVista · Apache Solr · Ask.com · Baidu · Bing · Cuil · DataparkSearch · Dir.com · DuckDuckGo · Exalead · Gigablast · Google · Grub · Guruji · Heritrix · Hi-Algérie ! · Hotbot · Ht-//Dig · Ixquick · Junglekey · KartOO · Krozilo · Lucene · Lycos · mnoGoSearch · Mozbot · Naver · Panguso · Piria · PubGene · Qwant · Seeks · Sharelook · Soso · Technorati · Theseus · Verticrawl · WebCrawler · WolframAlpha · Xapian · YaCy · Yahoo! · Yandex · Yatedo · Yauba · Zettair	
Spécialisés	Publications scientifiques : CiteSeerX · Google Scholar · ScientificCommons Shopping : Kelkoo · NexTag · Skyscanner Emploi : Indeed · Jobboom · Keljob.com · Monster · Workopolis	
Solidaires	Doona (humanitaire) · Ecosia (écologique) · GoodSearch (humanitaire) · HooSeek (écologique)	
Abandonnés	Accoona · Cuil · ElgooG · Lokace · Wikiseek	
Autres	Astalavista (sécurité informatique) · BabyGo (pour enfants) · Gameskoot (jeux vidéo) · IceRocket (blogs) · Spock (personne)	



Portail d'Internet

Catégories : [Moteur de recherche](#) | [Site web norvégien](#) | [Site web disparu](#) | [\[+\]](#)

Capture d'écran, Wikipédia, 13 décembre 2013

Enjeux des moteurs de recherche et des métamoteurs

Enjeux autour de cet outil.

Outil : enjeux de puissance : dimensions géopolitiques et politiques de ces technologies et de leur maîtrise.

Qui maîtrise les contenus accessibles ?

Une information qui n'est pas contenue dans les 5 premières pages a peu de chance d'être lue. Enjeux pour les différents acteurs de l'écosystème numérique.

Enjeux : confidentialité des données personnelles.

Questions de cybersécurité. Les débats se multiplient.

Enjeux pour l'élève

Trouver l'information recherchée dans la masse d'informations.

Apprendre à chercher.

Mener une recherche sûre, efficace et pertinente.

Prendre conscience des enjeux précédents.

