ethical web dev

EDRi

EUROPEAN DIGITAL RIGHTS

GUIDE FOR ETHICAL WEBSITE DEVELOPMENT AND MAINTENANCE

Security is actually quite simple, cleaning up after a breach is complicated.

Your website visitors deserve respect. This means not taking shortcuts that lead to their data being leaked to third parties and it means ensuring adequate security.

Respecting your visitors also means taking a big step towards compliance with data protection legislation.

This publication is a result of an extensive collective work, with inputs from experts of the EDRi network (Anders Jensen-Urstad, Walter van Holst, Maddalena Falzoni, Hanno "Rince" Wagner, Piksel), external contributions (Gordon Lennox, Achim Klabunde, Laura Kalbag, Aral Balkan), and the crucial help of Sid Rao, Public Interest Technologist and ex-Ford-Mozilla Fellow at EDRi. Special thanks to Joe McNamee who had the original idea for this booklet and steered the process to a successful conclusion.

Project coordination: Guillermo Peris, Community Coordinator, EDRi

Layout: Heini Järvinen, Senior Communications Manager, EDRi

Printed thanks to the support of the European Cultural Foundation

European Cultural Foundation

Contents

0. Introduction **2**

1. The Basics **3**

Expectations 3
Trustworthiness and Security 3
Personal data and compliance with the GDPR 3

2. General recommendations **5**

3. Security recommendations **6**

ISO 6
DNSsec 6
Tor 7
HTTPS 7
CSP 7
JavaScript 8
Securing sensitive data 8
Protection against DDoS attacks 9
Static websites are back 9

4. Alternatives to costly "free" third-party services **10**

Analytics 10
Videos 10
Maps 11
Fonts and icons 11
Social widgets 12
CAPTCHAs 12
And more 13

Event organisation tools 13
Making websites accessible for all 14
Protect the web from untangling 14
Ethical advertising 14
Website search 14

5. Glossary **15**

0. Introduction

A website is almost like a living thing. Most of the time the basic site itself is not static and in addition to its own dynamic features, its environment is also subject to continuous change which in turn leads to even more changes.

Visitors of a website can also be very diverse. The technologies they use and their expertise may vary widely.

Many websites themselves also rely on a variety of external services and resources. These also continue to evolve.

As website developers have to cope at the same time with the increasing expectations of users and the limited resources most organisations devote to website development, there is a growing tendency to use more external services and resources.

For example, it has become more and more common for web developers to take "free" resources, such as fonts and scripts and use them on the websites that they design. While these are "free" for the developer, they can have undesirable side effects for the users and the organisations that provide the website. For example some resources and services, particularly those provided by certain data hungry internet companies, can undermine user privacy. Others can have adverse affects on security. In both cases, the reputation of the website owner may suffer, or it may even face legal challenges.

This warrants attention. However, there is a general lack of awareness of this problem, and these practices have already become quite pervasive. The purpose of this document is to clarify the problems and, where possible, identify some usable solutions.

This guide is aimed at web developers and maintainers who have a strong understanding of technical concepts. Links are provided to background information, where necessary, in order to keep the document brief and maximise its usability.

We hope that this will assist developers and maintainers in bringing the web back to its roots – a decentralised tool that can enhance fundamental rights, democracy and freedom of expression.

It is also our internet. We all need to take responsibility.

1. The Basics

EXPECTATIONS

People who visit a website have a range of legitimate expectations.

They want to be safe. They do not want any harm to come to their own equipment or the data on their equipment. They want their privacy to be respected. Some of these requirements are codified in legislation such as the EU's General Data Protection Regulation (GDPR).

Previously, this was simply about a service being technically simple and reliable. But a service that is a bit broken, a bit clunky, a bit unpredictable, means that users have to accept that it does not meet their general expectations. They may not continue to use the site. More importantly in this context, however, if they decide to continue using the service, they will learn to tolerate things that maybe should not be tolerated and that are really signs of significant problems.

TRUSTWORTHINESS AND SECURITY

In the current social and political climate, many are also concerned about correctness and truthfulness. How readily should a user trust software downloaded from a site? How readily should they trust links provided on a site?

