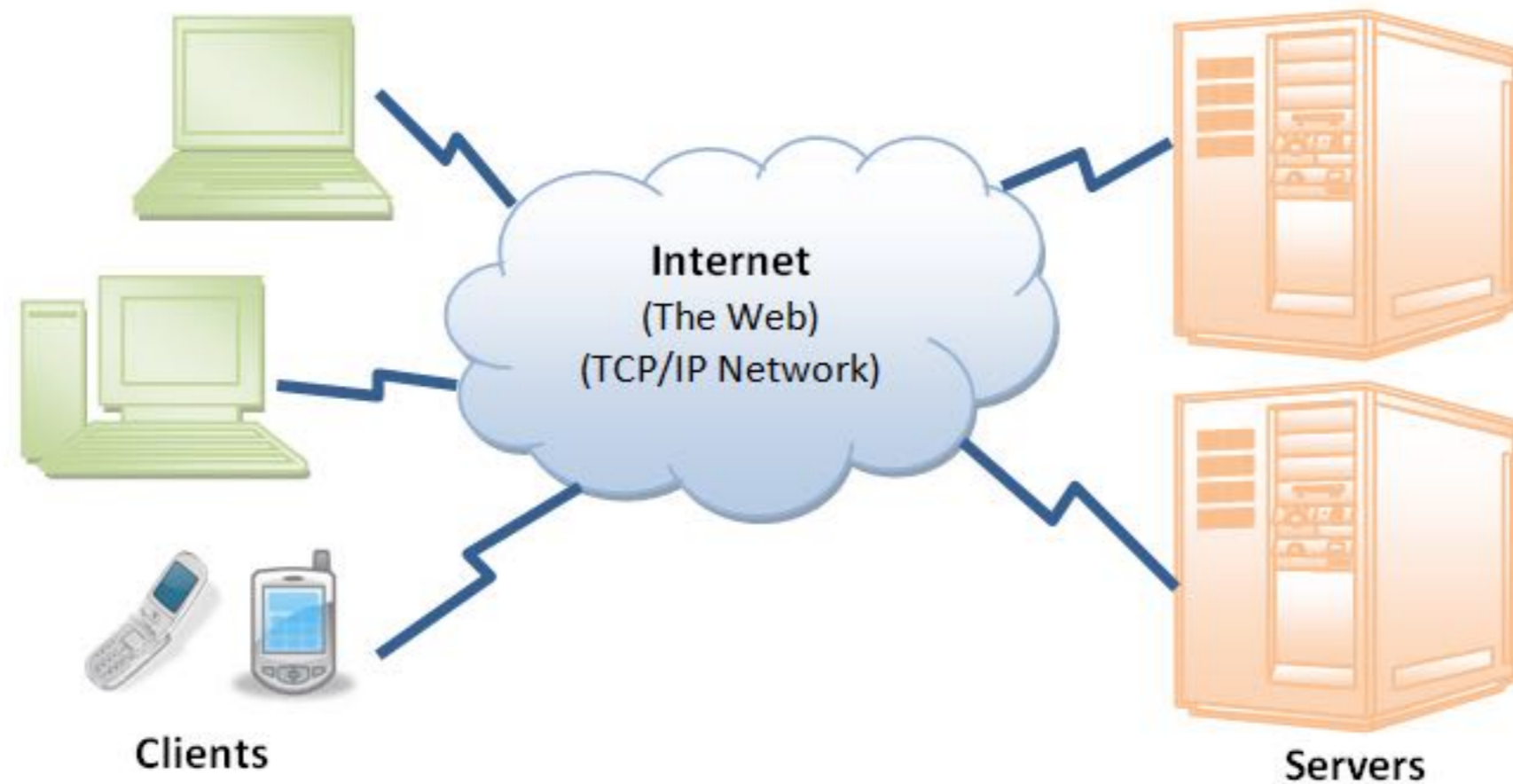


# IFT 1005

# Introduction à Internet

Guy Lapalme  
Université de Montréal

# Internet



# Différents protocoles

- Multiples protocoles
  - FTP: File transfer
  - SMTP, POP, IMAP : courriel
  - Telnet : accès à un ordinateur à distance
  - HTTP : web
- s'appuient tous sur
  - TCP/IP
  - Domain name service