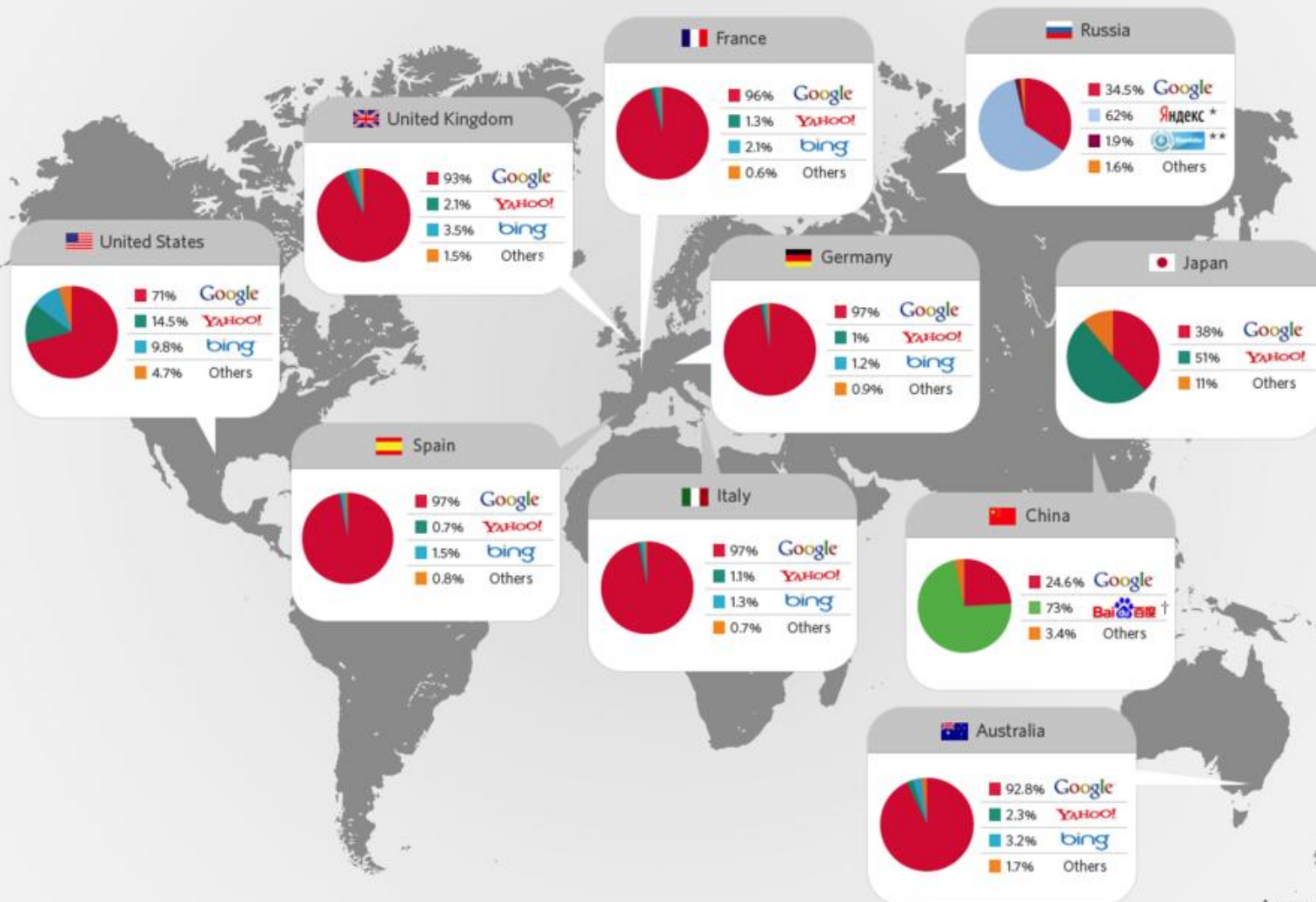
Enjeux pour le professeur

Amener les élèves à progresser dans la maîtrise de la recherche documentaire et donc de l'outil.

Développer une culture du numérique.

Développer l'esprit critique dans l'utilisation des outils.

Search engine market shares around the world - Q4 2010



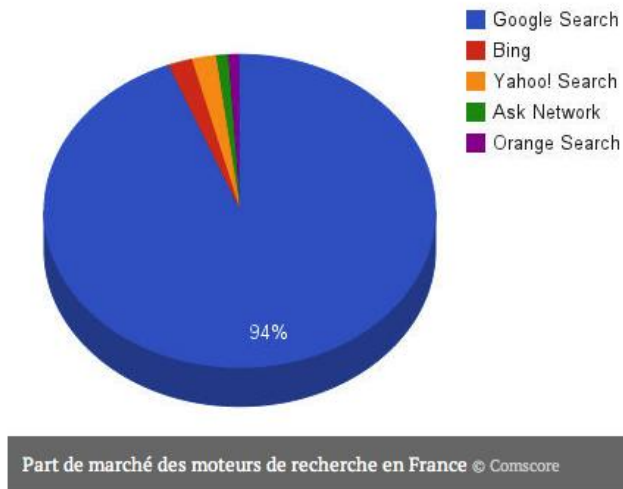
Part de marché des moteurs de recherche dans le monde en 2010

Source : Greenlight

Parts de marché des moteurs de recherche en France

La part de marché des moteurs de recherche en France en décembre 2012.

Selon ComScore, en France, 94% des requêtes effectuées sur des moteurs de recherche passent par Google. Bing et Yahoo! représentent chacun 2% de parts de marché dans l'Hexagone. Ask Network et Orange Search ne sont utilisés chacun que dans 1% des requêtes.








» [Tous les chiffres clés de l'Internet](#)

Un constat : un marché dominé par Google.

Objectif réaliste : apprendre à maîtriser le mieux possible la recherche sur les principaux moteurs en particulier Google.

Présenter des alternatives, leurs atouts, leurs limites.

