

BTSF

Websites, web browsers and basic searches — Diego RIVERO

Contract number 2017 96 05 – New Food Investigation Techniques –
Phase II - *Course 2a: E-Commerce of food standard*

© European Union 2020

Unless otherwise noted the reuse of this presentation is not authorised. For any use or reproduction of elements that are owned by the EU, permission may need to be sought directly from the respective right holders. All statements and references in this presentation come from the Training coordinator and tutors and do not represent the official position of the European Commission.



BTSF What are web pages? Just pieces of computer data

- The web page refers to a computer file, usually written in HTML or comparable markup language.
- Web browsers coordinate the various web resource elements for the written web page, such as style sheets, scripts, and images, to present the web page.
- The importance of links/hyperlinks.

BTSF

BTSF What are web pages? Just pieces of computer data

- When web pages are stored in a common directory of a web server, they become a website.
- A website will typically contain a group of web pages that are linked together, or have some other coherent method of navigation.
- The most important web page to have on a website is the index page (index.html)

BTSF

BTSF What are web pages? Just pieces of computer data

- Most web pages content are dynamic:
 - In the way they are generated (server and client software, java scripts)
 - In their content itself (text, links, images, etc.) and
 - In how the content is shown (user customisation).
- That's why it's very important that we capture and save this data while doing the investigation, as it is the best way to ensure we have the best possible evidence.

NOTE: while receiving a consumer's complaint it is relevant to receive their digital evidence

BTSF HTML

HTML – Hypertext Markup Language

- The basic original computer language that most web pages are created in.
- A markup language is a set of markup tags
- HTML documents are described by HTML tags.
- Each HTML tag describes different document content.
- The HTML data is downloaded via the Internet and then interpreted via the web browser.

BTSF

BTSF HTML – main elements

HTML elements are delineated by *tags*, written using angle brackets.

Syntax: `<tagname>content</tagname>`

- a start tag
- content
- end tag

Let's try!

```
<!DOCTYPE html>
<html>
  <head>
    <title>This is a title</title>
  </head>
  <body>
    <p>Hello world!</p>
  </body>
</html>
```

BTSF Example: Basic HTML code

- The **DOCTYPE** declaration defines the document type to be HTML
- The text between **<html>** and **</html>** describes an HTML document
- The text between **<head>** and **</head>** provides information about the document
- The text between **<title>** and **</title>** provides a title for the document
- The text between **<body>** and **</body>** describes the visible page content
- The text between **<h1>** and **</h1>** describes a heading

BTSF Example: Basic HTML code

- The text between `<p>` and `</p>` describes a paragraph
- The text between `<a>` and `` describes a link
- The `` tag describes images

HTML elements can be nested (elements can contain elements)

All HTML documents consist of nested HTML elements.

BTSF HTML versions

Version	Year
HTML	1991
HTML 2.0	1995
HTML 3.2	1997
HTML 4.01	1999
XHTML	2000
HTML5	2014

Like everything related to computers, html has got several versions...