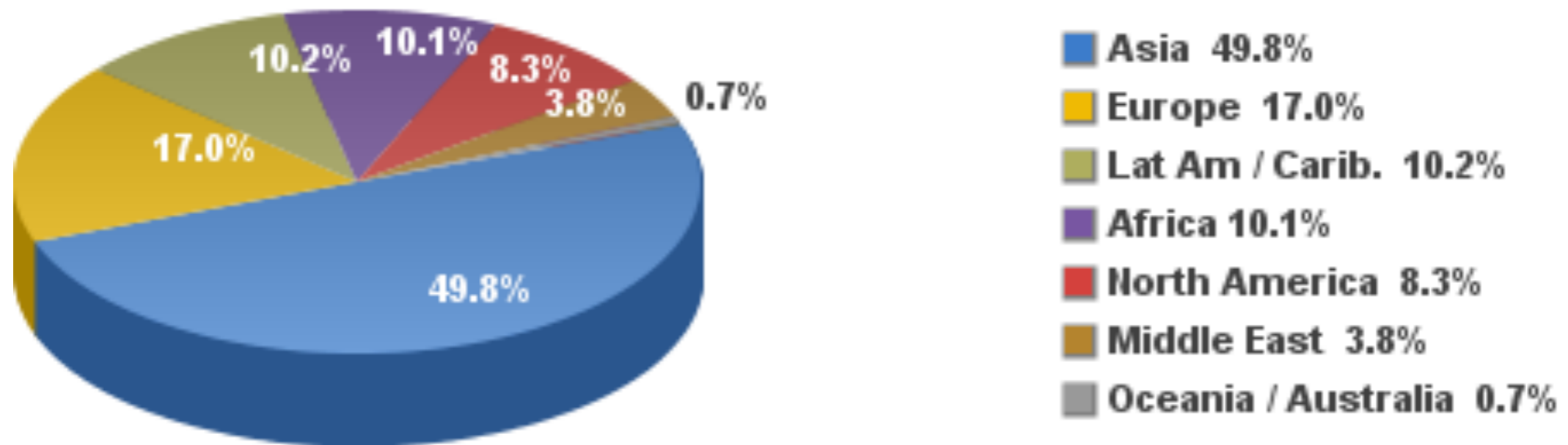
# Raisons de la popularité d'internet dans les années 90

- Permission d'avoir une activité commerciale sur internet
- Développement du World Wide Web
- Développement des browsers graphiques
- Ordinateurs personnels plus abordables
- Fournisseurs de services offrent des services à des prix *raisonnables*

# Popularité d'internet

<http://www.internetworldstats.com/stats.htm>

## Internet Users in the World by Regions - 2017 Q2



Source: Internet World Stats - [www.internetworldstats.com/stats.htm](http://www.internetworldstats.com/stats.htm)