Finally there is the old catchall: security. If safety is about trying to make sure a system does not do harm, security is about protecting a system from harm, protecting a system from accidents and attacks. So we have the triad: availability, integrity and confidentiality. Systems, services and data need to remain available despite component failures and distributed denial of service (DDoS) and other attacks. Measures need to be taken to preserve service and data integrity. Is a damaged database really still useful as a database? Data should of course only be visible to those who have the appropriate rights.

PERSONAL DATA AND COMPLIANCE WITH THE GDPR

Organisations in (or targeting the citizens of) the European Union have to comply with the European Union's General Data Protection Regulation (GDPR). This Regulation on data protection and privacy for all individuals in the EU (or by EU companies) aims primarily to give control to individuals over their personal data and simplifies the regulatory environment. It also addresses the

export of personal data to jurisdictions outside the EU.

The GDPR is however a complex piece of legislation. There are many companies offering consultancy, training courses, interpretive guides and checklists. In addition, each EU Member State has its own data protection authority which has responsibilities when it comes to interpretation and enforcement.

Personal Data (also referred to, with somewhat different meaning, as Personal Identifiable Information or PII, in the United States) is any information relating to an identified or identifiable natural person ('data subject'). An identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.

According to the GDPR, an IP address is personal data. This is because an IP address may be connected to an individual if it is combined with other information. When third-party solutions are used to provide functionality or content such as JavaScript and images, a website's visitor's IP address may be available to those third party providers. This is why, under the GDPR, a data processing agreement with each of those third party providers may be needed.

Privacy by default means that, where you provide a service with a variety of levels of data processing, the default setting should be the one that most respects privacy. The user can then be given the option to change this setting, should they wish. This needs to be done in the most user-friendly way possible.

In any case, data subjects retain rights over their personal data and organisations have specific responsibilities in the event of security incidents or data breaches.

GDPR compliance is therefore a specialist area on its own. However, following the principles in this document will be a step towards legal compliance as well as ensuring an ethical approach. In addition, there are plenty of genuinely free resources available to help with ensuring compliance – which is good for you and good for your website visitors. See, for example, the European Data Protection Supervisor's guidelines for web services:

https://edps.europa.eu/data-protection/our-work/publications/guidelines/web-services_en

2. General recommendations

- Allow as much data processing on an individual's device as possible.
- Where you must deal with user data, **use encryption**. Facilitate end-to-end encryption for all relevant communications. The aim there is that you simply cannot see private information.
- Where possible also **use data minimisation methods** – only ever process data that actually needs to be processed.
- The best solution is often to **use first-party resources** (i.e. your resources that you host) and **avoid third-party solutions** as much as possible. In other words, try to host everything on your own server. This includes third party code and content such as:
  - Cookies
  - CSS files
  - Images
  - Media such as video files, audio files
  - JavaScript, if you decide to use it
  - Frames with third-party content
  - Font files, if you need them
- If downloading a resource, such as a JavaScript or font file, is not allowed by the terms of its provider, then they may not be privacy-friendly and should therefore be avoided. If you do have to load third-party resources, then the use of **Subresource Integrity (SRI)** can be a good idea.

3. Security recommendations

Security is a process, and it needs to take into consideration the entire stack. Security decisions need to take account of differing, potentially contradictory, objectives, and they cannot be taken in isolation. Security related decisions may affect such features as the speed and user-friendliness of a service and have a significant impact on the resources needed, i.e. money and working time spent.

Generally speaking, you ought to be able to achieve the following:

- the site ought not to be a source of harm for the rest of the internet;
- the integrity of the site itself should be maintained;
- all communications should be secure;
- the security and privacy of visitors should be protected.

Achieving these objectives tends to rest on two pillars: sharing and standards. As threats and responses change it is important to share information, such as via specialist mailing lists. At the very least one should be aware of one's local Computer Emergency Response Team (CERT) / Computer Emergency Readiness Team and Computer Security Incident Response Team (CSIRT). In addition, adopting appropriate current standards ensures that you are not the weak link in any chain.