Moteurs de recherche	Particularités	Fonctionnalités	Qualités, atouts	Défauts, limites
Google 	Moteur de recherche américain, siège social à Mountain View. Créé en 1997. Leader mondial. Leader en France.	Capacité de tri exceptionnelle. Curation améliore résultats.	Bon degré de pertinence. Le meilleur moteur à ce jour, le plus rapide. Outils de recherche avancée permettent d'affiner la recherche. 112 langues.	Règles de référencement obscures. A la fois Moteur de recherche et moteur de positionnement. Affichage contingenté par page. Utilisation des données personnelles : Chaque clic est enregistré. Google traque tous ses utilisateurs.
Qwant 	Présenté comme un moteur de recherche français	Affichage des sites par image. Nationalité de l'article.	Différents modes d'affichage. Résultats contextualisés. Rubriques thématiques. Affichage	En fait un agrégateur de résultats, qui pioche ses résultats chez Bing
Ixquick 	Métamoteur. Créé en 1998. Siège social Harleem, Pays-Bas. Crée Startpage en 2009 qui n'interroge que google.	Interroge simultanément 10 moteurs et sites web : Aol, Althetweb (qui redirige vers yahoo depuis 2011), Altavista, Ask/Teoma, Bing, Gigablast, Google, Open directory, Wikipédia. En 18 langues.	Respect de la vie privée. Affirme ne conserver aucune trace des recherches, même pas l'adresse ip. pertinent.	Peu connu
Yahoo 	Moteur de recherche américain. 1995. Siège : Santa Clara	Recherche multilingue et traduction automatique en plusieurs langues. Subdivision des sites par arborescence.	Degré de pertinence élevé. Proche de google. Donne aussi des réponses en image ou vidéo.	Un affichage chargé, parfois confus.
Bing 	Moteur de recherche Microsoft, 2009	Rubriques thématiques Essaie de tester la curation manuelle avec groupe d'experts.	Moteur puissant.	Moins pertinent que google dans sa version française.

Malgré la très grande **variété** des moteurs le secteur est **hyperconcentré et dominé par quelques géants**. (Cf part de marché des moteurs de recherche). On se centre donc ici sur **google, le leader mondial**.

Les problèmes liés aux difficultés d'usage

Indexation mouvante : d'un moment à l'autre ou d'une adresse ip à l'autre l'indexation change. Donc il y a différentes hiérarchisations des sites en fonction du poste informatique utilisé.

En effet Google traque vos clics et adapte sa réponse à l'utilisateur constituant ainsi ce qu'on nomme « **sa bulle** » (à savoir ce qui est supposé par google être ses centres d'intérêt compte tenu de ses recherches précédentes).

Accès à des documents primaires : question de la fiabilité n'est pas assurée, pas de résumé, pas de commentaires sur la pertinence du document. Demande de l'esprit critique.

Recherche trop rapide ? 75 % des internautes ne dépassent pas les 5 premières pages. Pour beaucoup d'élèves ils ne dépassent pas même les 5 premières références de la première page. Google trie 30 trillions de documents. (30 000 milliards) et traite 3,3 milliards requête /j.

Les difficultés liés aux modèles économiques

Le modèle économique est basé essentiellement sur la **publicité** (92 % des recettes de Google). Un moteur de recherche organise l'échange d'informations. Il échange des informations sur les lectures contre de l'informations sur les lecteurs. Puis le moteur échange les informations sur les lecteurs contre de la publicité. Les entreprises peuvent aussi acheter une position visible dans **le référencement**. Les règles dans ce domaine restent obscures. Le véritable enjeu des moteurs de recherches réside dans la diffusion des contenus et des très nombreux services qui leurs sont associés (clouds, agendas, actualités, enseignement à distance etc..). Plus le moteur est utilisé, plus il s'enrichit.

Les enjeux économiques sont très importants :

Par exemple pour la presse et les entreprises du numérique.

Pour **la presse**, parce que Google utilise les informations, les articles, pour les afficher dans certaines de ses applications, sans passer par les sites web des auteurs, ce qui est un manque à gagner en matière de recettes publicitaires. C'est pour cette raison que les journaux européens, comme Le Monde, font pression sur les gouvernements pour instaurer une « taxe Google ».