Basis: 3,835,498,274 Internet users in June 30, 2017

Copyright © 2017, Miniwatts Marketing Group

# Popularité d'internet

<http://www.internetworldstats.com/stats.htm>

<b>WORLD INTERNET USAGE AND POPULATION STATISTICS</b>						
<b>JUNE 30, 2017 - Update</b>						
<b>World Regions</b>	<b>Population ( 2017 Est.)</b>	<b>Population % of World</b>	<b>Internet Users 30 June 2017</b>	<b>Penetration Rate (% Pop.)</b>	<b>Growth 2000-2017</b>	<b>Internet Users %</b>
<b><u>Africa</u></b>	1,246,504,865	16.6 %	<b>388,104,452</b>	31.1 %	8,497.0%	10.1 %
<b><u>Asia</u></b>	4,148,177,672	55.2 %	<b>1,909,408,707</b>	46.0 %	1,570.5%	49.8 %
<b><u>Europe</u></b>	822,710,362	10.9 %	<b>650,558,113</b>	79.1 %	519.0%	17.0 %
<b><u>Latin America / Caribbean</u></b>	647,604,645	8.6 %	<b>392,215,155</b>	60.6 %	2,070.7%	10.2 %
<b><u>Middle East</u></b>	250,327,574	3.3 %	<b>146,972,123</b>	58.7 %	4,374.3%	3.8 %
<b><u>North America</u></b>	363,224,006	4.8 %	<b>320,059,368</b>	88.1 %	196.1%	8.3 %
<b><u>Oceania / Australia</u></b>	40,479,846	0.5 %	<b>28,180,356</b>	69.6 %	269.8%	0.7 %
<b><u>WORLD TOTAL</u></b>	<b>7,519,028,970</b>	<b>100.0 %</b>	<b>3,835,498,274</b>	<b>51.0 %</b>	<b>962.5%</b>	<b>100.0 %</b>

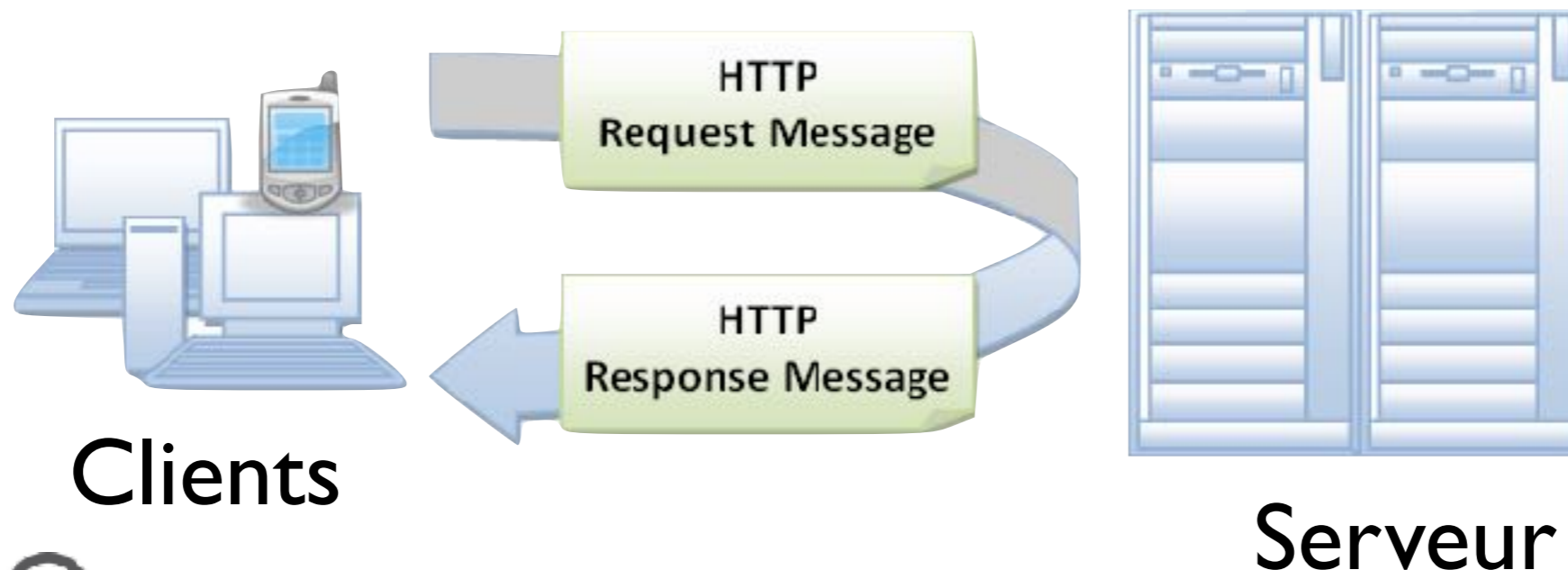
NOTES: (1) Internet Usage and World Population Statistics updated as of June 30, 2017. (2) CLICK on each world region name for detailed regional usage information. (3) Demographic (Population) numbers are based on data from the [United Nations Population Division](#). (4) Internet usage information comes from data published by [Nielsen Online](#), by ITU, the [International Telecommunications Union](#), by [GfK](#), by local ICT Regulators and other reliable sources. (5) For definitions, navigation help and disclaimers, please refer to the [Website Surfing Guide](#). (6) Information from this site may be cited, giving the due credit and placing a link back to [www.internetworldstats.com](http://www.internetworldstats.com). Copyright © 2017, Miniwatts Marketing Group. All rights reserved worldwide.