ISO

ISO has published a variety of standards which could be of interest, including the following on quality and security:

- ISO / IEC / IEEE 90003:2018 Software engineering
- ISO / IEC 27000 family of Information security standards

These standards provide, among other things, a common vocabulary which can be very useful. One place they can be useful is in demonstrating GDPR compliance.

DNSSEC

Authenticating responses to DNS queries regarding your domain name is a good thing.

TOR

In order to protect their privacy and security, people may prefer to use the Tor Browser. When choosing a hosting provider for your application, make sure Tor connections are allowed. For example, using the **Onionshare** tool, one can easily publish an anonymous and uncensorable version of the exisiting website with only a few clicks.

https://onionshare.org/

HTTPS

To avoid middleman attacks and eavesdropping, do not allow unencrypted connections to your website and always encourage connection through **HTTPS**. The principal motivation for HTTPS is authentication of the accessed website and protection of the privacy and integrity of the exchanged data. Consider using **HSTS – HTTP Strict Transport Security** - to enforce this.

If you have root access on the server – it is your server - you can use **Let's Encrypt** and get a certificate for all the domains you require. If you have shared hosting, choose a provider that offers the option to have a certificate provided by Let's Encrypt.

https://letsencrypt.org/

When you use https, make sure that you allow only secure encryption methods. Some of the older https versions are no longer secure because they use outdated encryption.

There are free resources on the web for testing the security of your web server, such as:

https://www.ssllabs.com/ssltest/

CSP

Implement **Content Security Policy (CSP)**. This is a Worldwide Web Consortium (W3C) security layer that you can enable on your web server, in order to block external resources, such as scripts, style sheets, and images, from being loaded on your website. CSP helps to detect and mitigate certain types of attacks, including Cross Site Scripting (XSS) and data injection attacks. These attacks are used for everything from data access to site defacement to distribution of malware. You can customise the rules of your policy, white-listing some URLs.

If you do not have access to the web server configuration, you can still enable a CSP through an **.htaccess** file or directly in the header of your website.

You can test your configuration of CSP through a free service provided by Mozilla:

https://observatory.mozilla.org/

JAVASCRIPT

There is a view that JavaScript ought to be avoided because it may exclude some users or may result in less accessibility for some users. However, JavaScript in and of itself does not mean content necessarily becomes inaccessible. Care is, however, required.

https://webaim.org/techniques/javascript/

Implementing JavaScript tends to imply needing to build a non-JavaScript version. The best suggestion is to design a website without a JavaScript requirement, and build JavaScript-based features on top of it. This can make the browsing experience much more accessible for every user. Include a **<noscript>** tag to define alternate content that offers a version for users who have disabled JavaScript in their browser or who have a browser that still does not support JavaScript.

SECURING SENSITIVE DATA

If the website stores personal data of its users, for example by allowing them to have their personal accounts with login:

- Enforce the usage of strong passwords;
- Never store passwords and other sensitive information in plaintext. Always use hashing and encryption of sensitive data at rest;
- Support two-factor authentication (2FA);
- If the website administrative capacities are limited, you can offload the above tasks to an identity provider and enable single sign on. Please remember that while choosing an identity provider, this may have privacy costs.

If personalised accounts are not needed, in some cases, commenting on blogposts could be a desired feature. Integrate with third-party open source plugins such as **Discourse** or the **Coral project**:

https://www.discourse.org/

https://coralproject.net/

For additional details, refer to https://darekkay.com/blog/static-site-comments/

PROTECTION AGAINST DDOS ATTACKS

Distributed denial of service attacks (DDoS) are attacks typically accomplished by flooding the targeted machine with requests from many different sources, in an attempt to overload systems. If you are building an application for an NGO, an activist initiative or a civil society group, consider protecting it against DDoS. Some DDoS mitigation techniques use a third party that operates between a website's visitor and the website's hosting provider, acting as a reverse proxy for websites. However, some DDoS mitigation techniques will not require the services of a third party service, but can handle the attack on a network level.

It may be worthwhile looking at **Deflect**, which is a free Open Source solution. It is developed by a digital security non-profit organisation.

https://docs.deflect.ca/en/latest/about_deflect.html

However all DDoS mitigation services mean handing over control of your traffic to a third-party, whether or not this is done very quickly or the service is in-line all the time. The risks, reputational and otherwise, therefore need to be carefully evaluated.