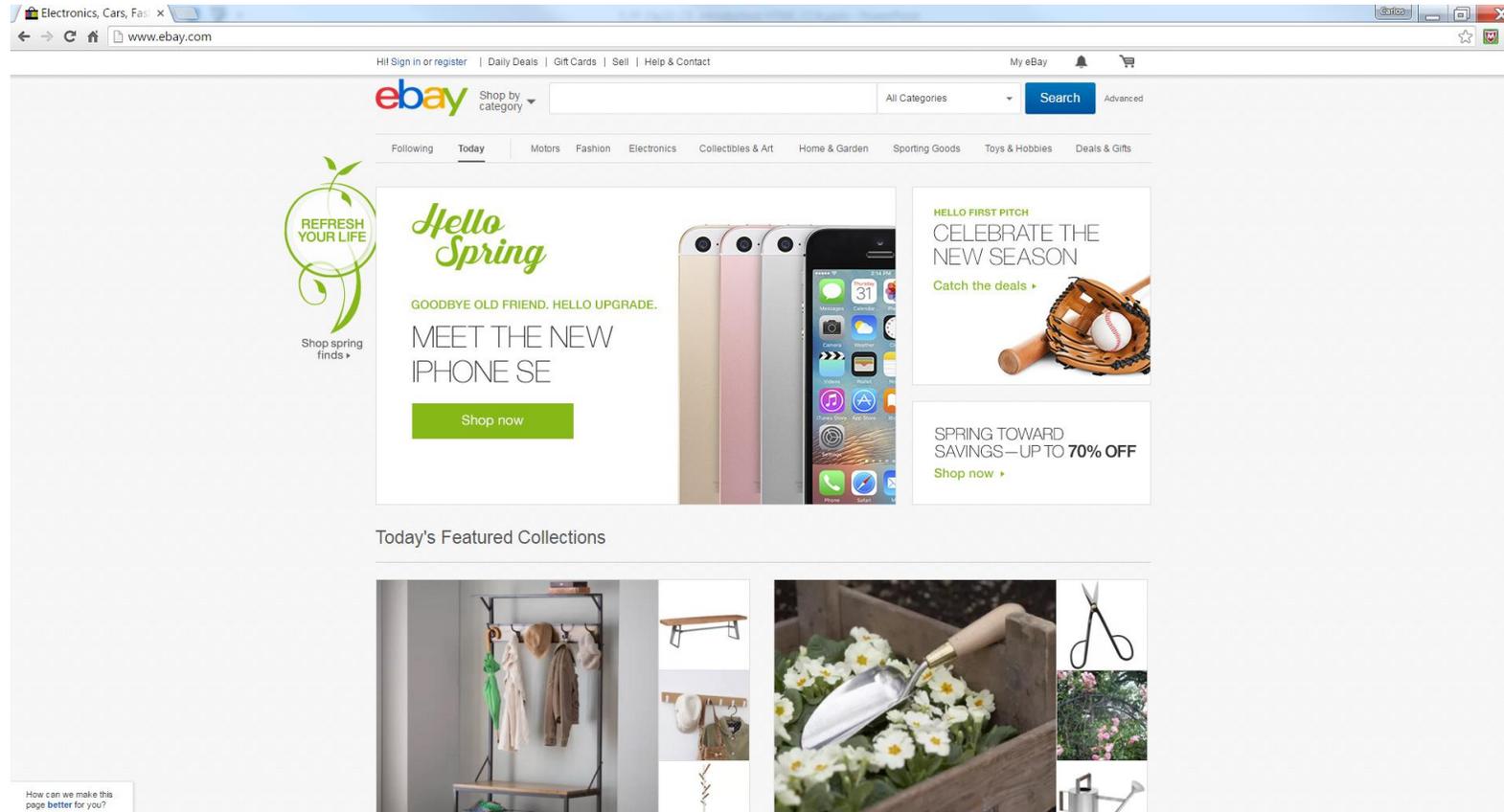
BTSF

BTSF HTML and browsers

- Browsers do not display the HTML tags, but use them to **interpret** the content of the page.
- HTML can embed [scripts](#) written in languages such as [JavaScript](#) which affect the behavior of HTML web pages.

BTSF

BTSF Example: HTML interpreted by web browsers



BTSF Some things to take into account



- External resources
- Dynamic content: cookies and profiling

BTSF

BTSF HTML Tutorial



Interested in learning more about html?

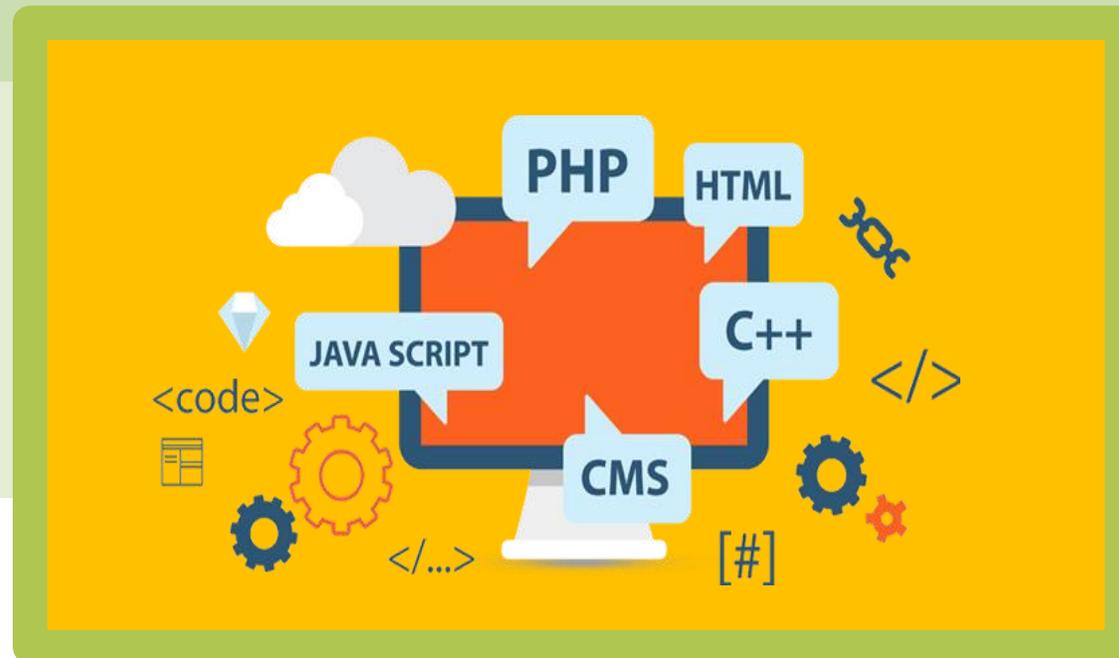
An easy to follow html tutorial:

<http://www.w3schools.com/html/default.asp>

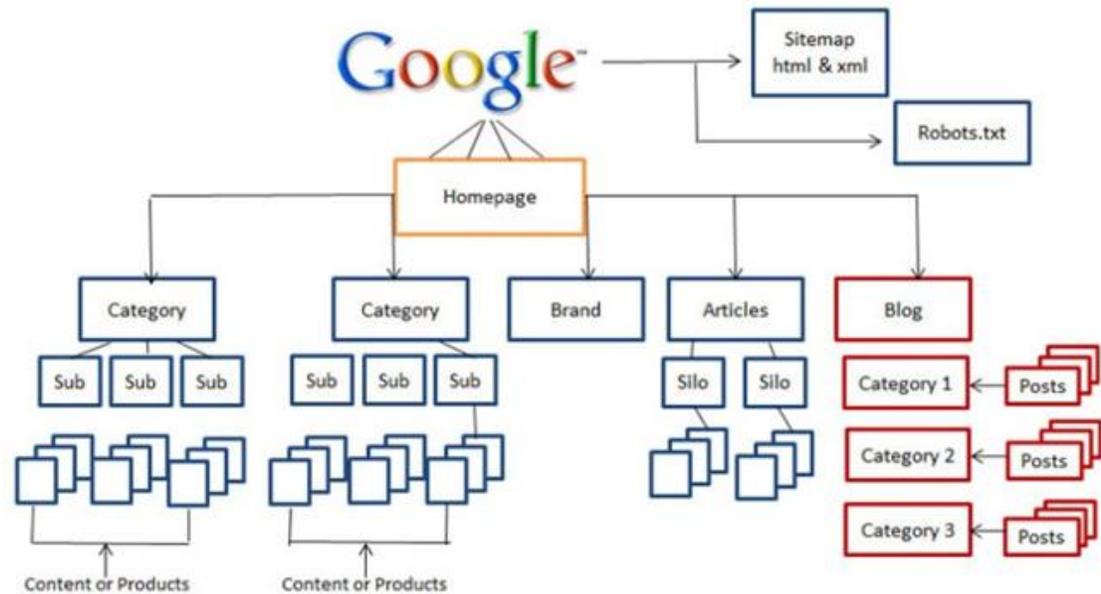
BTSF

Website structure

- A website is just an ordered compilation of several (even hundreds) documents (html, pictures, files, etc.) called resources.



