

BTSF

Web pages capture — 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.

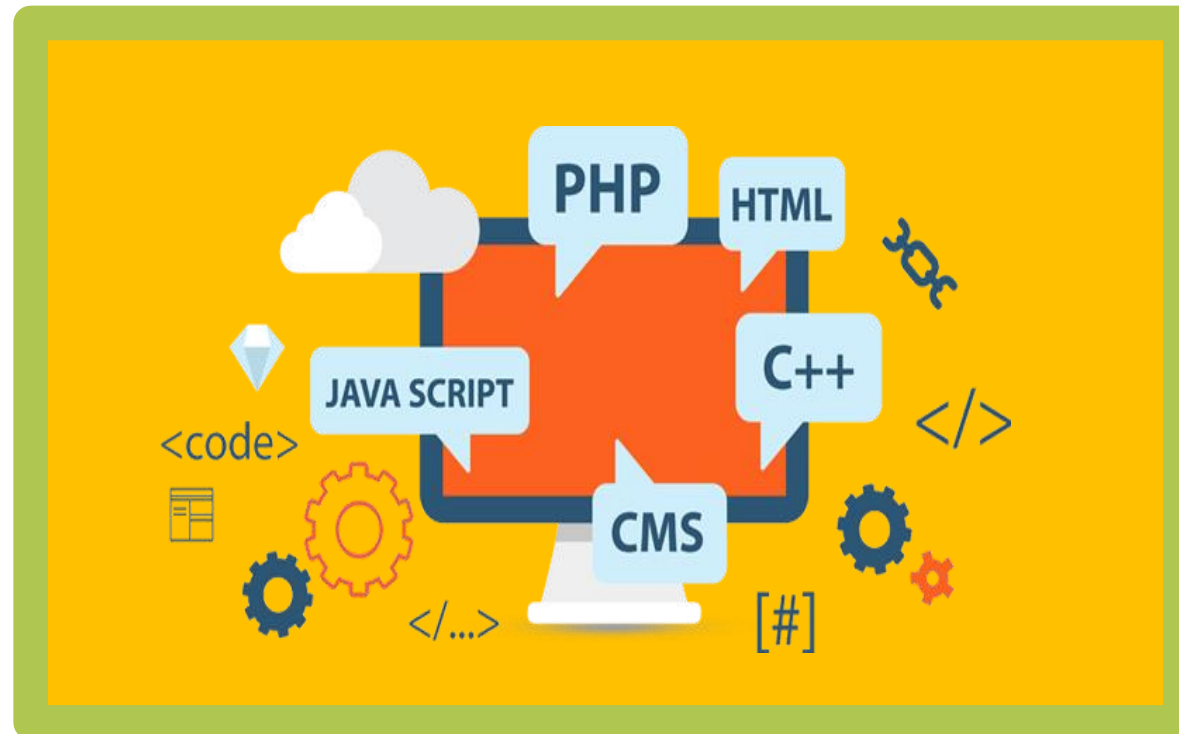
BTSF Web pages capture

- Website structure
- Saving source code and webpages
- Saving webpages with third parties
- Saving complete websites
- Video captures



BTSF Website structure

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



BTSF

How do we save a web page?

- Many investigations will centre around one web page.
- It's very useful/essential to capture this as best evidence.
- Let's practice...

BTSF

BTSF Saving web pages

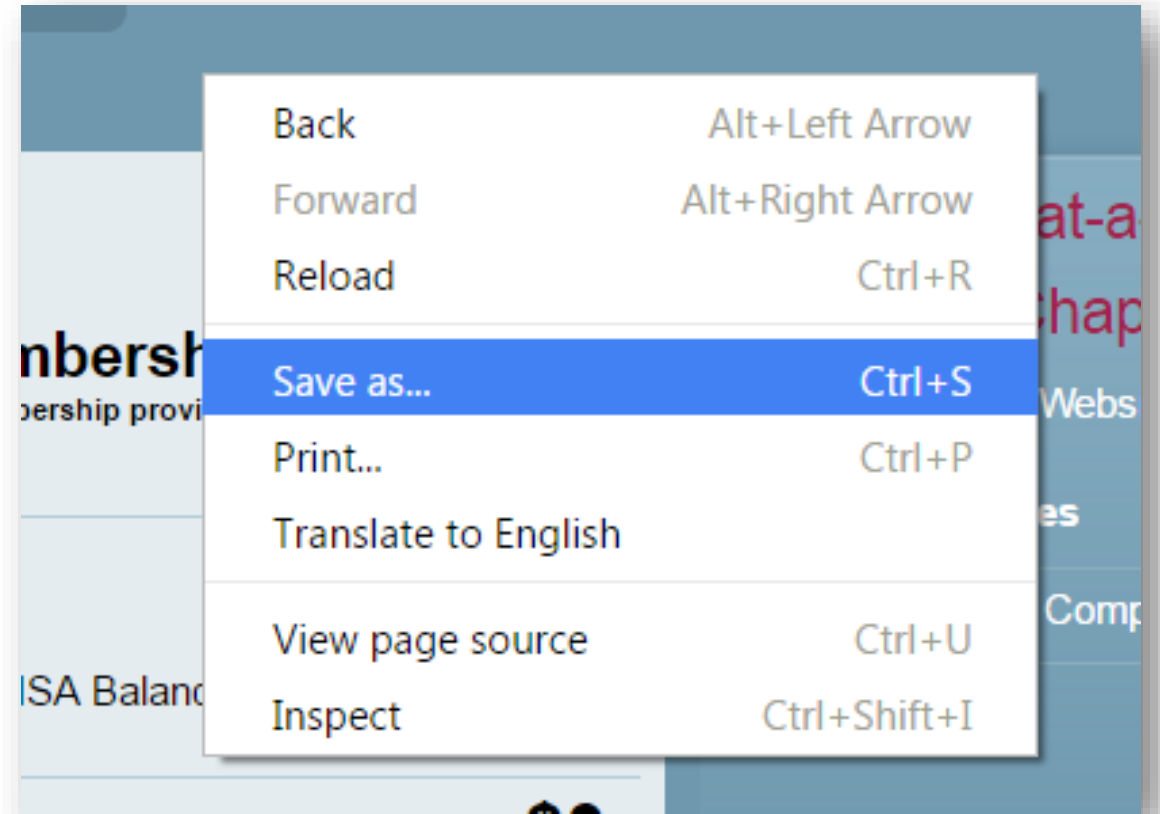
- ✓ While viewing a web page, a copy of it is saved locally; this is what is being viewed.
- ✓ Depending on the browser settings, this copy may be deleted at any time, or stored indefinitely, sometimes without the user realizing it.
- ✓ Most browsers provide options for saving a web page more permanently.

BTSF

BTSF

Let's practice... Saving web pages and source code

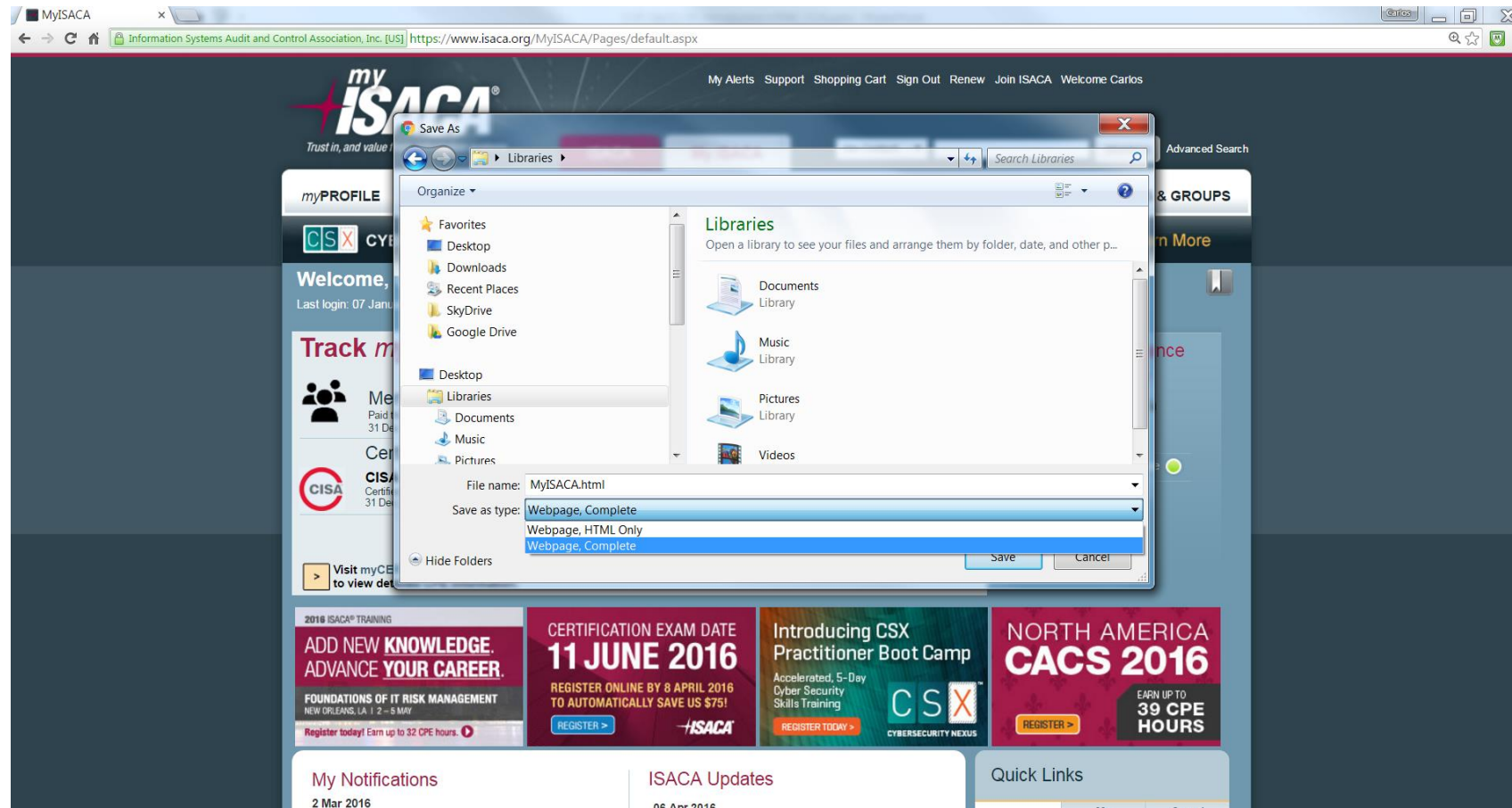
- Google Chrome browser... (but more or less the same with every browser)
- Right-click + Save as
- Ctrl + S
- Save webpage complete