Pour **les entreprises du numérique** ensuite, parce que dans les recherches spécifiques à un service (vidéo, mail, stockage de données par exemple), Google propose en premier lieu sa propre application dans ce type de service (youtube, gmail, google drive etc). Les start-up qui développent ce genre d'application ont donc une visibilité très déséquilibrée vis à vis de Google.

En somme Google dispose avec son gigantisme d'un véritable droit de vie ou de mort sur une grande partie des acteurs de l'écosystème numérique.

Les difficultés matérielles

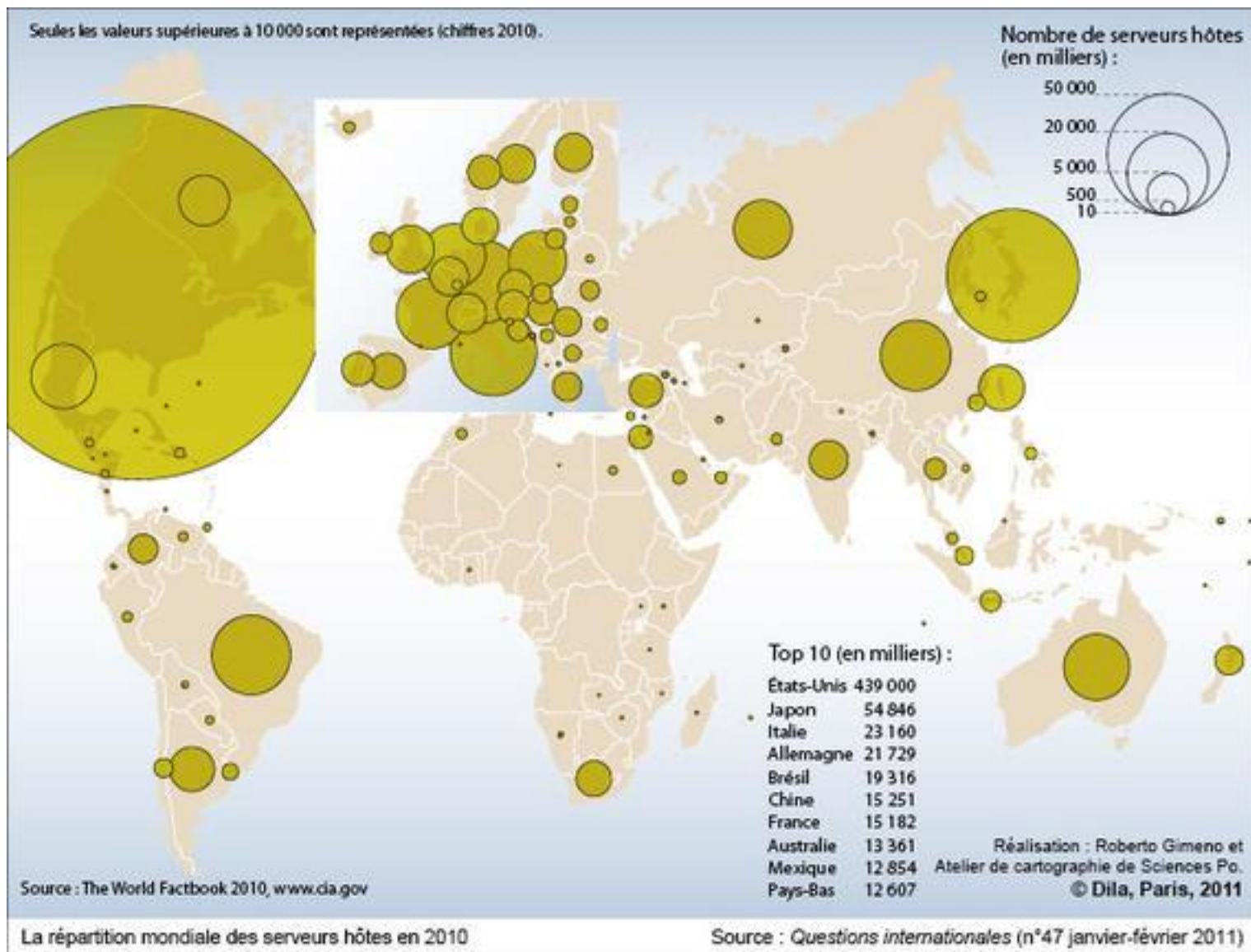
L'information est stockée matériellement dans des lieux avec un ancrage territorial fort : les centres de traitement des données, les data center. Ils concentrent équipements électroniques, ordinateurs, systèmes de stockage et équipements de télécommunications. Ils abritent et stockent les bases de données de quantités d'entreprises et bien sûr celles de Google même.