BTSF Website structure

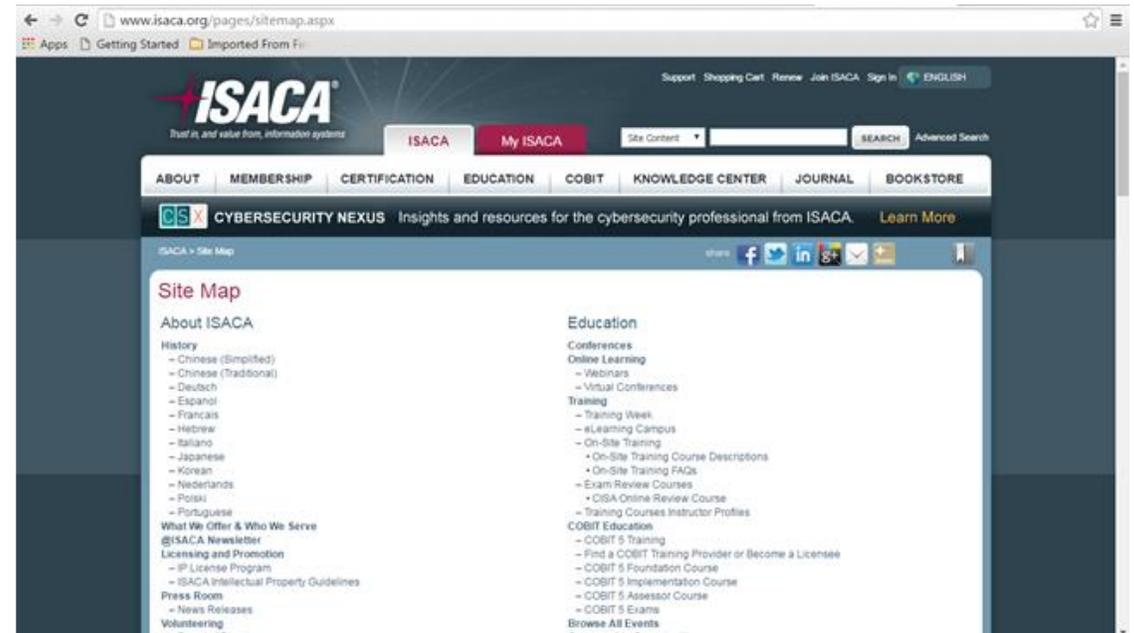


The size of the website will vary depending on its purpose. Its structure is traditionally represented like organigram charts.

BTSF

BTSF Example: Site Map

- Some sites provide their “site map”:
- <https://www.isaca.org/pages/sitemap.aspx>



BTSF Recap

- **Web page** is “a [computer file](#)”, usually written in [HTML](#) or comparable [markup language](#) commonly with hyperlinks.
- Web pages files usually requires of other files/resources (images, videos, style sheets, scripts) to fully show the desired content.
- Most web pages content is dynamic.
- When web pages are stored in a common directory of a web server, they become a [website](#).



BTSF Recap

- HTML – Hypertext Markup Language
 - A markup language is a set of markup tags.
 - HTML documents are described by HTML tags and every tag describes different document content.
- Syntaxis: `<tagname>content</tagname>`
- HTML can be edited using a simple text editor or with specific software.
- Browsers interpret the content of the page without displaying the HTML tags.
- Websites have a structure. It is common to navigate throughout the website pages through the webpages hyperlinks.



BTSF Internet browsers



- What are they?
- How many are there?
- How do they work?
- How to configure?

BTSF

BTSF Internet browsers

- Just a computer program for bringing information resources (retrieving), allowing to view the information (display or presenting), and access other information (navigating, following links).
- Although designed to be used on the [World Wide Web](#), they can also be used to access information provided by [web servers](#) in [private networks](#) or files in [file systems](#).



BTSF Major Internet browsers

[Firefox](#), [Internet Explorer/Microsoft Edge](#) (Windows' default), [Google Chrome](#), [Opera](#), and [Safari](#) (Mac's default) and their equivalent in mobile devices