BTSF

The screenshot shows the ISACA MyISACA website interface. The browser address bar displays the URL: <https://www.isaca.org/MyISACA/Pages/default.aspx>. The page features the ISACA logo with the tagline "Trust in, and value from, information systems". Navigation tabs include "ISACA", "My ISACA", and "Site Content". A search bar is located on the right. Below the navigation, there are tabs for "myPROFILE", "myMEMBERSHIP", "myCERTIFICATIONS", "myDOCUMENTS", "myPURCHASES", and "myCONNECTIONS & GROUPS". A "CYBERSECURITY NEXUS" banner is visible. The main content area welcomes the user, Carlos, and shows their last login on 07 January 2016. A "Track mySTATUS" section displays membership details, including a renewal notice and a CISA certification status. A context menu is open over the CISA section, showing options like "Back", "Forward", "Reload", "Save as...", "Print...", "Translate to English", "View page source", and "Inspect". The bottom of the page features promotional banners for "2016 ISACA® TRAINING", "CERTIFICATION EXAM DATE 11 JUNE 2016", "Introducing CSX Practitioner Boot Camp", and "NORTH AMERICA CACS 2016". A footer section includes "My Notifications", "ISACA Updates", and "Quick Links".

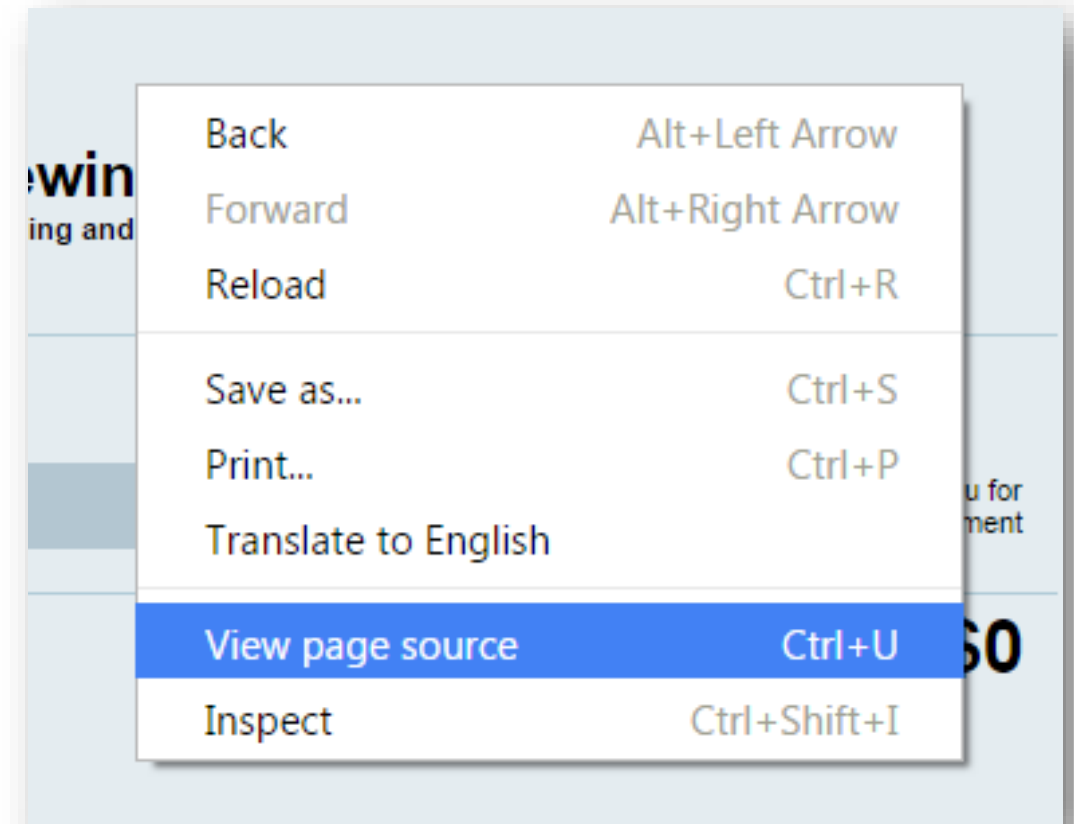
BTSF



BTSF

Let's practice... Saving web pages and source code

- Google Chrome browser
- Viewing and saving page source code
 - Right-click + View page source
 - Ctrl + U
- Or save webpage HTML only



BTSF

The screenshot shows the MyISACA website interface. At the top, the ISACA logo is displayed with the tagline "Trust in, and value from, information systems". Navigation links include "My Alerts", "Support", "Shopping Cart", "Sign Out", "Renew", "Join ISACA", and "Welcome Carlos". A search bar is located on the right.

The main navigation menu includes: myPROFILE, myMEMBERSHIP, myCERTIFICATIONS, myDOCUMENTS, myPURCHASES, and myCONNECTIONS & GROUPS.

A banner for "CYBERSECURITY NEXUS" is visible, with the text "Insights and resources for the cybersecurity professional from ISACA." and a "Learn More" link.

The user profile section for "Carlos" shows a welcome message, last login date (07 January 2016), and messages (0). A "Remember My Preferences" checkbox is also present.

The "Track mySTATUS" section displays membership and certification details:

- Membership:** Paid through: 31 Dec 2016. A message says "Thank you for renewing" with a link to "Continue to enjoy the resources, training and View your member benefits".
- Certification:** CISA Certified through: 31 Dec 2016. A message says "2015 CISA CPE Hours Remaining Due: 30".

A context menu is open over the "Track mySTATUS" section, showing options: Back, Forward, Reload, Save as..., Print..., Translate to English, View page source (highlighted), and Inspect.

Below the status section, there are several promotional banners:

- 2016 ISACA® TRAINING:** ADD NEW KNOWLEDGE. ADVANCE YOUR CAREER. FOUNDATIONS OF IT RISK MANAGEMENT. NEW ORLEANS, LA | 2 - 5 MAY. Register today! Earn up to 32 CPE hours.
- CERTIFICATION EXAM DATE 11 JUNE 2016:** REGISTER ONLINE BY 8 APRIL 2016 TO AUTOMATICALLY SAVE US \$75! REGISTER >
- Introducing CSX Practitioner Boot Camp:** Accelerated, 5-Day Cyber Security Skills Training. REGISTER TODAY >
- NORTH AMERICA CACS 2016:** EARN UP TO 39 CPE HOURS. REGISTER >

At the bottom, there are sections for "My Notifications" (2 Mar 2016), "ISACA Updates" (06 Apr 2016), and "Quick Links".

BTSF

```
<!DOCTYPE html PUBLIC
<html __expr-val-dir="
<head><meta name="GENE
content="IE=EmulateIE8
MyISACA
</title><script src="
<script type="text/jav

<script src="
<script src="
<script src="
<script type="
<script type="
<!--<script ty
<script type="
<script src="
<script type="

<link rel="stylesheet"
href="/_layouts/1033/s
rev=843zq671I5zQ9zaPe
type="text/css" href="/_layouts/1033/s
type="text/css" href="/_layouts/1033/s
type="text/css" href="/_Style Library/

<!-- Include P
<link rel="sty
<link rel="sty

<script type="text/jav
<script type="text/jav
<script type="text/jav
<script type="text/jav
<!--[if gte IE
<script
<!--[endif]-->
<!--[if lt IE
<script
<script
<script
<link
<!--[endif]-->

<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
<html __expr-val-dir="ltr" dir="ltr">
<head><meta name="GENERATOR" content="Microsoft SharePoint" /><meta http-equiv="Content-Type" content="text/html; charset=utf-8" /><meta http-
equiv="Expires" content="0" /><meta http-equiv="X-UA-Compatible" content="IE=EmulateIE8" /><META NAME="ROBOTS" CONTENT="NOHTMLINDEX"/><title>

MyISACA

</title><script src="/_layouts/1033/init.js?rev=TFgMAEbR8Rj7bIFPqCsR5w%3D%3D"></script>
<script type="text/javascript" language="javascript" src="/_layouts/1033/non_ie.js?rev=yfNry4hY0Gwa%2FPDNGrqXVg%3D%3D"></script>

<script src="/9ffa4600af84abc9026dae89076197a82e6ef025/satelliteLib-1e5fc580126ed2f59d81e58814b5de256b211d82.js"></script>
<script src="/_layouts/Ontollica/jquery-1.3.2.js" type="text/javascript"></script>
<script src="/Style Library/ajax/jquery/jquery.min.js" type="text/javascript"></script>
<script src="/Style Library/ajax/jqueryui/jquery-ui.min.js" type="text/javascript"></script>
<script type="text/javascript" src="/Style Library/scripts/BCCDefault-min.js"></script>
<script src="/_layouts/community/peoplepicker/definejsvars.aspx" type="text/javascript"></script>
<!--<script type="text/javascript" src="https://cloud.moovweb.net/api/public/js_redirection/moovweb/isaca-prod.js"></script>-->
<script type="text/javascript" src="/Style Library/MoovWeb/moovweb_redirection.js"></script>
<script src="/Style Library/scripts/scripts_main.js" type="text/javascript"></script>
<script type="text/javascript" src="/Style Library/TestandTarget/mbox.js"></script>

<link rel="stylesheet" type="text/css" href="/_layouts/1033/styles/ISACACPE/CpeCycleSummaryProgress.css?rev=J4c615531PUA1RAFRhp7nQ%3D%3D" /><link
rel="stylesheet" type="text/css" href="/_layouts/1033/styles/ISACACPE/CpeDashboard.css?rev=4J%2FX3z4tmU6Tan1dxm039Q%3D%3D" /><link rel="stylesheet"
type="text/css" href="/_layouts/1033/styles/ISACACPE/Form.css?rev=843zq671I5zQ9zaPeOVLg%3D%3D" /><link rel="stylesheet" type="text/css"
href="/_layouts/1033/styles/HtmlEditorCustomStyles.css?rev=8SKxtNX33FmoDhbbfB27UA%3D%3D" /><link rel="stylesheet" type="text/css"
href="/_layouts/1033/styles/HtmlEditorTableFormats.css?rev=guYGdUBUxQit03E2jh5dvA%3D%3D" /><link rel="stylesheet" type="text/css"
href="/_layouts/1033/styles/ISACACPE/CpeDetails.css" /><link rel="stylesheet" type="text/css" href="/_layouts/1033/styles/susqtech/controls.css"
/><link rel="stylesheet" type="text/css" href="/_layouts/1033/styles/susqtech/core.css" /><link rel="stylesheet" type="text/css" href="/Style
Library/ISACADefault.css" /><link rel="stylesheet" type="text/css" href="/Style Library/ISACAMaster-MyISACA.css" /><link rel="stylesheet"
type="text/css" href="/Style Library/ISACASearchBox.css" />

<!-- Include Print CSS -->
<link rel="stylesheet" href="/Style Library/ISACADefault-Print.css" type="text/css" media="print" />
<link rel="stylesheet" type="text/css" href="/Style%20Library/ISACALayout-MyISACADashboardLanding.css" />
<link rel="stylesheet" type="text/css" href="/Style%20Library/ISACALayout-Home.css" />

<script type="text/javascript" src="/_layouts/1033/CustomScripts/namefix.js" defer="true"></script>
<script type="text/javascript" src="/_layouts/1033/SusQtechScripts/dsfix.js" defer="true"></script>
<script type="text/javascript" src="/_layouts/1033/CustomScripts/dsfix.js" defer="true"></script>
<script type="text/javascript" src="/_layouts/1033/CustomScripts/wpfix.js" defer="true"></script>
<script type="text/javascript" src="/_layouts/1033/SusQtechScripts/namefix.js" defer="true"></script>
<!--[if gte IE 7]>
<script type="text/javascript" src="/Style Library/scripts/ISACADefault-IE-min.js"></script>
<!--[endif]-->
<!--[if lt IE 7]>
<script type="text/javascript" src="/Style Library/scripts/ISACADefault-IE-min.js"></script>
<script defer="defer" type="text/javascript" language="javascript" src="/Style Library/scripts/pngfix-min.js"></script>
<script type="text/javascript" src="/Style Library/scripts/iepngfix_tilebg-min.js"></script>
<link rel="stylesheet" type="text/css" href="/Style Library/ISACADefault-IE.css" />
```