# HyperText Transfer Protocol - HTTP

initialement RFC 2616

maintenant RFC 7230 à RFC 7235

- protocole de transfert asymétrique (pull)  
**requête**/réponse ou **client**/serveur
- sans état
- permet la négociation des types de données et de leur représentation



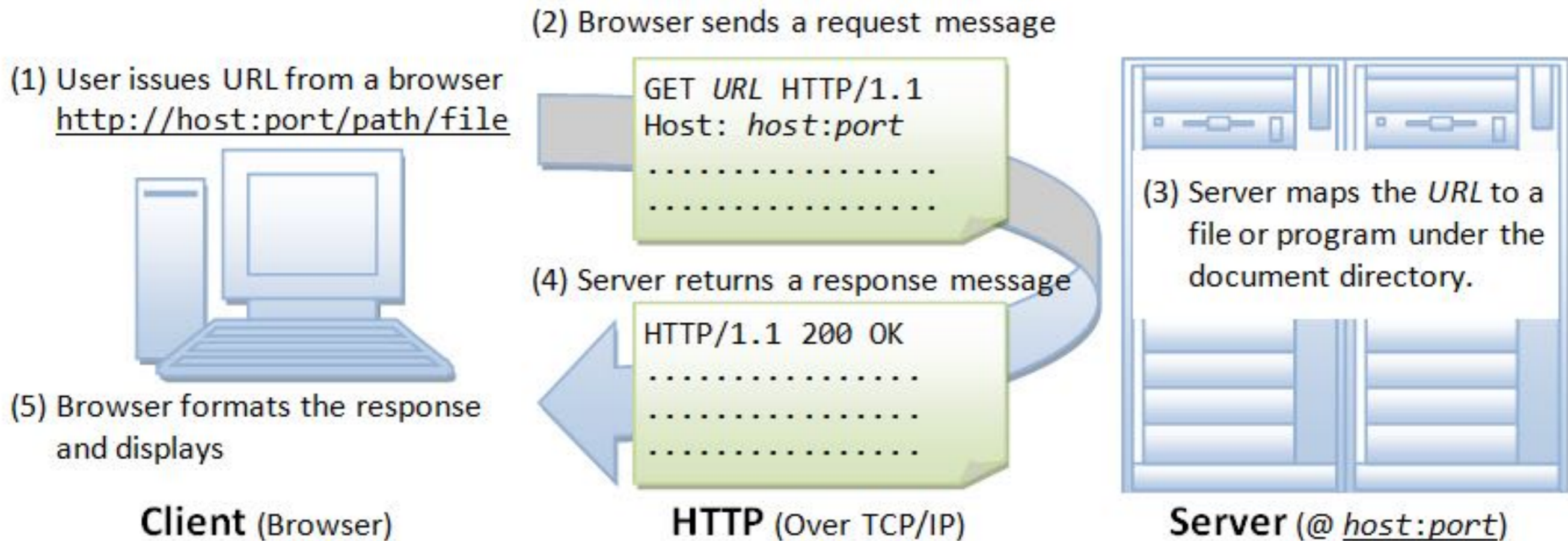
# Uniform Resource Locator - URL

scheme://host/path[?query-string][#frag]  
http://monsite.com/g1/test?prenom=guy&nom=lap#p1

- autres types d'url
  - ftp://ftp.is.ca.ca/rfc/rfc1808.txt
  - ldap://[2001.db8::7]/c=GB?objectClass?one
  - mailto:lapalme@iro.umontreal.ca
  - tel:+1-514-343-6111
  - telnet:132:204.24.179
  - isbn:978-2-1234-5680-3