STATIC WEBSITES ARE BACK

Do you really need a dynamic website? If you do not require a database, chances are that current web standards such as HTML5 allow you to create a state-of-the-art website using static resources (HTML, CSS and perhaps JavaScript and font files) only.

This does not mean that you have to hand-code navigation bars etc. Modern static site generators such as Jekyll (made popular by GitHub Pages), Hugo or Pelican let you create your websites in Markdown or basic HTML and convert them into full, interlinked HTML for you. If you are using GitLab Pages, conversion and deployment can be automated easily.

Static websites achieve unrivaled uptimes because they do not need to connect to a database. All they need is a plain web server to serve the static files. Consequently they are less ridden by security gaps, and they are lightning fast.

4. Alternatives to costly "free" third-party services

Using tools from Google, Facebook, Amazon and other data giants is often seen as incompatible with protecting the privacy of visitors to online services. Ethical developers who care about their users' privacy tend to try not to use, directly or indirectly, their services or those of their subsidiary companies. There are dozens of ethical, easy-to-use and privacy-conscious alternatives out there and many of them can be found in **Prism Break**.

https://prism-break.org/

Unfortunately, we know that third-party services are sometimes very useful, so we have assembled below a non-comprehensive list of data minimisation methods and alternatives for common services.

ANALYTICS

It's important to note that analytics are generally flawed and are blocked by many tracker blockers, so using them does not provide accurate information. But if you must... instead of using Google Analytics, or any other third-party analytics service, you can track the traffic on your website with **Matomo** (formerly **Piwik**). It is a privacy compliant analytics platform, very easy to install and can be configured to automatically anonymise data so you process a minimal amount of personal data. In this way, you should be in compliance with the GDPR. If you decide to process personal data, Matomo provides you with various features to more easily comply with the GDPR obligations.

https://matomo.org/

VIDEOS

YouTube belongs to Alphabet Inc., the multinational conglomerate parent of Google and other former Google subsidiaries. When embedding a video from YouTube, there is the option to enable privacy-enhanced mode. This means that YouTube will not place tracking cookies on the devices of the visitors of your website, unless they play the video. The 'Show suggested videos when the video finishes' should also be disabled. If you prefer to use a plugin for your content management system, look for one that takes in account these options.

An even better solution is to consider using, or even running your own, **Peertube** instance. For an explanation of **Peertube**, see:

https://framatube.org/videos/watch/9db9f3f1-9b54-44ed-9e91-461d262d2205

This has the added benefit of being less vulnerable to malicious take-down requests than in the case of YouTube. Furthermore, there will be no advertisement material added to your content. Keep in mind that if you rely on the video hosting website exploiting personal data to make money for you and for the video hosting website (such as YouTube), **Peertube** is not a viable option. Of course if you do rely on this approach, be aware that YouTube can cancel this at any moment, if Google, its parent company, suddenly decides you are not advertiser-friendly or arbitrarily changes the monetisation rules from one day to the next.

Also **Vimeo** allows Pro users to use their video files without the **Vimeo** player (which includes Google Analytics) so this is a useful stopgap if you are not able to use **Peertube**.

MAPS

If you use the Google Maps API to embed a map in your website, you are forcing users to accept the Google privacy policy. As an alternative you can use **OpenStreetMap**.

https://www.openstreetmap.org/

FONTS AND ICONS

Instead of using Google Fonts, which force users to accept Google privacy policy, you can use:

- **Fork Awesome** a very large library of open-source icons.

  https://forkawesome.github.io/Fork-Awesome/

- **Fontello** which allows you to create your own selection of icons, if you just want to load some of them.

  http://fontello.com/
  https://github.com/fontello/fontello/

- **Fontspring**, a font resource where they have a "trust, not trackers" policy, where you self-host your fonts.

  https://www.fontspring.com/fair-fonts

- **FontSquirrel** provides fonts that are 100% free for commercial use, and provides fonts for self-hosting.

http://www.fontsquirrel.com/

Lastly, you can resort to Google fonts downloaders, that take a Google Fonts URL, and produce an entirely local set of CSS and font files to use them. That still allows developers to use all of the fonts on Google Fonts with little-to-no extra effort, but without actually loading them from Google. You can download the fontfile you need and host it on your own webserver.