BTSF

Saving web pages using third parties

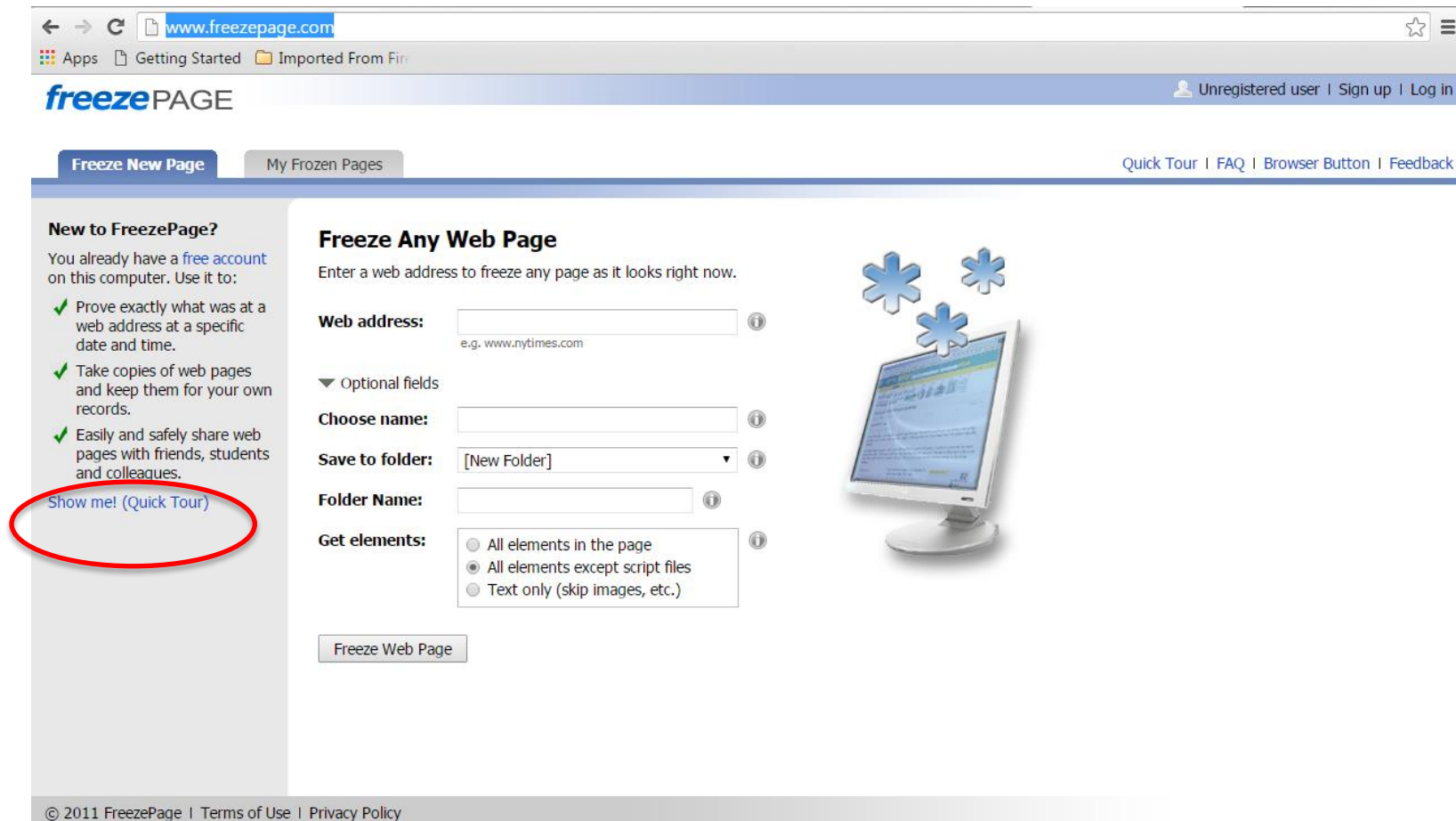
There are several tools or sites that offer the possibility to save web pages.
Useful when evidence can be compromised using manual approach.

Example:

- <http://www.freezepage.com/>
- <https://web.safestamper.com/> (not free)
- <https://archive.org/web/>

BTSF

BTSF Saving web pages using third parties



The screenshot shows the FreezePage website interface. The browser address bar displays www.freezepage.com. The page header includes the "freezePAGE" logo and user status: "Unregistered user | Sign up | Log in". Navigation tabs are labeled "Freeze New Page" and "My Frozen Pages". A secondary navigation bar contains links: "Quick Tour | FAQ | Browser Button | Feedback".

New to FreezePage?
You already have a [free account](#) on this computer. Use it to:

- ✓ Prove exactly what was at a web address at a specific date and time.
- ✓ Take copies of web pages and keep them for your own records.
- ✓ Easily and safely share web pages with friends, students and colleagues.

[Show me! \(Quick Tour\)](#)

Freeze Any Web Page
Enter a web address to freeze any page as it looks right now.

Web address:

▼ Optional fields

Choose name:

Save to folder:

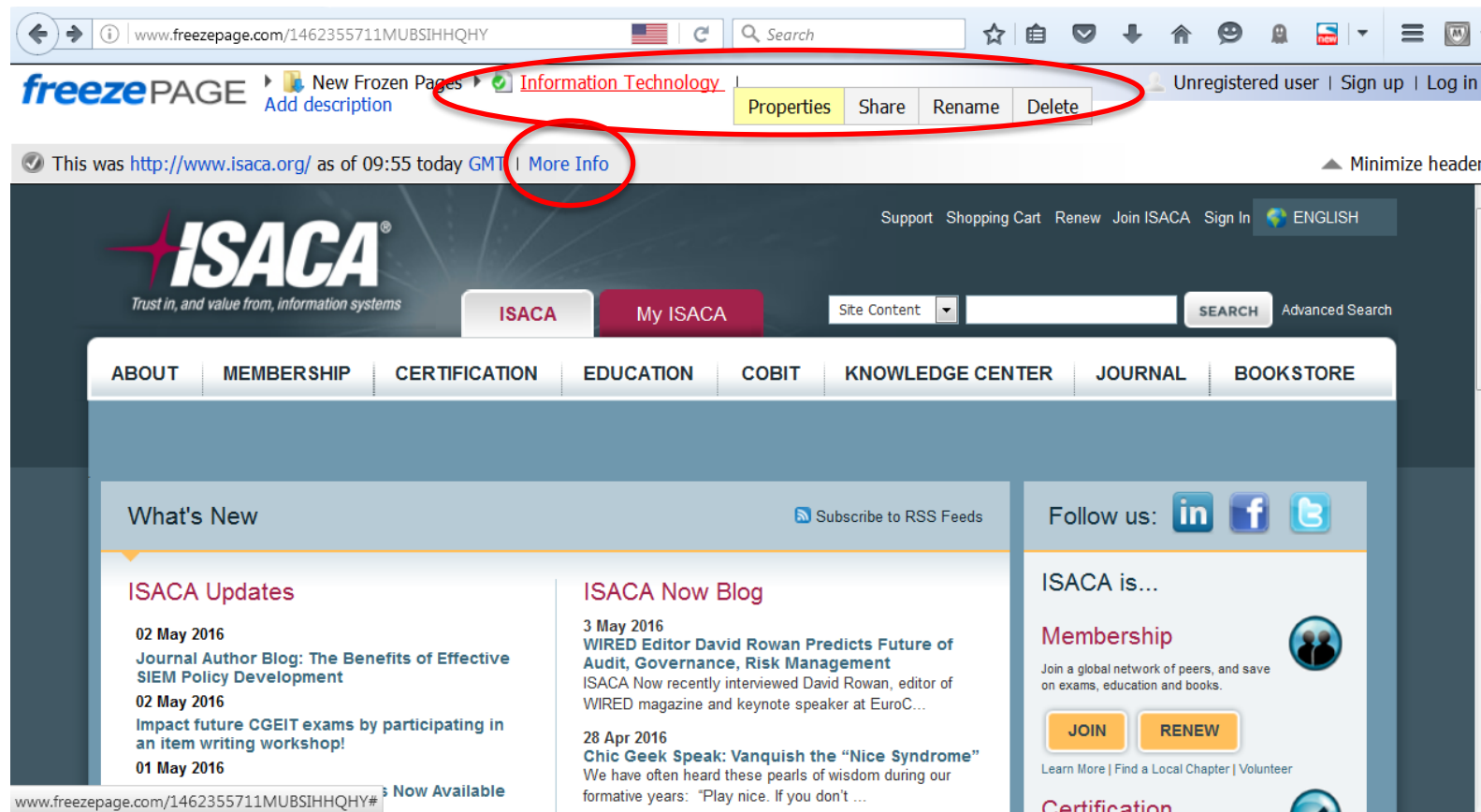
Folder Name:

Get elements:

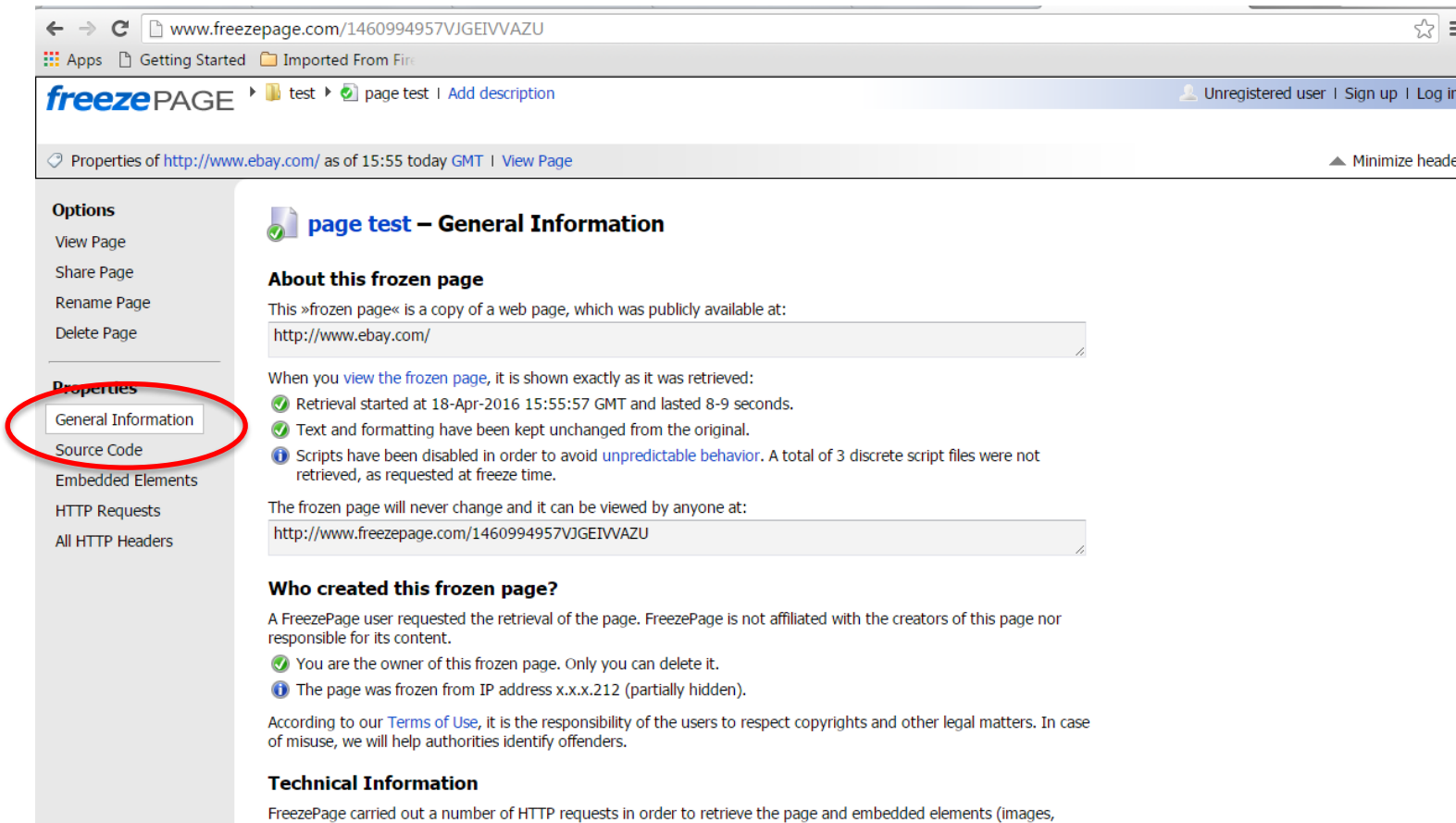
- ☐ All elements in the page
- ☒ All elements except script files
- ☐ Text only (skip images, etc.)

© 2011 FreezePage | [Terms of Use](#) | [Privacy Policy](#)

BTSF Saving web pages using third parties



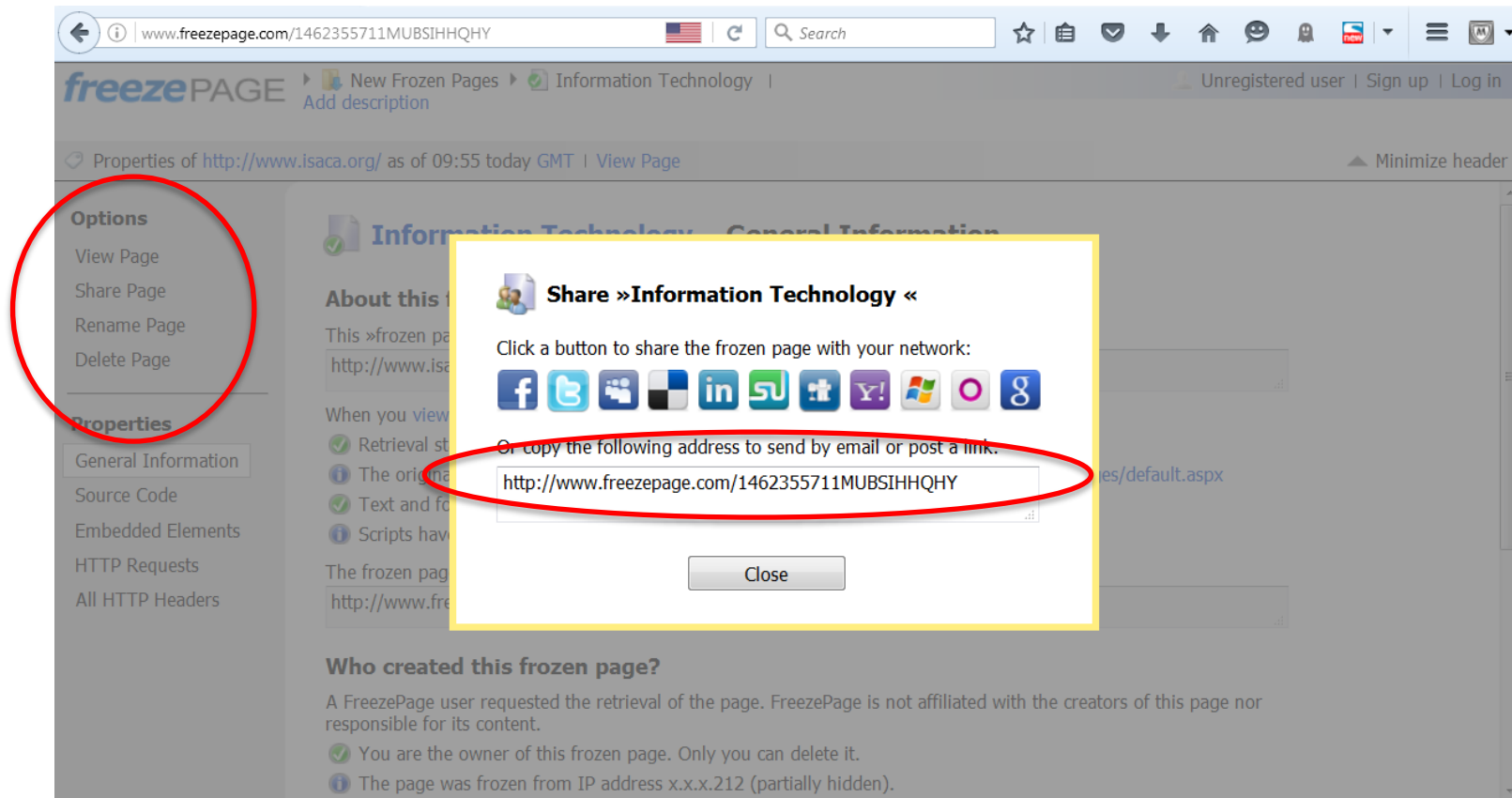
BTSF Saving web pages using third parties



The screenshot shows a web browser window with the URL www.freezepage.com/1460994957VJGEIVVAZU. The page title is "page test" and it is categorized as "page test". The user is logged in as "Unregistered user". The page is a "Properties" page for a frozen page, with the "General Information" tab selected. The "General Information" tab is highlighted with a red circle. The page content includes:

- Options:** View Page, Share Page, Rename Page, Delete Page.
- Properties:** General Information (selected), Source Code, Embedded Elements, HTTP Requests, All HTTP Headers.
- page test – General Information**
- About this frozen page**
 - This »frozen page« is a copy of a web page, which was publicly available at: <http://www.ebay.com/>
 - When you [view the frozen page](#), it is shown exactly as it was retrieved:
 - ✓ Retrieval started at 18-Apr-2016 15:55:57 GMT and lasted 8-9 seconds.
 - ✓ Text and formatting have been kept unchanged from the original.
 - ⓘ Scripts have been disabled in order to avoid [unpredictable behavior](#). A total of 3 discrete script files were not retrieved, as requested at freeze time.
 - The frozen page will never change and it can be viewed by anyone at: <http://www.freezepage.com/1460994957VJGEIVVAZU>
- Who created this frozen page?**
 - A FreezePage user requested the retrieval of the page. FreezePage is not affiliated with the creators of this page nor responsible for its content.
 - ✓ You are the owner of this frozen page. Only you can delete it.
 - ⓘ The page was frozen from IP address x.x.x.212 (partially hidden).
- According to our [Terms of Use](#), it is the responsibility of the users to respect copyrights and other legal matters. In case of misuse, we will help authorities identify offenders.
- Technical Information**
 - FreezePage carried out a number of HTTP requests in order to retrieve the page and embedded elements (images,

BTSF Saving web pages using third parties



F

BTSF Saving web pages using third parties

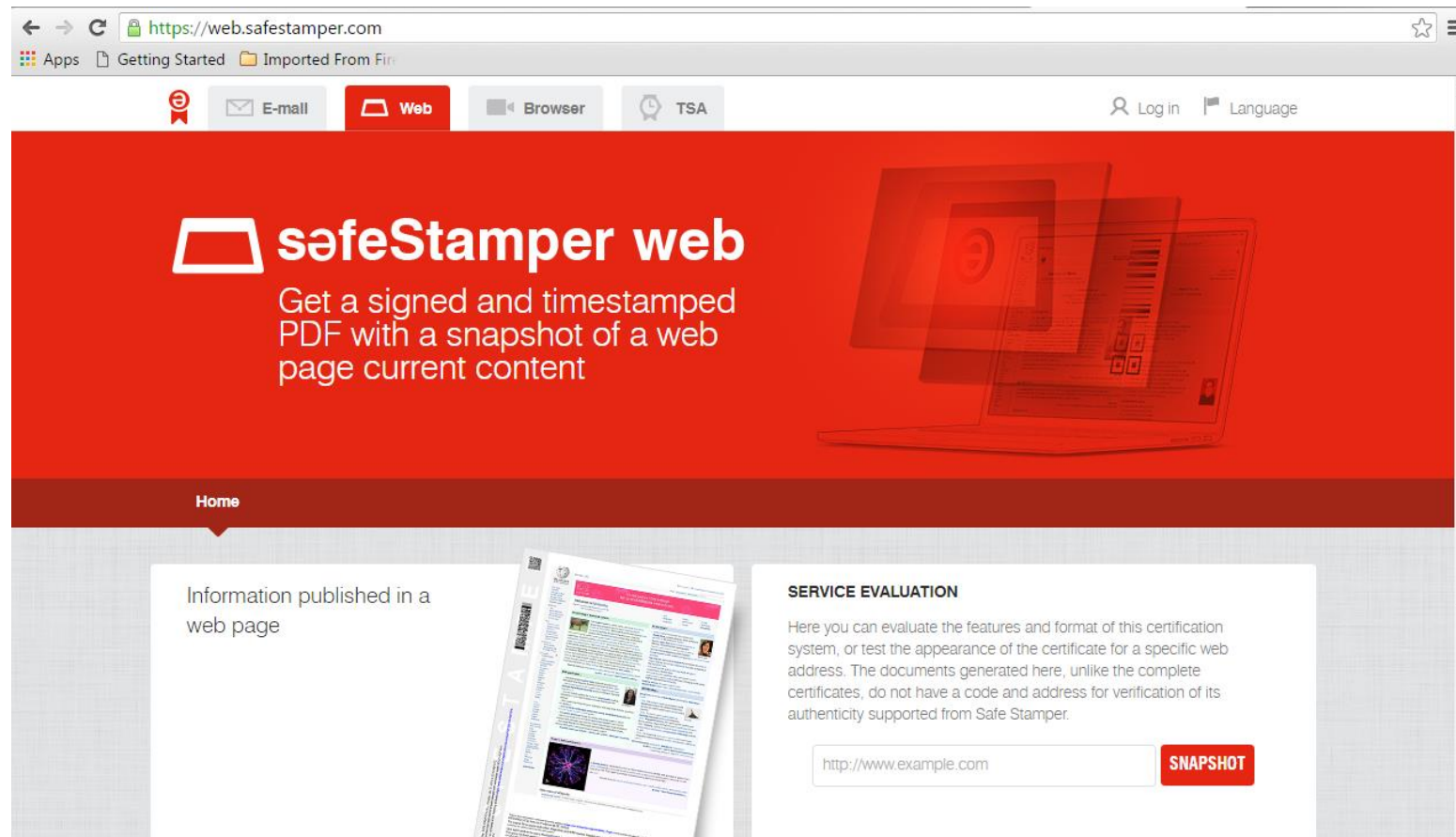
The screenshot shows the FreezePage website interface. The browser address bar displays 'www.freezepage.com/user'. The website header includes the 'freezePAGE' logo and navigation links for 'Unregistered user', 'Sign up', and 'Log in'. Below the header, there are tabs for 'Freeze New Page' and 'My Frozen Pages', along with links for 'Quick Tour', 'FAQ', 'Browser Button', and 'Feedback'.

The main content area is divided into two sections. On the left, under 'Registration Benefits', it states 'Registration is free and gives you:' followed by three bullet points: 'Mobility' (Log in and access your frozen pages from anywhere on the Internet), 'More Space' (Unregistered users have 10 MB available for their account, whereas members get 20 MB of free storage space), and 'More Time' (Unregistered accounts are closed after three days of inactivity, whereas registered users only need to log on every month).

The right section, titled 'Sign Up With FreezePage', is circled in red. It contains the following text: 'If you are new to FreezePage, please sign up here. All data can be modified later if necessary. If you have already signed up, [log in here](#).' Below this text are input fields for 'Email address:', 'Reenter email:', 'Choose password:', and 'Reenter password:'. Each field has an information icon (i) to its right. Under the 'General preferences' section, there are dropdown menus for 'User ID (public):' (set to 'U357001'), 'Time format:' (set to 'Select a time format...'), and 'Time zone:' (set to 'Greenwich Mean Time (GMT)'). There is also a 'This device:' section with two radio buttons: 'Keep me logged in on this device' (selected) and 'Log me out when I close the browser'. A 'Sign Me Up' button is located at the bottom of the registration form.

At the bottom of the page, the footer text reads: '© 2011 FreezePage | Terms of Use | Privacy Policy'.

BTSF Saving web pages using third parties



BTSF Saving web pages using third parties

Signed by SAFE CREATIVE, S.L. On date: May 4, 2016 10:35:00 AM UTC
This is a test document generated for evaluating Safe Stamper web page certification service
so it does not have a validation service.
Page 1 of 2

This is the information retrieved from the address <http://www.isaca.org> when queried on May 4, 2016 10:35:00 AM (UTC) from the IP address 54.247.118.247.

The page source code (index.html) has been included as attachment to this document.

User agent used on the query: Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko)

This query has been performed by web.safestamper.com service from Safe Creative Ltd. which records and digitally signs its result in this PDF document on May 4, 2016 10:35:00 AM (UTC)

Page 1 of 2

BTSF

Saving web pages using third parties

[Adobe Acrobat](#) and [Microsoft Office](#) have built-in time stamping and digital signature facilities. They are likely to meet the requirements of an advanced electronic signature.

Further information about time stamping, digital signature and providers in Europe:

<https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/eSignature>

BTSF

Capturing the Internet history: The Internet Library

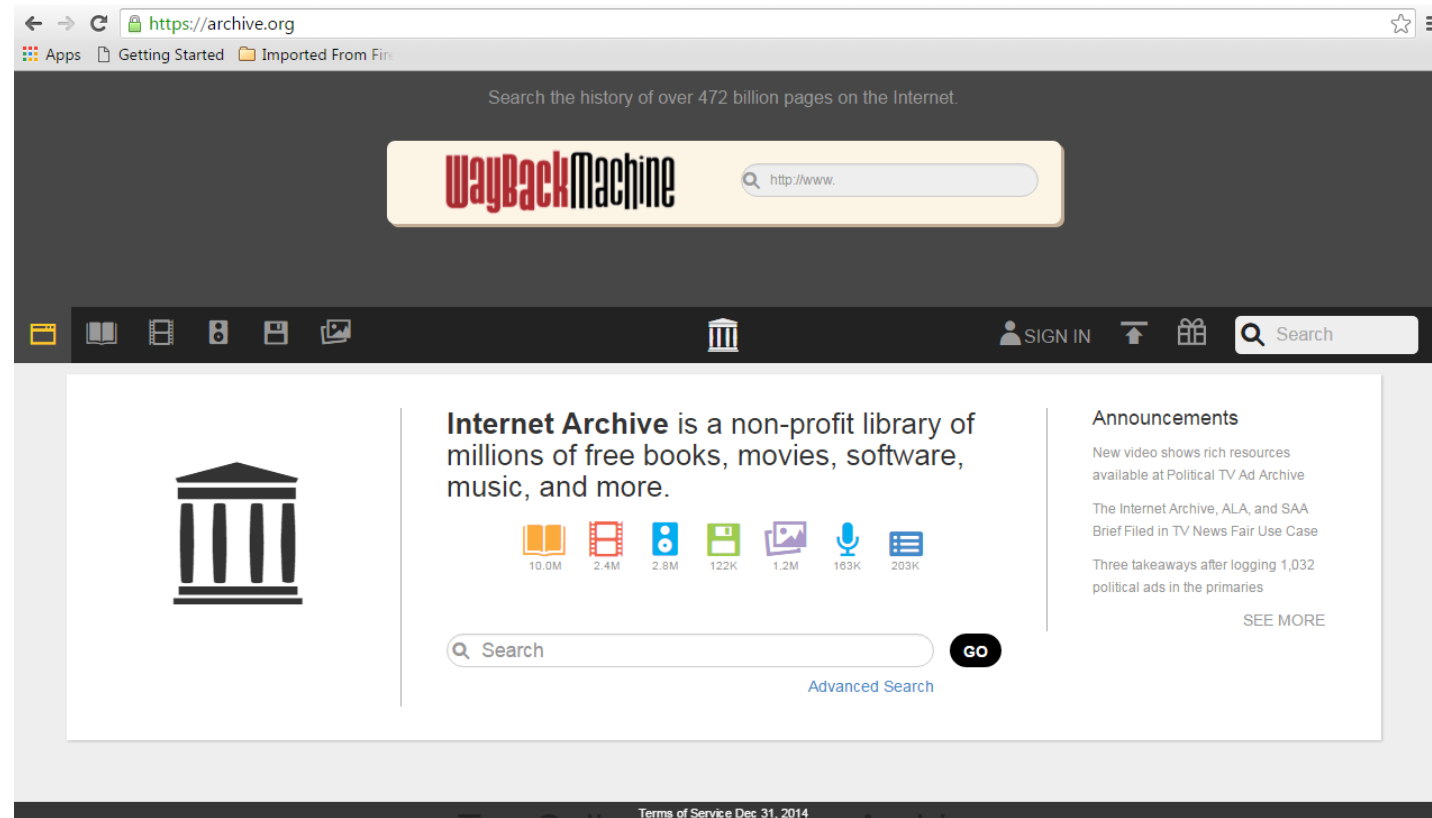
- The Internet Archive is a non-profit organization founded to build an Internet library.
- Its purposes include offering permanent access for researchers, historians, scholars, people with disabilities, and the general public to historical collections that exist in digital format.
- Now the Internet Archive includes: texts, audio, moving images, and software as well as archived web pages.