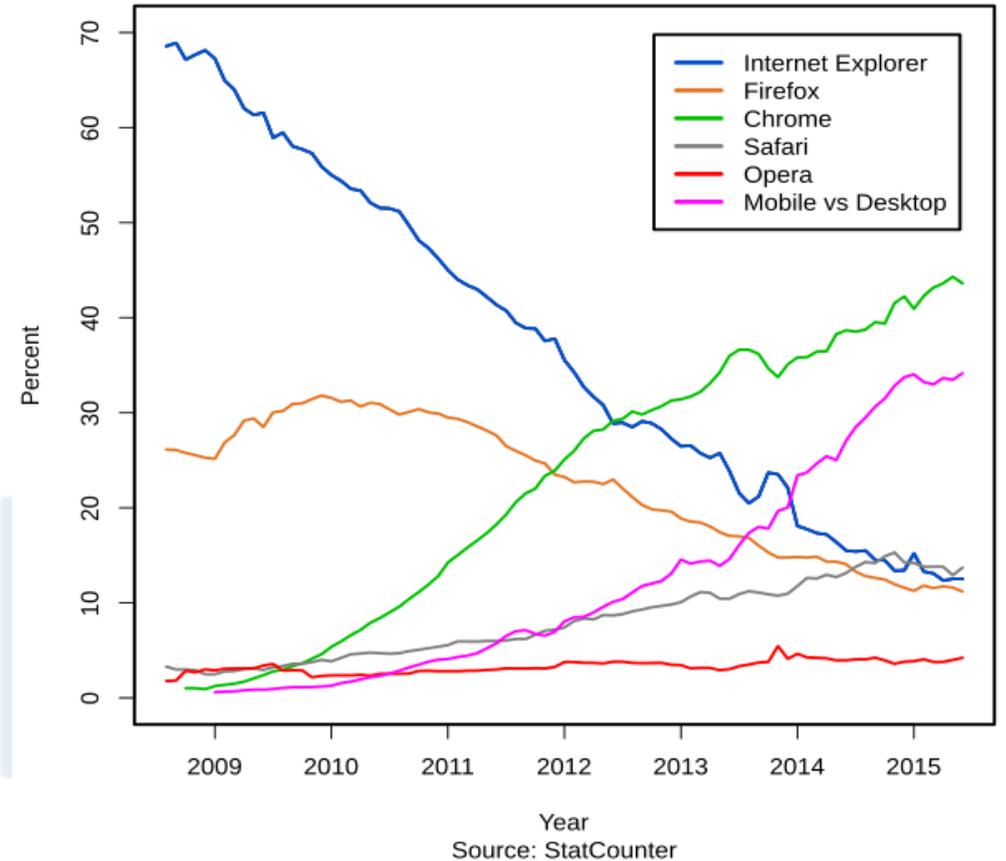
[Web-browser market share - Europe - Feb 2019](#)

[Desktop web-browser market share - Europe - Feb 2019](#)

[Tablet web-browser market share - Europe - Feb 2019](#)

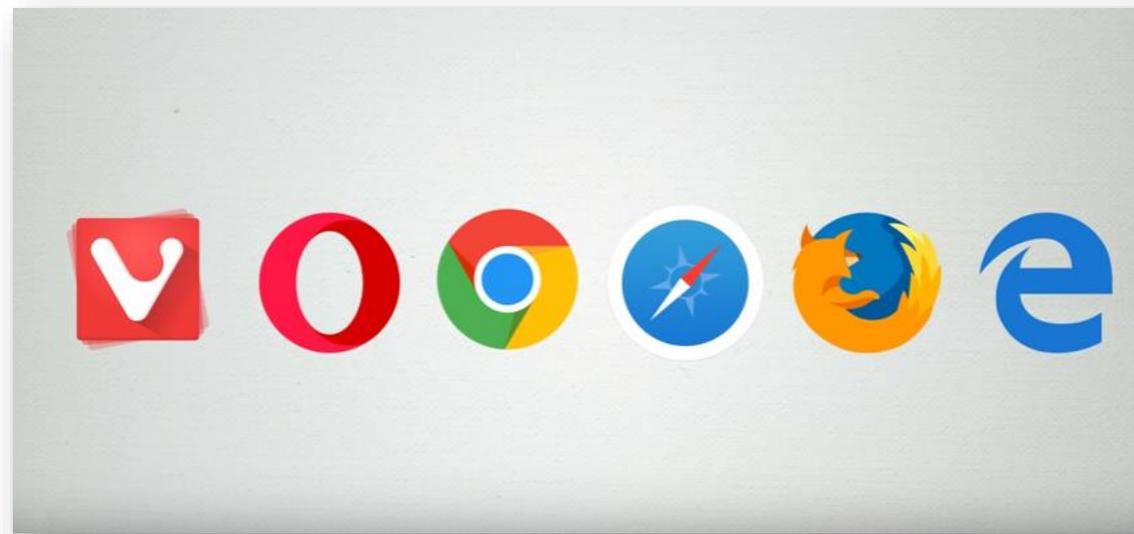
[Mobile web-browser market share - Europe - Feb 2019](#)

Usage share of web browsers



BTSF How web browsers work

- <https://youtu.be/z0HN-fG6oT4>



BTSF How web browsers work

- The main function of a browser is to present the web resource you choose, by requesting it from the server and displaying it in the browser window.
- The resource is usually an HTML document, but may also be a PDF, image, or some other type of content.
- The location of the resource is specified by the user using a URL (Uniform Resource Locator).

BTSF How web browsers work

- Some web browsers block or cannot display some types of files on the Internet, notably some kinds of multimedia files, such as sound, video, and animation files.(e.g Flash files)
- To view or play these files, you need what is called a helper application or plug-in.
- You must configure your web browser or operating system to launch these plug-ins whenever you click an object that needs them to be viewed.

BTSF How web browsers work

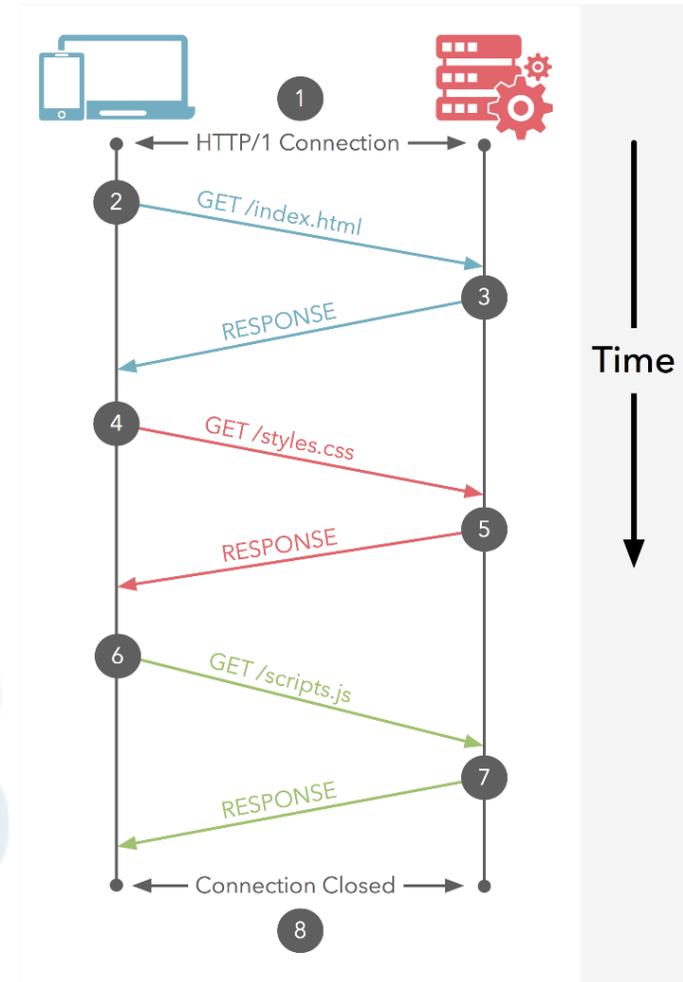
They interpret and display the URL depending on its prefix:

- *http: or https: – Hypertext Transfer Protocol (Secure)*
- *ftp: or ftps: - File Transfer Protocol (Secure)*
- *file: - for local files*

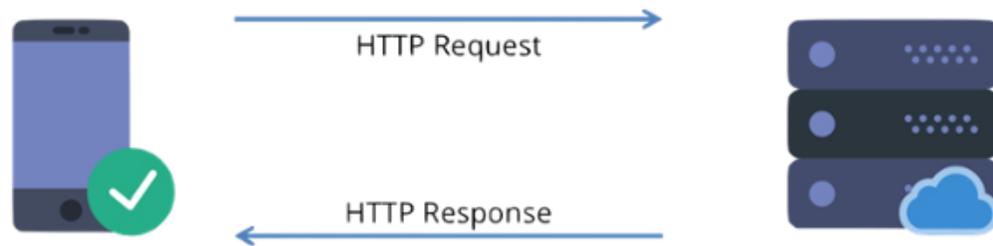
Modern web browsers support a combination of standards-based and de facto HTML and XHTML.

BTSF How web browsers work

- As soon as we type a web address, some communication starts to happen.
- The browser send an HTTP request through the Internet to where the server that runs the website is located.
- That server sends back an HTTP response that contains the website in HTML form.
- The browser interprets and displays the HTML code.



BTSF How web browsers work



The process repeats every time we click a link

Further explanation of how HTTP communication works:

<https://www.youtube.com/watch?v=DuSURHrZG6I>

BTSF

How web browsers work

- Navigation, Tabs & Address bar
- History & Bookmarks
- Downloads
- Find
- Settings: Cache, Cookies, clear browsing data, search engines, privacy & security)
- Add-ons / Extensions

BTSF

BTSF Recap

- A **browser** connects to web servers which are hosting the web resources using request-response communication.
- Browsers present the web resources as indicated in the web page code (typically HTML code).
- A **plug-in** is a piece of software which complements other (i.e. web browser) to provide an extra feature such as playing those files (typically multimedia), that browsers don't support by themselves.
- **Users** should know the main **features** and how to configure the main **settings** to make the most of the browser capabilities.



BTSF Search Engines basics

- We can use standalone search engine websites:
 - <https://www.google.com>
 - <https://www.bing.com>
 - <https://www.yahoo.com>
 - <http://search.carrot2.org>
 - <https://duckduckgo.com>
- Or we can access those search engines using the feature built-in in our web browser

BTSF

BTSF

Search Engines basics

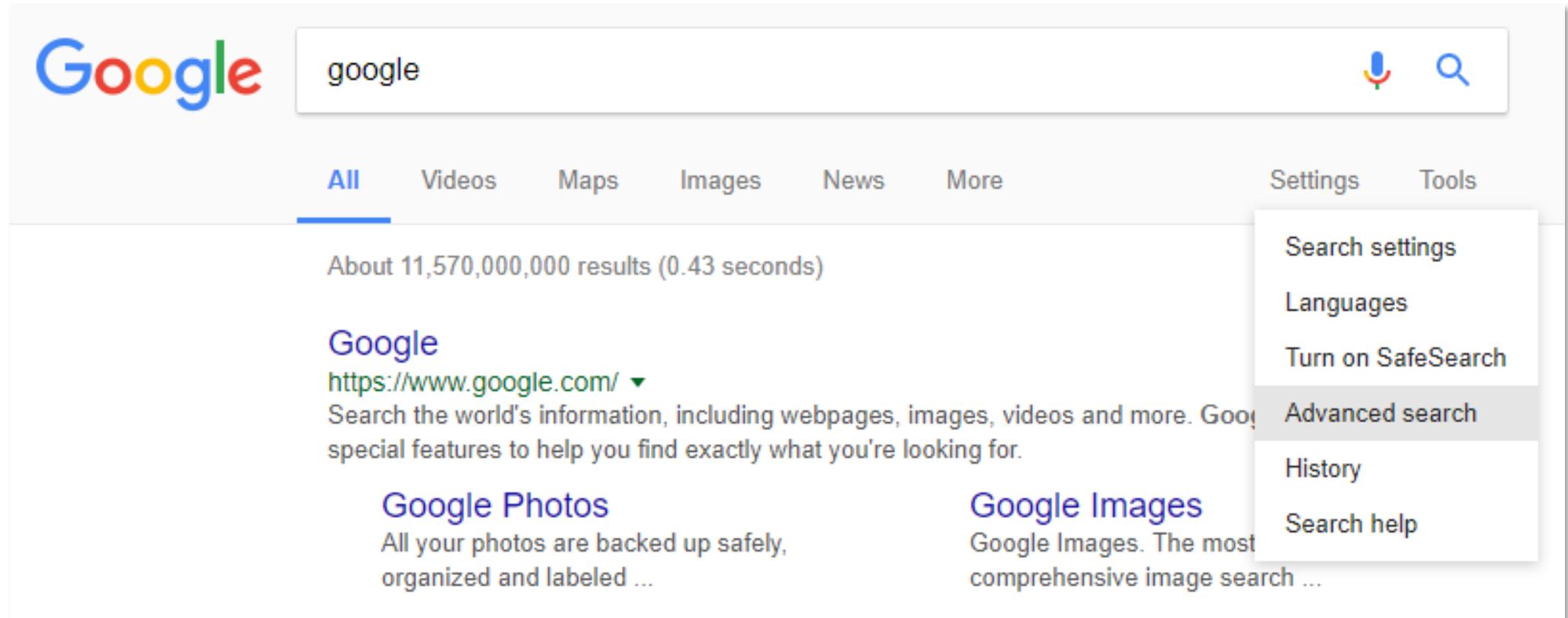
- How search works
https://www.youtube.com/watch?v=LVV_93mBfSU
- Advanced search in Google
 - <https://youtu.be/FtSeAwRhPKs>
- Tips to refine searches:
<https://support.google.com/websearch/answer/2466433>