Certains de ces centres consomment plus d'électricité qu'une ville de 100 000 habitants et nécessitent des systèmes de refroidissement et de climatisation performants. Le maintien d'une température de 20 degrés est essentielle pour éviter que les appareils électriques qui génèrent beaucoup de chaleur ne surchauffent.

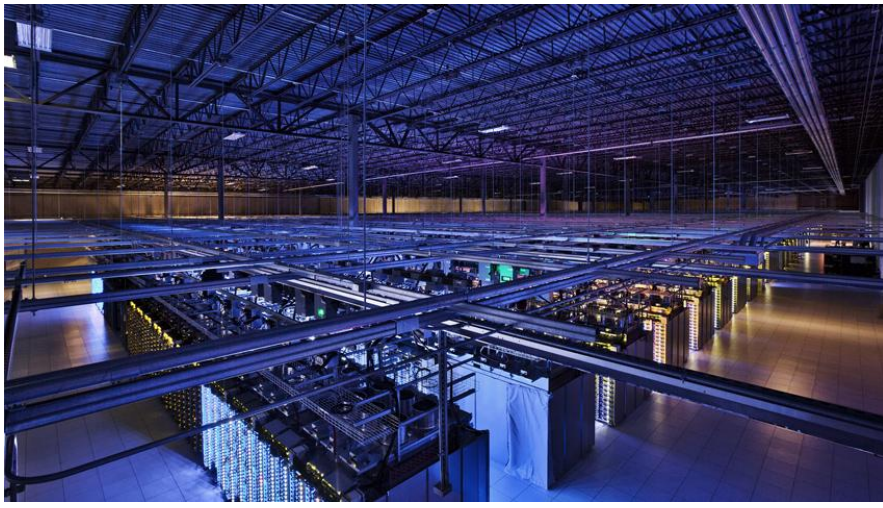
La localisation de ces usines numériques est devenue un enjeu crucial avec le nouvel impératif de la souveraineté des données : choisir où les données se trouvent, les transférer d'un pays à l'autre.

Google dissémine ses data center à travers le monde. Leur localisation est parfois secrète. Les bâtiments s'inscrivent, malgré leur taille, de façon souvent la plus discrète possible dans les territoires. Ils sont toujours étroitement surveillés pour des raisons de sécurité (clôtures électrifiées, services de sécurité 24h/ 24, accès très restreint, contrôle de la poussière, extinction automatique des incendies...). L'entreprise posséderait à elle seule plus d'un million de serveurs et 36 centres de données au moins. Mais ces chiffres restent sujet à discussion.