# Browsers



- Cet échange est effectué pour chaque élément différent (e.g. image, script, css) d'une page web
- Une même page peut combiner des informations provenant de plusieurs serveurs

# World Wide Web Consortium (W3C)

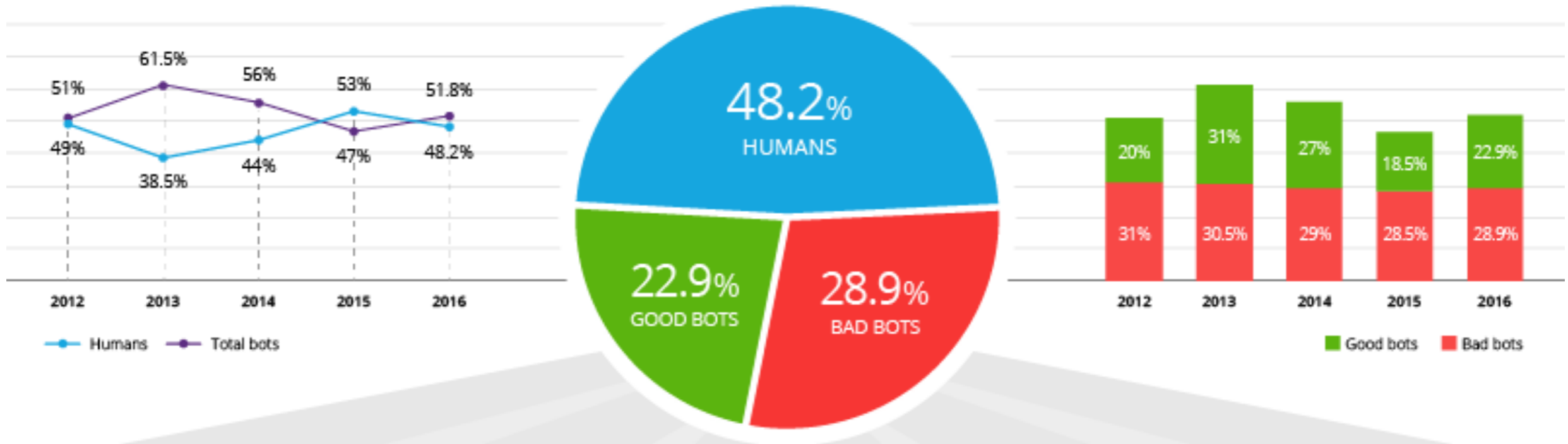
- Développe des spécifications (recommandations) Request For Comment
- Standardisation des échanges
- Web Accessibility Initiative (WAI)
  - accommoder les personnes avec handicaps visuels, auditifs ou physiques
- Web Content Accessibility Guidelines

# BOT TRAFFIC REPORT 2016

BOTS ONCE AGAIN COMPRISE THE MAJORITY OF ONLINE TRAFFIC AMID AN INCREASE IN GOOD BOT ACTIVITY.

BOT ACTIVITY IS IN AN UPTREND, after a three year decline.

INCREASE IN GOOD BOT ACTIVITY, which went up by 4.4 percent.



1.2%

**MONITORING BOTS**

Health checkers that monitor website availability and the proper functioning of various online features.



2.9%

**COMMERCIAL CRAWLERS**

Spiders used for authorized data extractions, usually on behalf of digital marketing tools.



6.6%

**SEARCH ENGINE BOTS**

Bots that collect information for search engine algorithms, which they use to make ranking decisions.



12.2%

**FEED FETCHERS**

Bots that ferry website content to mobile and web applications, which they then display to their users.



24.3%

**IMPERSONATORS**

Bots that assume false identities to bypass security solutions. They are commonly used for DDoS assaults.



1.7%

**SCRAPERS**

Bots used for unauthorized data extraction and the reverse engineering of pricing models.