BTSF Search Engines - advanced search

- Advanced search will help you to refine your search avoiding results not relevant for your investigation.
- Access through:
 - Advanced Search interface of the search engines:
 - Google: https://www.google.com/advanced_search
 - Yahoo: <https://search.yahoo.com/web/advanced>
 - Search Symbols & Operators:
 - Symbols such as "" - & ..
 - Operators followed by : such as "site", "typefile", "intitle", "related", "info" or "cache"

BTSF

BTSF Search Engines – Google advanced search



The image shows a screenshot of the Google search homepage. The search bar contains the word "google". Below the search bar, the navigation tabs include "All", "Videos", "Maps", "Images", "News", "More", "Settings", and "Tools". The "Tools" dropdown menu is open, showing options: "Search settings", "Languages", "Turn on SafeSearch", "Advanced search" (highlighted), "History", and "Search help". The search results for "google" are displayed, showing "About 11,570,000,000 results (0.43 seconds)". The first result is "Google" with the URL "https://www.google.com/". Below this, there are two featured links: "Google Photos" and "Google Images".

Google

google

All Videos Maps Images News More Settings Tools

About 11,570,000,000 results (0.43 seconds)

Google

<https://www.google.com/>

Search the world's information, including webpages, images, videos and more. Google has special features to help you find exactly what you're looking for.

Google Photos

All your photos are backed up safely, organized and labeled ...

Google Images

Google Images. The most comprehensive image search ...

Search settings

Languages

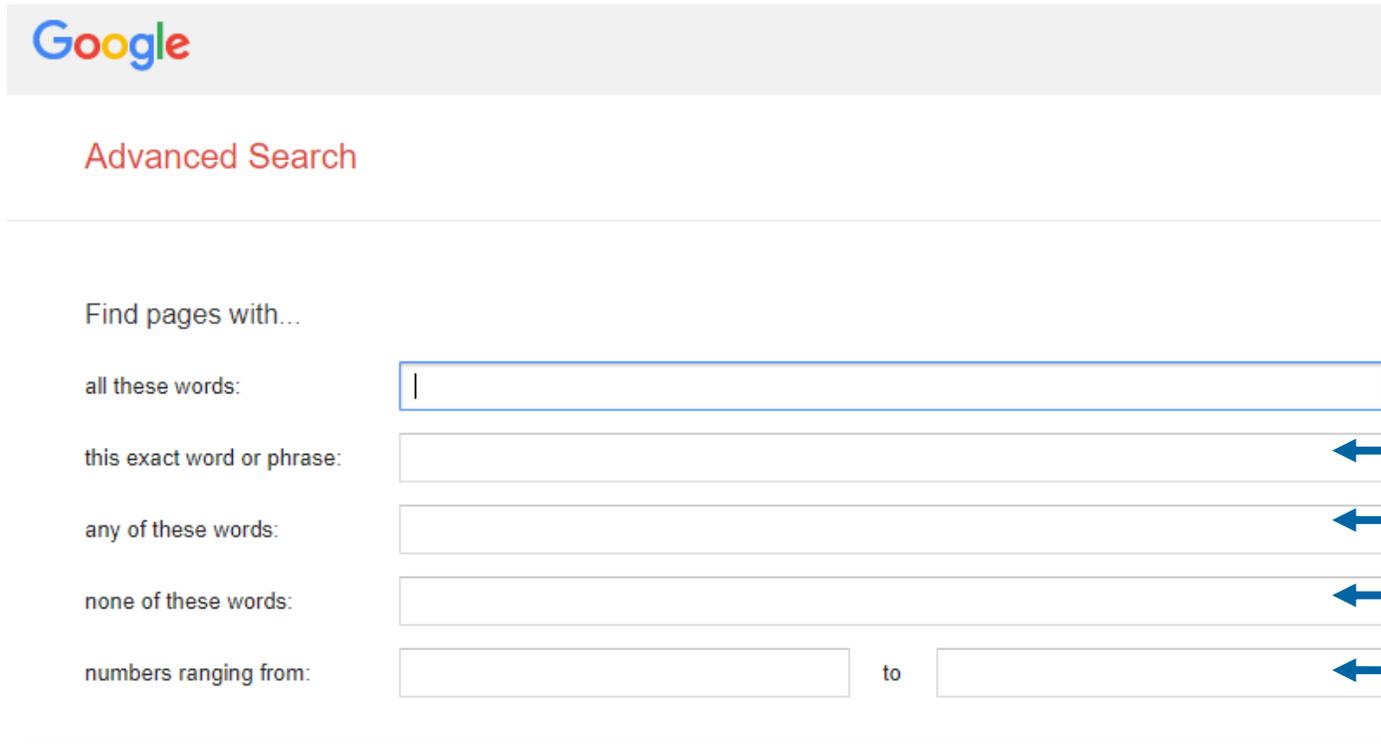
Turn on SafeSearch

Advanced search

History

Search help

BTSF Search Engines – Google advanced search



Google

Advanced Search

Find pages with...