BTSF

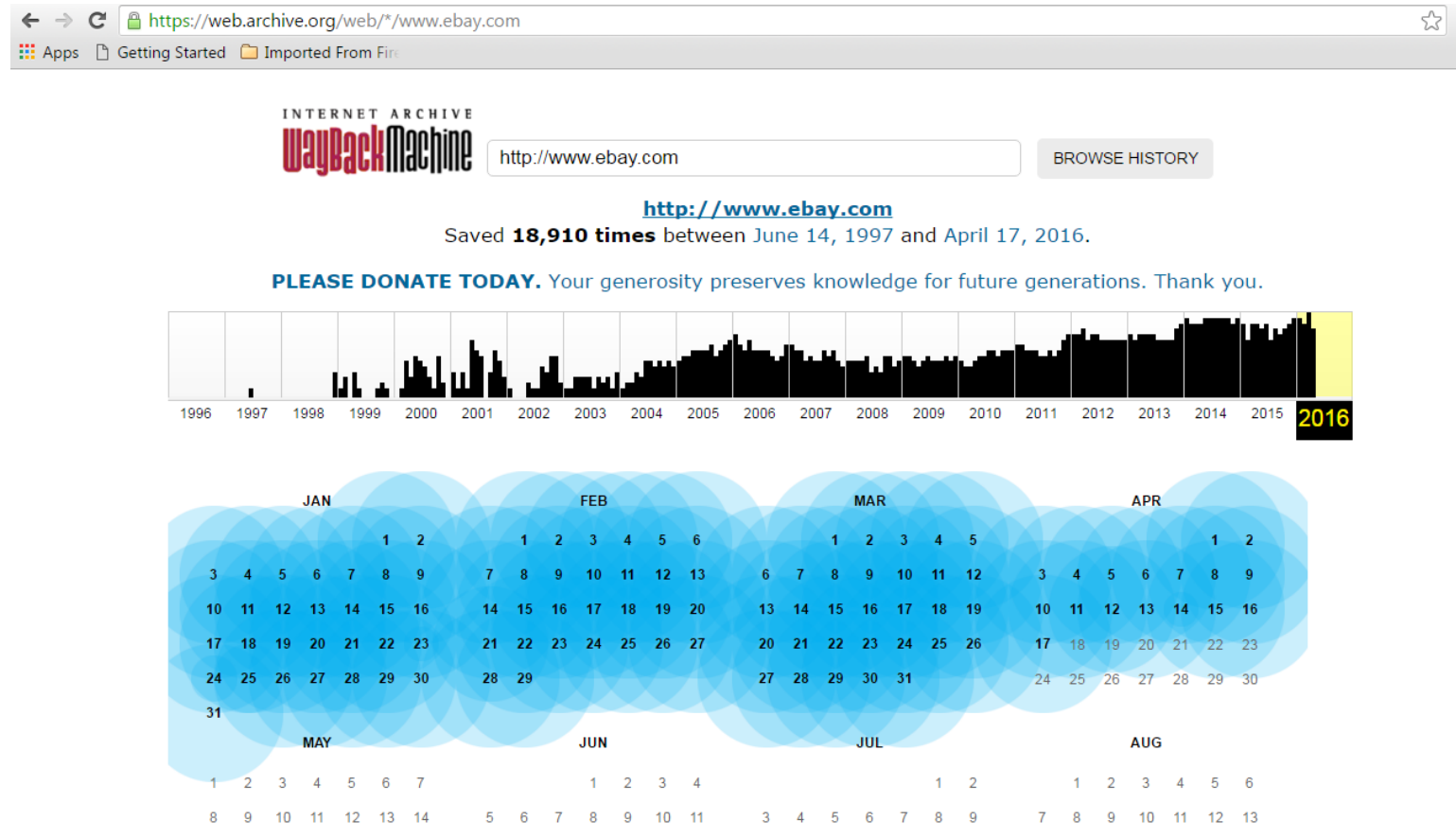
Capturing the Internet history: The Internet Library

- Basically, the Internet Archive saves copies of web pages published in different times.
- We can access to different views of the same site in different times to compare them.
- Useful in investigations where the history can provide useful evidence.
- The Internet Archive web engine doesn't index all websites everyday, so this is not a complete compilation of every piece of content upload into the Internet.

BTSF The Internet Library: archive.org



BTSF The Internet Library: archive.org



BTSF

Capturing a whole website

- Sometimes you may want the whole site.
- Only recommended for small sites.
- Automated tools work their way around a site, saving the data for you.
- These are called “offline browsers”.
- HTTrack is a free tool.
- Depending on the website size it's a time consuming process.

BTSF

Capturing a whole website - **WARNING**

- Website copying is not recommended in all cases.
- If it is noticed that a user/connection is going systematically through all the pages from a site within a short time period, this may look more suspicious than accessing occasionally some pages from an anonymous IP over a longer time period.
- Some websites may block automatic web crawling. In these cases this tool will not work.

BTSF

Capturing a whole website - **WARNING**

- Downloading a site can overload it, which may raise suspicions.
- To avoid this, follow these simple guidelines:
 - Do not use too many simultaneous connections.
 - Configure the depth of downloading.
 - Use size limits.
 - Use time limits.
 - Try to not download during working hours.
 - Check your mirror transfer rate/size.