SOCIAL WIDGETS

The default embed code for social buttons provided by their platforms, such as the like or share buttons, send information back to these social networks even if the user does not click on them. There are many ways to avoid this, while still providing the same functionality. One of them is by using **Social Share Privacy**. A JavaScript and non-JavaScript option are available.

https://panzi.github.io/SocialSharePrivacy/

Another alternative out there to avoid both tracking and JavaScript is the social buttons generator **Sharingbuttons.io**. Due to the fact that they do not use JavaScript, they load very quickly and do not block your website from rendering.

https://sharingbuttons.io/

Another solution would of course be to build the buttons yourself.

CAPTCHAS

You may want to use a Captcha to prevent spam from occurring in any user generated content (i.e. comments). There are numerous reasons not to choose this option.

Captchas can pose serious barriers to users with disabilities. When attempting to distinguish actual users from 'bots', other skills are often demanded, such as hearing or seeing – and then responding to – certain challenges. Almost by definition, this will make the service difficult or impossible for visitors with sight or hearing problems. [See *"Making website accessible for all"* below.]

Currently, many captchas have privacy issues. Google Captchas are not only

annoying, but appear to be collecting personally identifiable additional data about their users.

https://www.businessinsider.com.au/google-no-captcha-adtruth-privacy-research-2015-2

There are better options to protect your website with simple captcha methods that do not require to load external JavaScript code. Some extensions have different options for settings. Take care of user experience and privacy when configuring them.

Some example of alternative Captchas are:

- Drupal

  https://www.drupal.org/project/captcha
- Wordpress For Contact form 7 with honeypot :

  https://wordpress.org/plugins/contact-form-7-honeypot/
- Wordpress Secure Image Captcha:

  https://wordpress.org/plugins/securimage-wp/
- Securimage:

  https://www.phpcaptcha.org/
- IndyCaptcha:

  https://github.com/dyne/indycaptcha

AND MORE

Event organisation tools

A lot of activism hinges on physical meetups and it is clearly important for activists not to let third parties look into attendance of such meetings.

- **Odoo** (formerly OpenERP) is a management software that includes several applications such as CRM, website/e-commerce, billing, accounting, manufacturing and event management among many others. The Community version is open source.

  https://www.odoo.com
  https://www.odoo.com/page/events
- **Attendize** is an open source ticket selling and event management platform.

  https://www.attendize.com/

Making websites accessible for all

Accessible website design defines whether disabled users have equal access to services. Accessibility is essential for developers and organisations that want to create high quality websites and web tools, and not exclude people from using their products and services. Organisations like the W3C have developed useful standards in that regard.

https://www.w3.org/WAI/

https://www.w3.org/standards/webdesign/accessibility

https://www.washington.edu/accesscomputing/sites/default/files/30-Web-Accessibility-Tips.pdf

Protect the web from untangling

External information that is available outside your control is often referenced through links. While these links interconnect the scattered information from different websites to the "web", the links that are outside your control eventually might vanish from the internet. Reasons for the links to fail could range from censorship to DDoS attacks or simply due to lack of maintainance of a website domain. However, ensuring that the users of your website can reliably access any linked content is your responsibility too.

One can use centralised archiving services (e.g. the **Internet Archive** or **perma.cc**) or host your own archive using the **Amber** tool. These services and open source tools create a snapshot of every page linked on website and preserve them.

http://amberlink.org/

https://perma.cc/

Ethical advertising

You may wish to consider avoiding the use of ad tech that track users and sell their data. There are alternatives to dodge this harmful business model.

https://docs.readthedocs.io/en/latest/advertising/ethical-advertising.html

Website search

If you use a search tool in your website, consider private search engines that do not track their users, such as **Startpage**.

https://startpage.com

https://choosetoencrypt.com/search-engines/private-search-engines-a-complete-guide/

5. Glossary

CMS (CONTENT MANAGEMENT SYSTEM): this kind of system supports the creation and management of digital content.