0.3%

**SPAMMERS**

Polluters that inject spam links into forums, discussions and comment sections.



2.6%

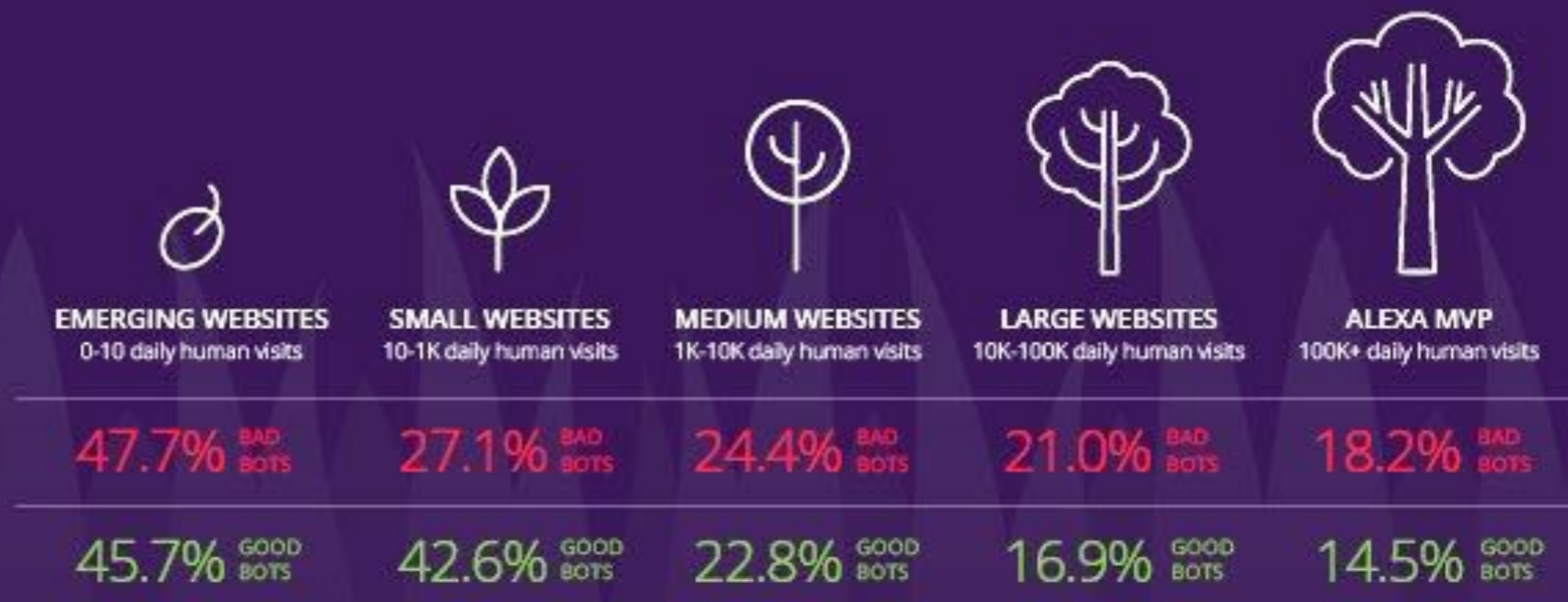
**HACKER TOOLS**

Scavengers that look for sites with vulnerabilities to exploit for data theft, malware injection, etc.

## WHAT'S YOUR BOT SITUATION?

Bot traffic patterns change depending on the popularity of a given website.