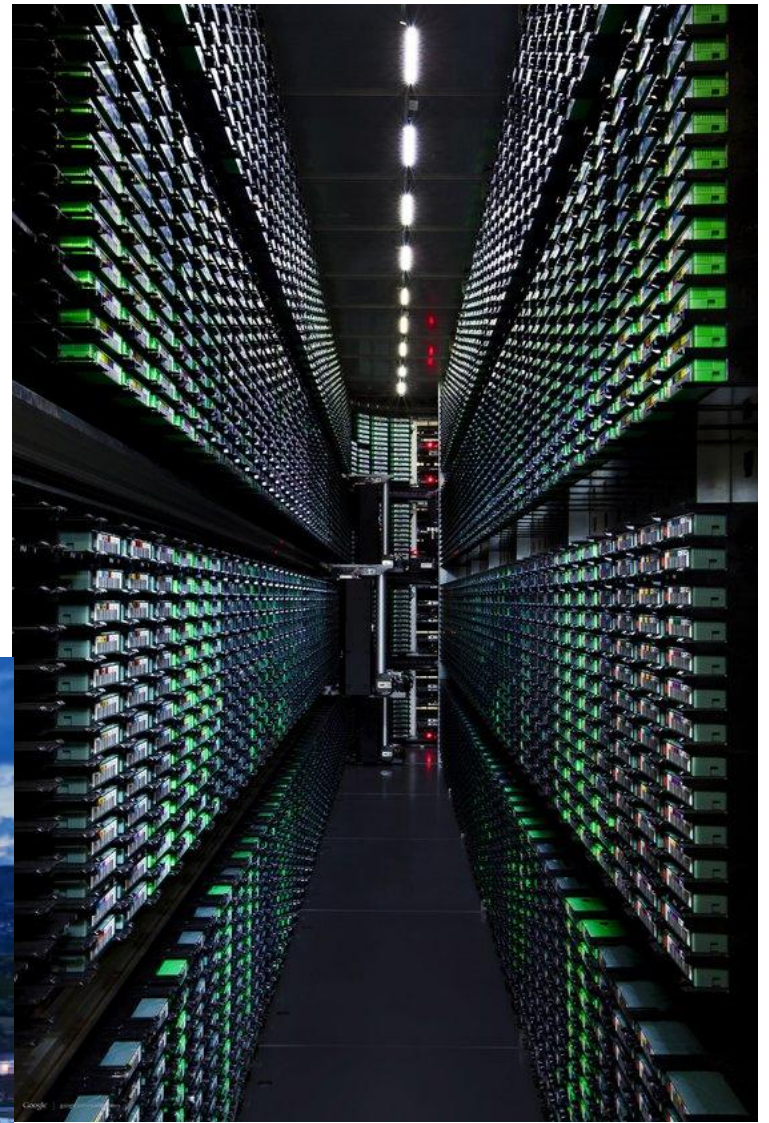
La répartition mondiale des serveurs hôtes en 2010



Source : Questions internationales : internet à la conquête du monde, La Documentation française, n°47, Paris, janvier /février 2011



Clubic Pro, Data Center, Google montre à quoi ressemble ses entrailles, 17 octobre 2012

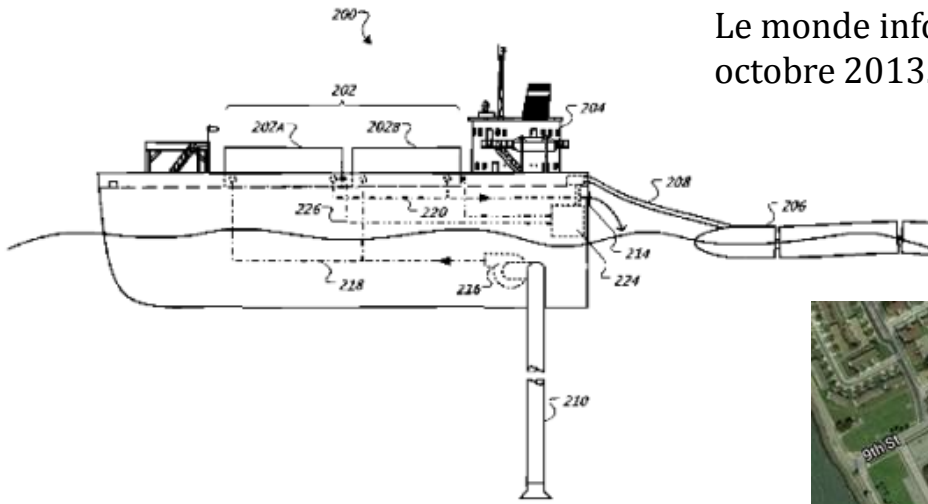




Les Data center de Google. Photos Google, Octobre 2012, Photos données à la presse par le centre de données de Google.

<http://www.google.com/about/datacenters/gallery/#/all>

Le monde informatique, Un data center flottant chez Google, 28 octobre 2013.



Objectif : un système énergétiquement autonome.



Inhotep, le blog innovation, numérique, Google serait en train de construire un data center flottant à San Francisco, mardi 29 octobre 2013

Un projet qui finalement a été abandonné pour faire de ce site un espace de démonstration des nouvelles technologies de Google.