all these words:

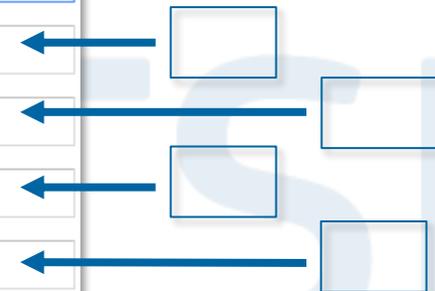
this exact word or phrase:

any of these words:

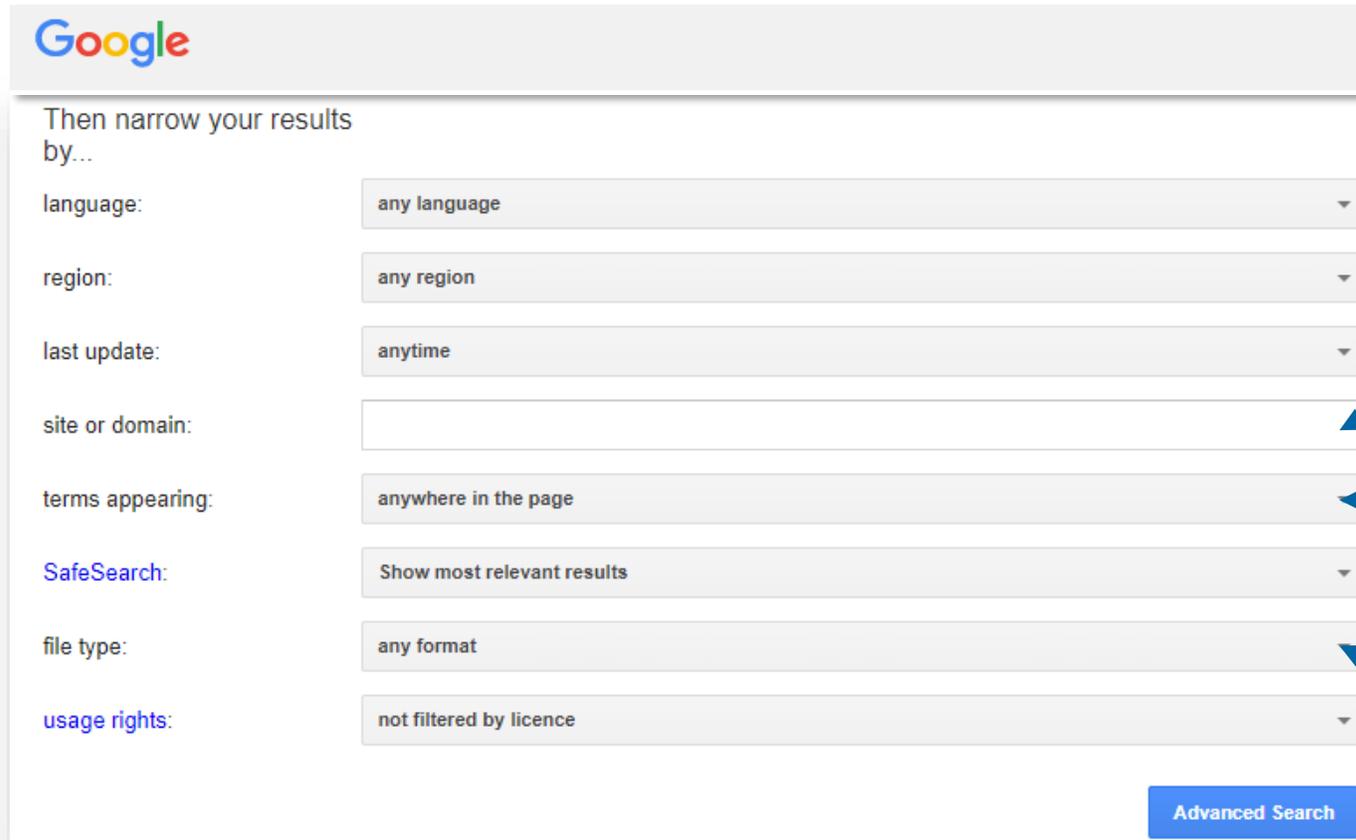
none of these words:

numbers ranging from: to

Equivalent symbol



BTSF Search Engines – Google advanced search



Google

Then narrow your results by...

language: any language

region: any region

last update: anytime

site or domain:

terms appearing: anywhere in the page

SafeSearch: Show most relevant results

file type: any format

usage rights: not filtered by licence

Advanced Search

Equivalent operator

BTSF Search Engines – Alternatives to text

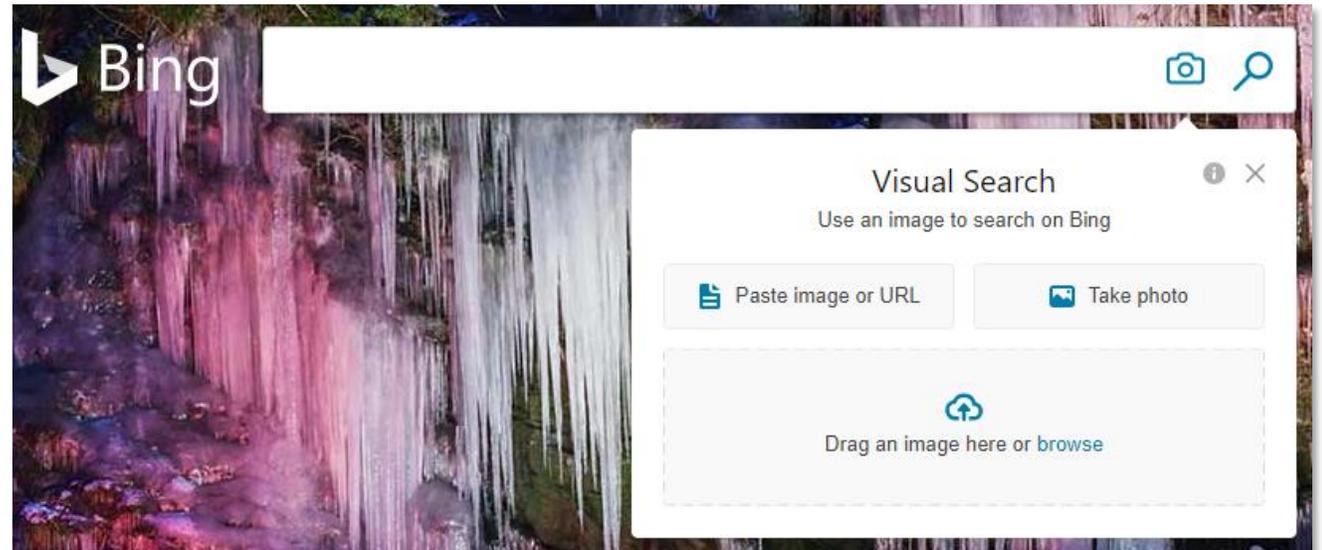
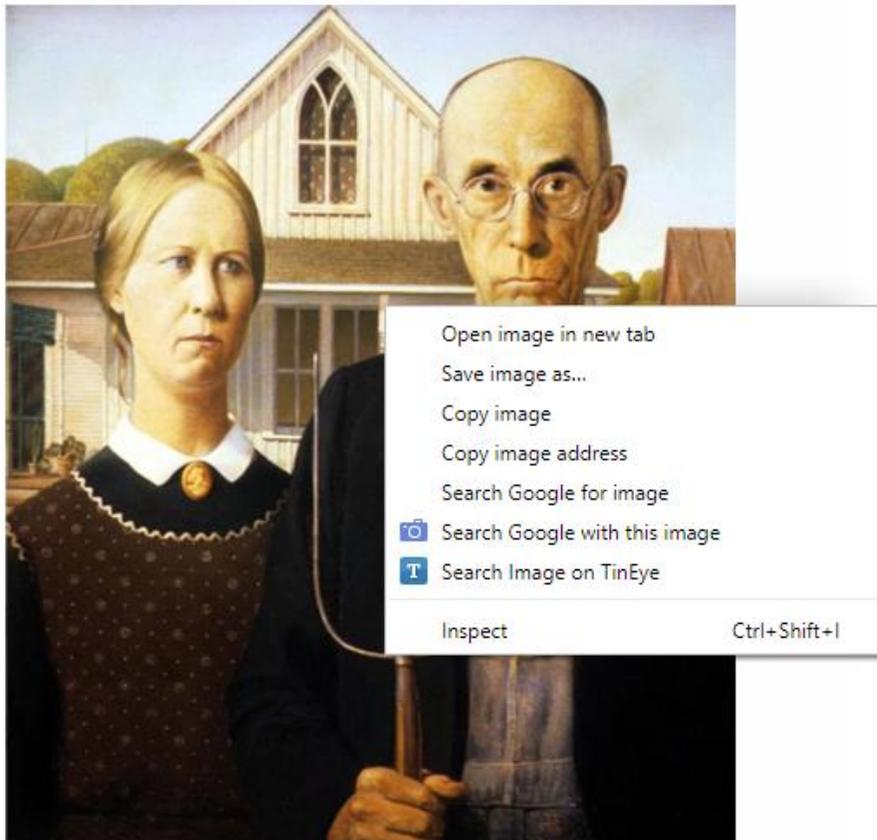
- Search by image: Reverse search image
 - It is an Internet search using an image/photo in order to find the same (or a similar) image on the web
 - When may you need to use a reverse image search?
 - When you search on the Internet for a specific product, but...
 - you cannot find it searching by the brand name;
 - you want to find similar products (same or different brand);
 - you want to find results from online shops in different languages/countries (in which the name of the product may be different);
 - you want to cross-check if you can find other results than those found when searching keywords in the search engine.

BTSF Search Engines – Alternatives to text

- Search by image: Reverse search image
 - How can you make a reverse image search?
 - Websites: e.g. [TinEye](#), [Yandex](#), [Bing](#) or [Google](#)
 - Paste Image
 - Image URL
 - Upload image
 - Browser add-on/extensions: Some image search engines (e.g. [TinEye](#), [Google](#) & [others](#)) provide add-on/extensions to use this service directly from your browser
 - Phone app: e.g. CamFind (available for [Android](#) & [iOS](#))

BTSF Search Engines – Alternatives to text

- Search by image: Reverse search image



BTSF

Web-browsers and add-ons

- Browsers such as Chrome and Firefox allow 'extensions' to be added to access features created by third-party developers.
- This features can be used for investigation or evidence gathering and to facilitate other tasks which may enhance enforcers' performance.
- Let's practice:
 - Chrome: [Fireshot](#) / [Awesome Screenshot](#)
 - Firefox: [Flagfox](#) / [Save as PDF](#)

BTSF

BTSF

Thank you

European Commission
Consumers, Health and Food Executive Agency
DRB A3/042
L-2920 Luxembourg

AENOR INTERNACIONAL
6, Genova street. 28004. Madrid, SPAIN
Tel: +34 91 432 61 25
Mail: 20179605NFIT@aenor.com
www.btsf-aenor.com

AENOR

© European Union 2020

Unless otherwise noted the reuse of this presentation is not authorised. For any use or reproduction of elements that are owned by the EU, permission may need to be sought directly from the respective right holders. All statements and references in this presentation come from of the Training coordinator and tutors and do not represent the official position of the European Commission.

Slide xx: [element concerned](#), source: [e.g. Fotolia.com](#); Slide xx: [element concerned](#), source: [e.g. iStock.com](#)