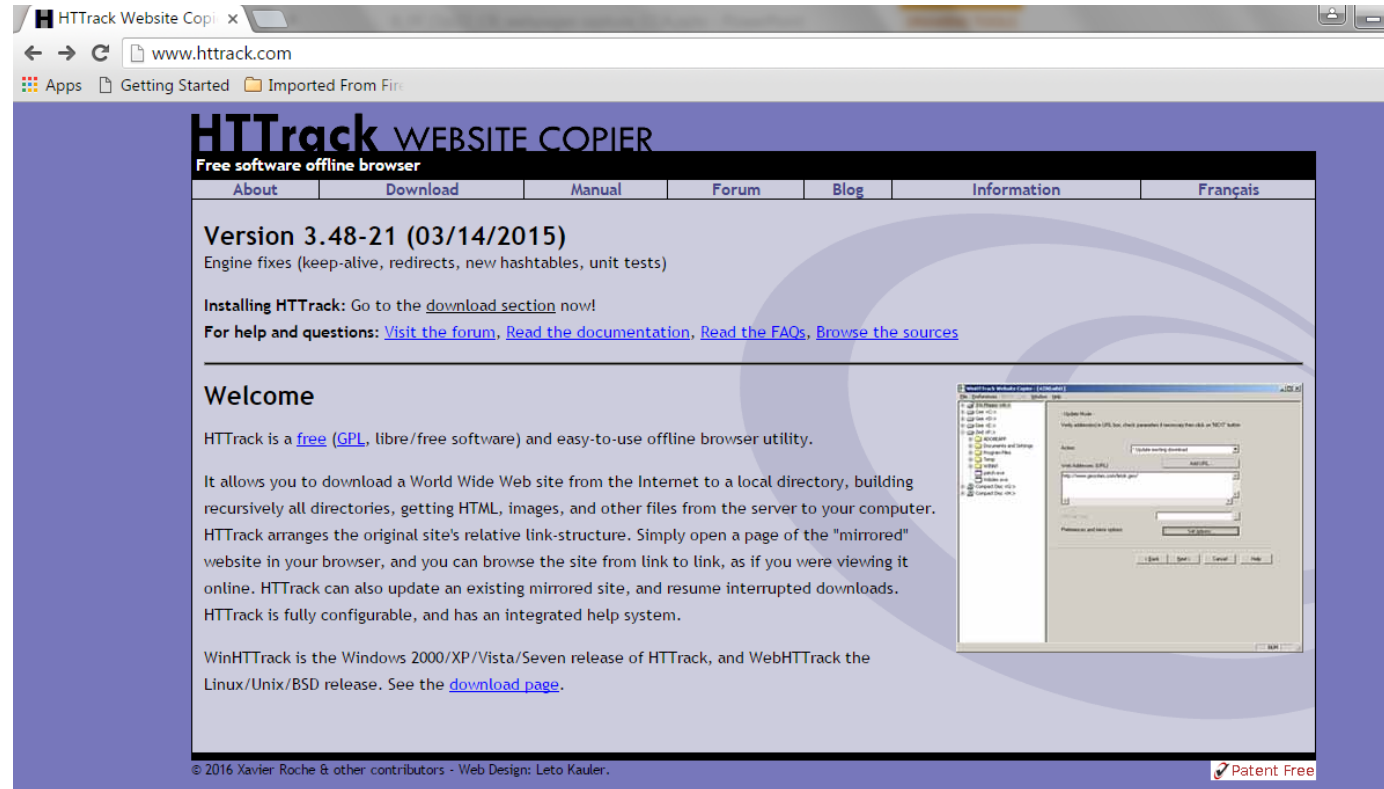
BTSF

Capturing a whole website - **WARNING**

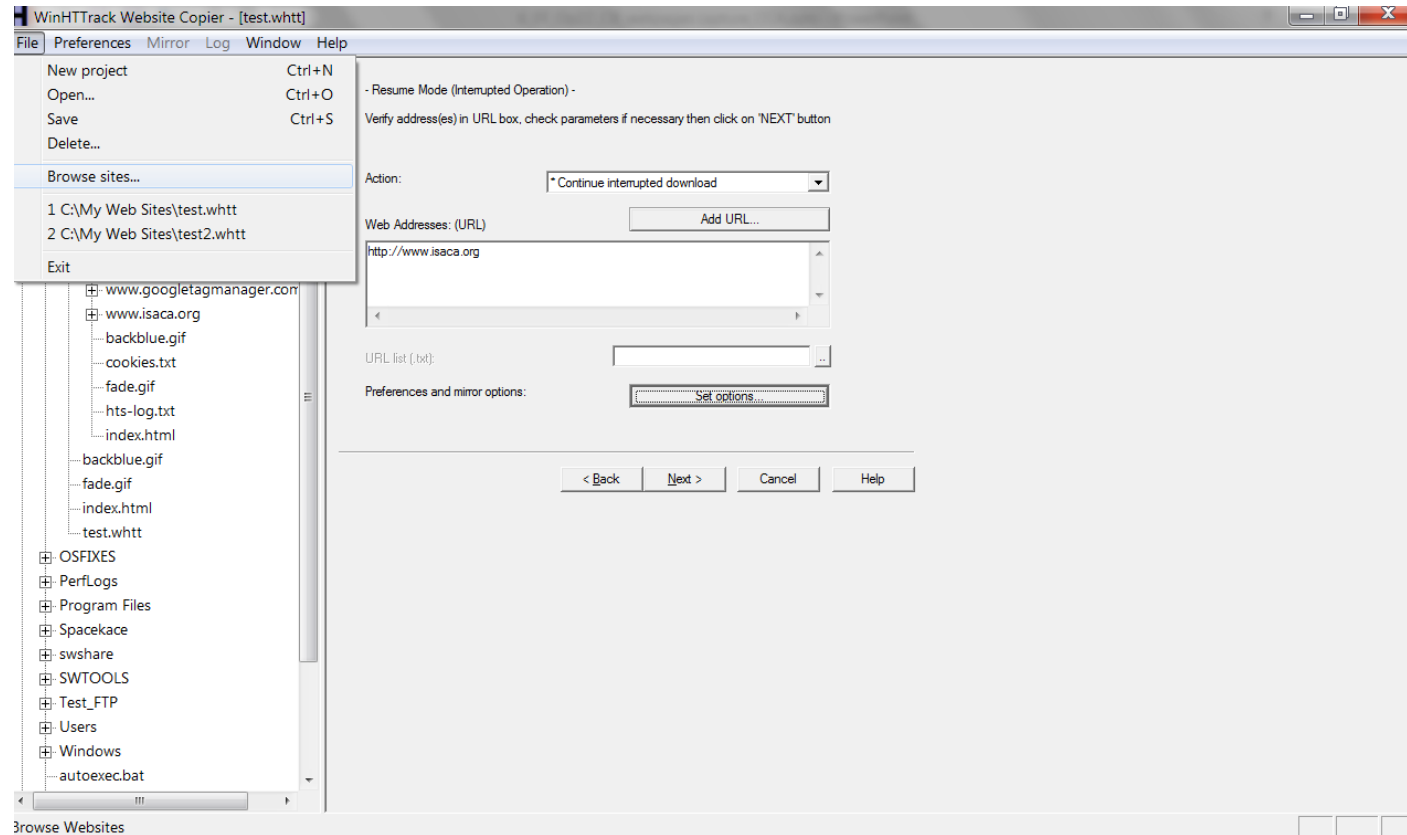
- Also, make sure that you can copy the website according to your national legislation:
 - Are the pages copyrighted?
 - Can you copy them only for private purpose?

BTSF

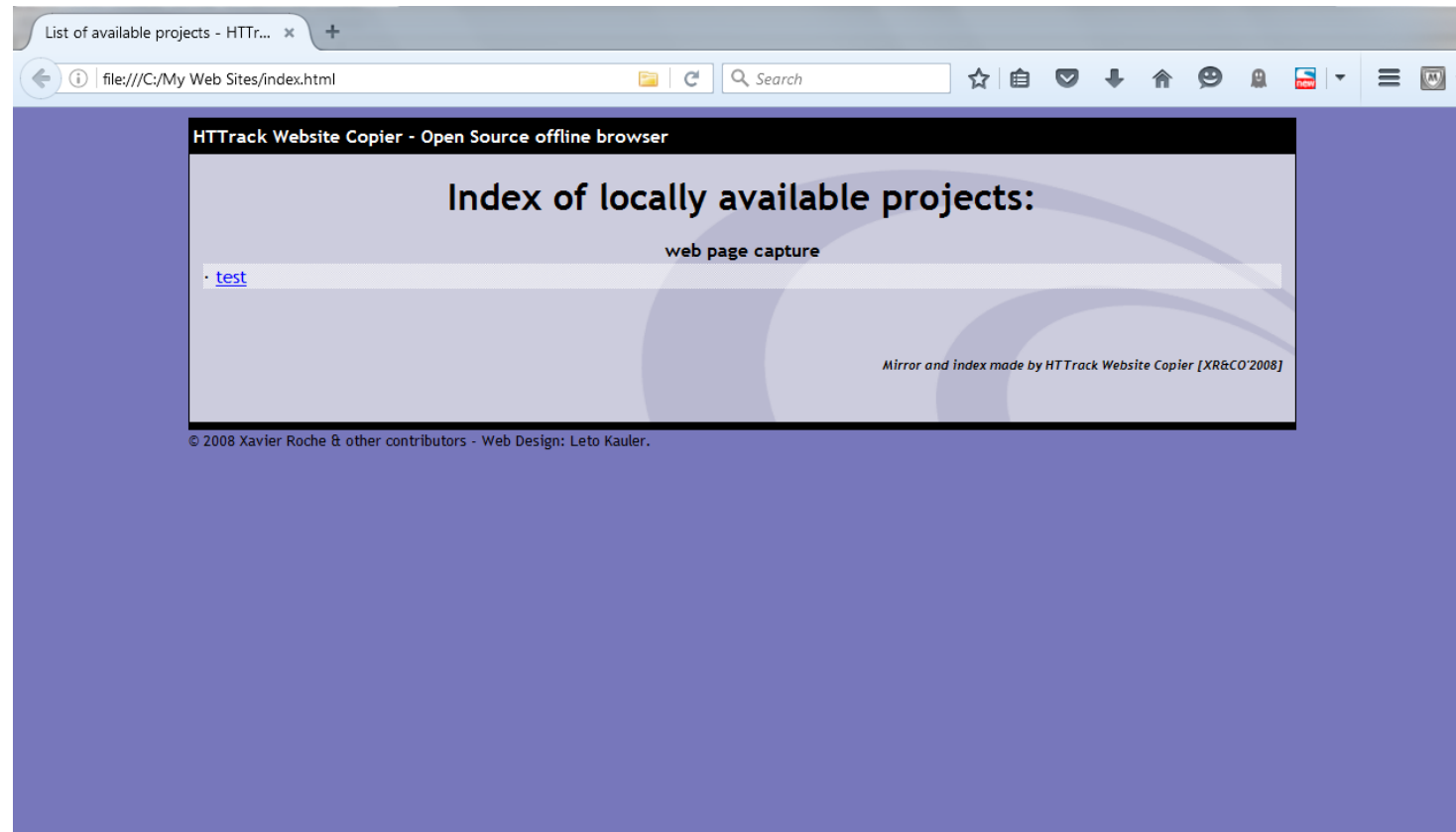
BTSF Exercise: Saving complete websites using an offline browser



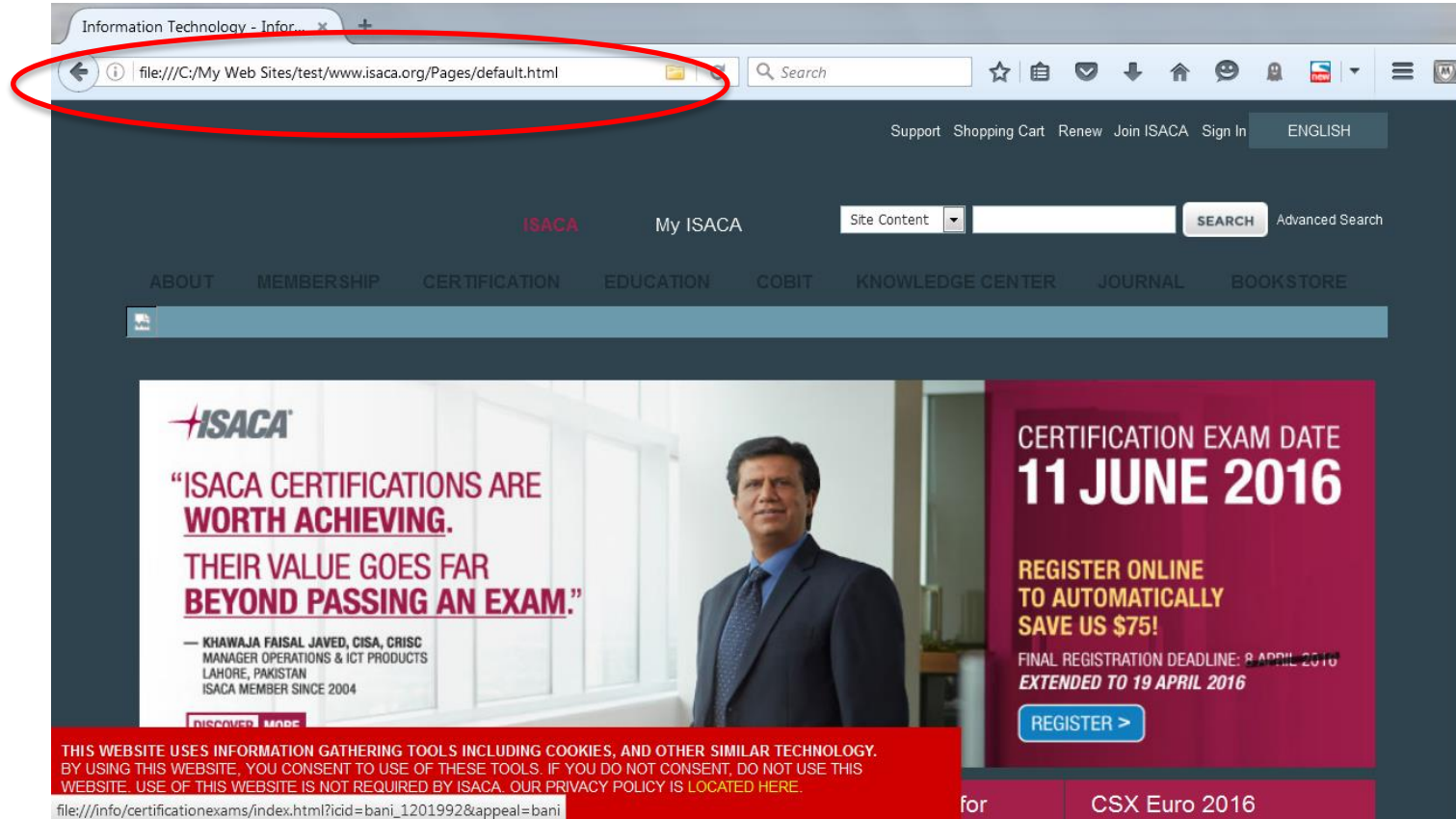
BTSF Exercise: Saving complete websites using an offline browser