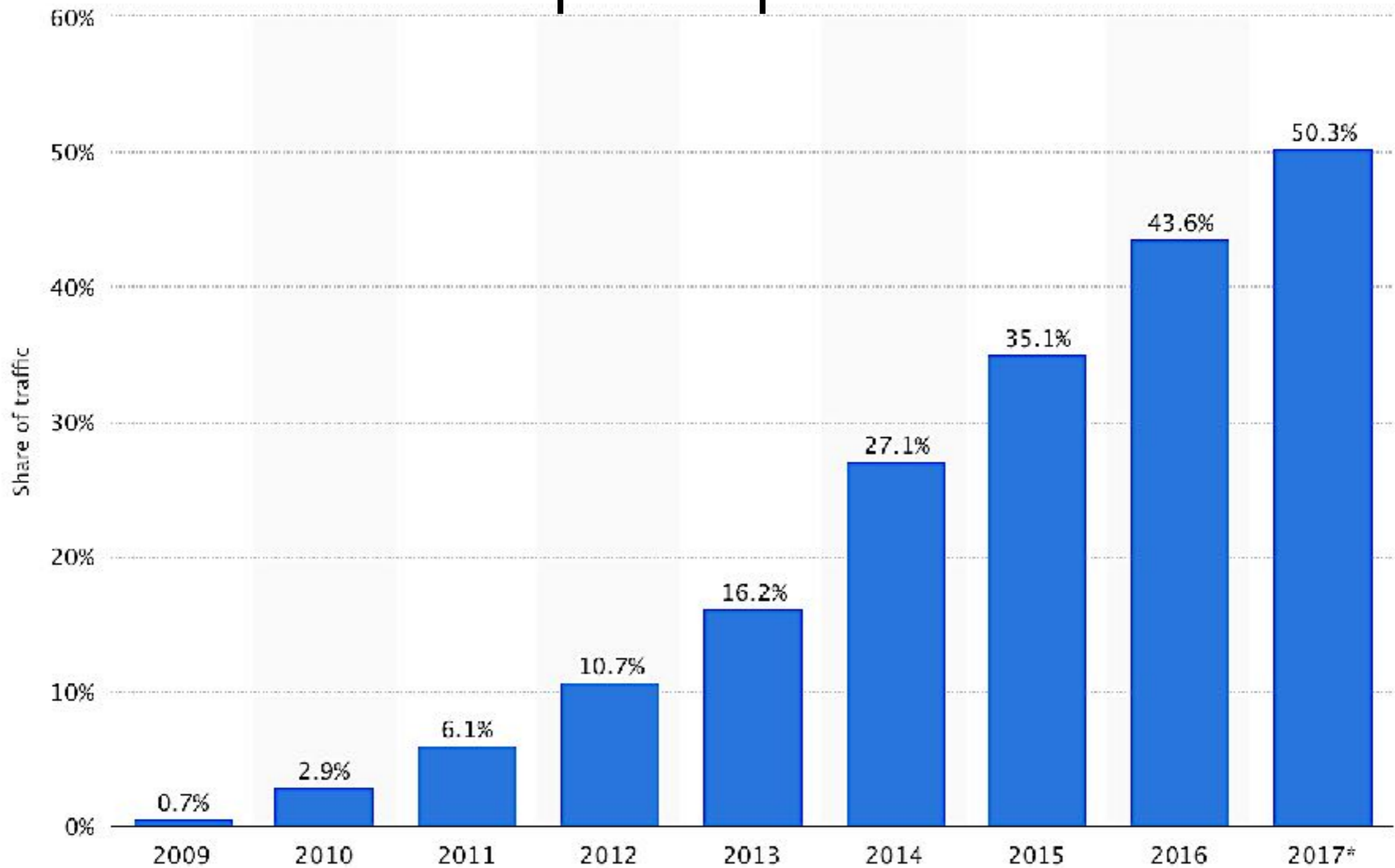
Here's the amount of bot traffic you can expect based on the number of human daily visitors to your site:



## 2012-2016 Traffic Breakdown (by visitor type)



# Proportion des pages webs lues sur téléphones portables



<https://www.statista.com/statistics/241462/global-mobile-phone-website-traffic-share/>

© Statista 2017

# Importance du balisage *sémantique*

- Étiquetage des informations d'une page
- **Pas seulement de l'affichage** pour les humains
- Important
  - pour les gens qui ne peuvent voir
  - pour que les moteurs de recherche *comprennent* vos pages
  - pour les applications puissent utiliser vos pages
- Respect des standards avec une validation
- Dans ce cours, l'**organisation interne des pages** est aussi importante que leur présentation