CSP (CONTENT SECURITY POLICY): computer security standard introduced to prevent cross-site scripting (XSS), clickjacking and other code injection attacks resulting from execution of malicious content in the trusted web page context.

CRONJOB: is a task launched by the software utility Cron which is a time-based job scheduler in Unix-like computer operating systems. People who set up and maintain software environments use Cron to schedule jobs (commands or shell scripts) to run periodically at fixed times, dates, or intervals. It typically automates system maintenance or administration.

DDOS (DISTRIBUTED DENIAL OF SERVICE ATTACK): an attack in which the perpetrator seeks to make network service unavailable to its intended users by temporarily or indefinitely disrupting one or more hosts. This is typically done by flooding the targeted environment with requests or other specific forms of traffic from many different sources in an attempt to overload the system and prevent some or all legitimate communication.

END-TO-END ENCRYPTION: system of communication where only the communicating parties can access the content. In principle, it prevents potential eavesdroppers – including telecom providers, internet access providers, transit providers and even the provider of the communication service – from being able to decrypt the content of the communication. End-to-end may imply user-to-service or user-to-user.

FLOSS (FREE LIBRE OPEN SOFTWARE): software freely licensed to be used, copied, studied, and changed in any way, and the source code openly shared so that people are encouraged to improve the design of the software. This is in contrast to proprietary software, where the software is under restrictive licensing and the source code is usually hidden from the users. There are also restrictions related to open-source software. These are necessary to preserve its fundamental nature as being free. Open source software also depends on an active and committed community to develop and maintain it.

GDPR (GENERAL DATA PROTECTION REGULATION OF THE EUROPEAN UNION): an EU Regulation on data protection that covers all individuals in the EU and that aims primarily to give control to individuals over their personal data and to simplify the regulatory environment. It also addresses the export of personal data to jurisdictions outside the EU.

INTEROPERABILITY: characteristic of a product or system, whose interfaces are designed to work with other products or systems, in either implementation or access, without significant restrictions.

NGO: Non Governmental Organisation.

PERSONAL DATA: any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person. (As defined by the GDPR).

PEER-TO-PEER (P2P): a distributed application architecture that divides tasks or workloads or indeed data between peers. Peers are equally privileged participants in the application. They are said to form a peer-to-peer network of nodes.

PRIVACY BY DEFAULT: the principle according to which an organisation (the data controller) ensures that only data strictly necessary for each specific purpose of the processing are processed by default, i.e. the most privacy protective setting is the default, even if the user has the option to choose less protective settings.

STOPGAP: temporary measure or short-term fix used until something better can be obtained.

THIRD-PARTY CONTENT: is at its simplest where the property rights related to some content are held by another party or dynamically imported from their equipment.

TOR: free software initially developed by the US military for enabling anonymous communication. Tor directs Internet traffic through a free, worldwide, volunteer overlay network consisting of more than seven thousand relays to conceal a user's location and usage from anyone conducting network surveillance or traffic analysis. The Tor Browser automatically starts Tor background processes and routes traffic through the Tor network.

European Digital Rights

Founded in 2002, EDRi is the biggest European network defending rights and freedoms online.

Currently 42 non-governmental organisations are members of EDRi and 30 observers closely contribute to our work.

Our mission is to promote, protect and uphold human rights and the rule of law in the digital environment, including the right to privacy, data protection, and freedom of expression and information.

Our vision is for a Europe where state authorities and private companies respect everyone's fundamental rights and freedoms in the online environment. Our overall aim is to build the structures where civil society and individuals are empowered to embrace technological progress in control of their rights.

MASS SURVEILLANCE.
RANDOM CENSORSHIP.
CONTENT RESTRICTIONS.

Companies and governments increasingly restrict our freedoms.

Donate NOW:
https://edri.org/donate

PRIVACY!
FREE SPEECH!
ACCESS TO KNOWLEDGE AND CULTURE!

We defend rights and freedoms online.

EDRi

EUROPEAN DIGITAL RIGHTS

https://edri.org

@edri

brussels@edri.org