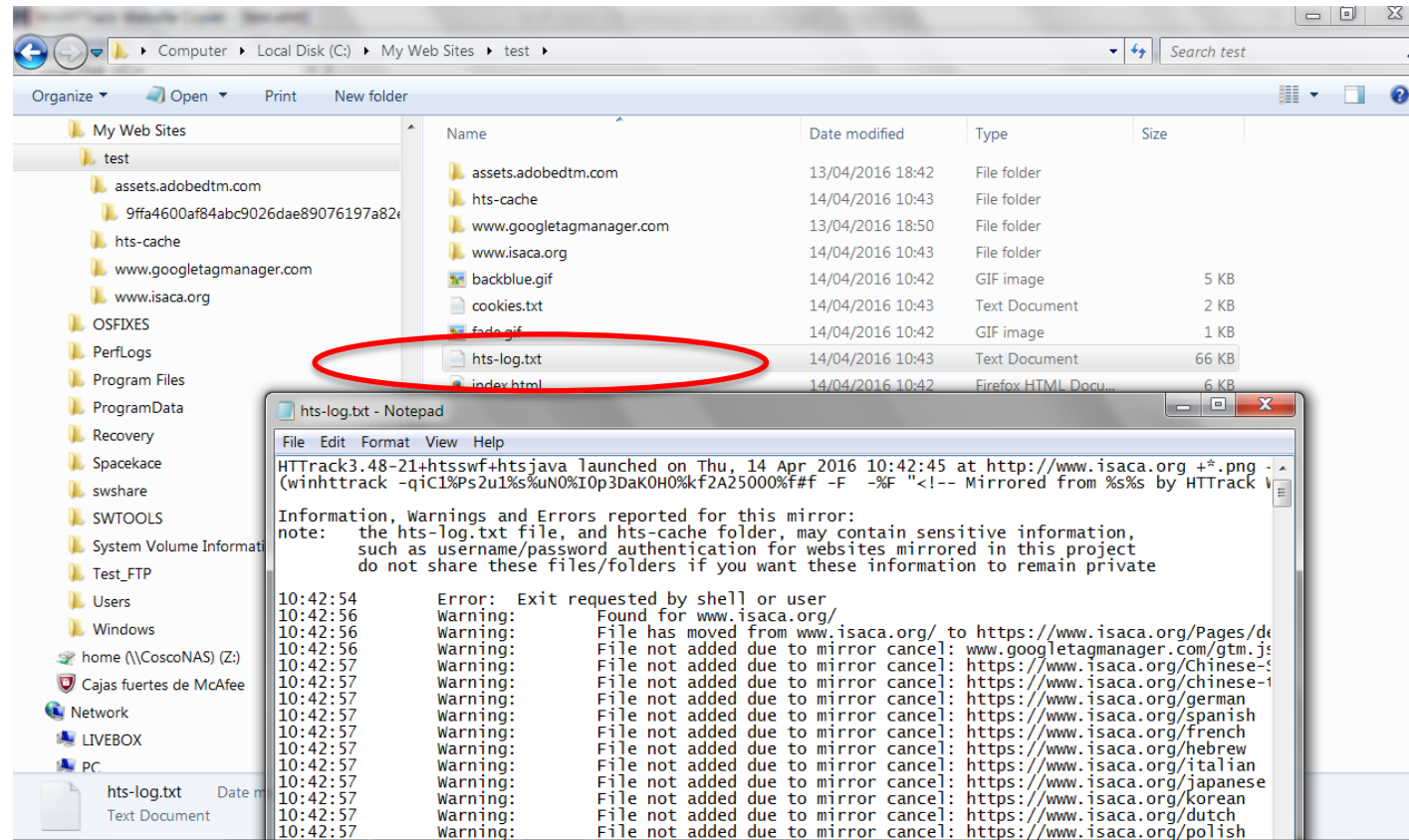
BTSF Exercise: Saving complete websites using an offline browser



BTSF Exercise: Saving complete websites using an offline browser



BTSF Exercise: Saving complete websites using an offline browser



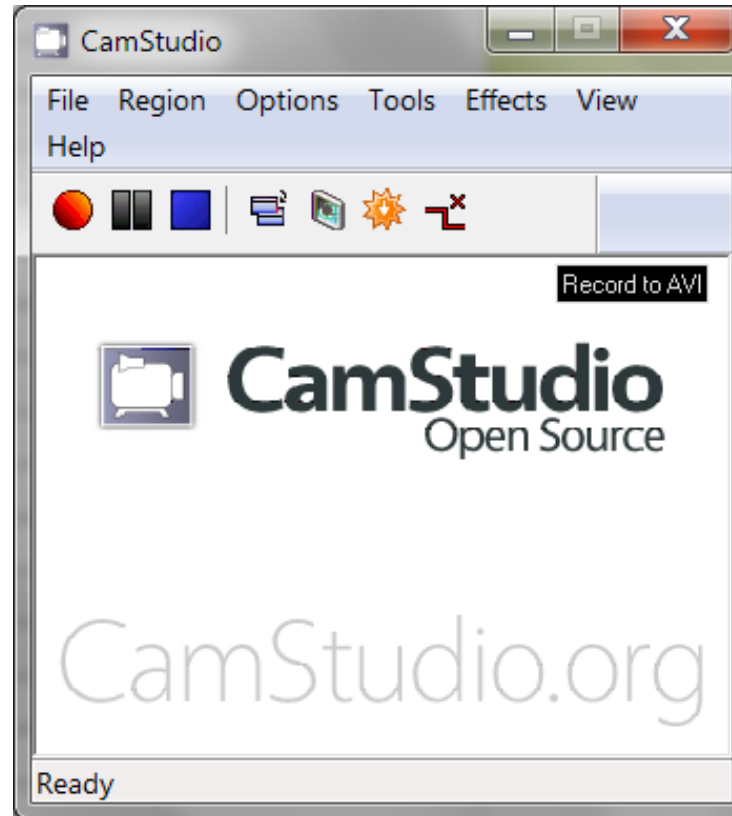
BTSF

Creating Video Captures

- Creating videos of your online activity is very useful
- It can include sound, animation or other details (pop-ups, etc.)
- Video capture software is easily available
- It is recommended to evidence the time/date of the recording.
 - **Tip:** check an official time reference online at the beginning of your investigation video recording

BTSF

Exercise: Setting up CamStudio and creating a video capture



TSSF

BTSF

Exercise: Setting up Hypercam and creating a video capture

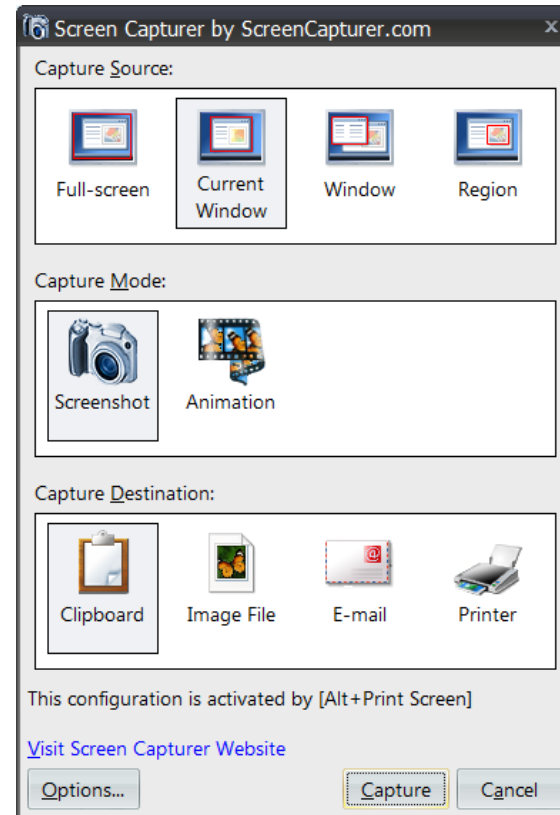
**HYPER
CAM**
SCREEN RECORDER



BTSF

BTSF

Exercise: Setting up Screen Capturer and creating a video capture



BTSF

Things to remember



Most web pages dynamically change their content, so design your strategy to preserve them.

- Manual preservation:
 - Using browser capabilities to save them.
- Using third party tools. They add extra valuable information about the time and date the web page was captured. They help us to custody the evidence.
- Video captures. Valuable to document the entire browsing session and describe with fully detailed all the work done.

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](#)