



The National Archives of the Netherlands and archiving government websites

Antal Posthumus

Archiwum Narodowe Holandii / National Archives of the Netherlands
antal.posthumus@nationaalarchief.nl, ORCID: 0000-0002-4559-7395

ABSTRACT

The National Archives of the Netherlands, as a permanent government agency and official archive for the central government (ministries and their agencies), has the legal duty, laid down in the Archiefwet, to secure the future of government records. Within this context, our role does not involve actively forming a collection of archived websites through selecting and harvesting these ourselves. This is a key difference between us and other national archives, national libraries and other (inter-)national heritage institutions. Guidelines and a central platform for archiving government websites. Such a mandate requires an environment in which the processes, in relation to one another, can take place in a controlled manner. A significant part of making it happen was the effort we've put (and continue to do so) into advising the producers of records – ministries and their agencies – as to how they should create, and eventually transfer, archived public websites that are a specific form of government records. One example of the type of support we offer was a very well received set of guidelines on archiving websites that we issued in 2018. Those guidelines were also used as part of the requirements in a public European tender (2021). The objective of the tender: implementation of a central harvesting platform to harvest approximately 1500 public websites of the Central Government. This article will also present our experiences and insights into integrating the processes of ingestion, storage, management and preservation of and providing access to archived public websites of the Dutch Central Government into the existing infrastructure and workflows of our trusted digital repository (e-depot in short).

KEYWORDS

guidelines on archiving government websites, WARC validation factsheet, The General Data Protection Regulation, publication rights, archiving by design, web archiving guidance, pywb (Python Wayback) implementation, ingestion, preservation and access of WARC-files, Central Platform for archiving websites for the Dutch Central Government

Archiwum Narodowe Holandii i archiwizacja rządowych stron internetowych

STRESZCZENIE

Archiwum Narodowe Holandii, jako stała agencja rządowa i archiwum rządu centralnego (ministerstw i ich agencji), ma prawny obowiązek zabezpieczania akt rządowych. Rola archiwum nie polega na aktywnym tworzeniu kolekcji zarchiwizowanych stron internetowych poprzez ich samodzielne wybieranie i gromadzenie. Różni się tym od innych narodowych archiwów, bibliotek i innych (między-)narodowych instytucji dziedzictwa zajmujących się archiwizacją stron internetowych. Archiwum w swoich działaniach skupia się na doradzaniu twórcom dokumentacji – ministerstwom i ich

SŁOWA KLUCZOWE

wytyczne archiwizacji rządowych stron internetowych, karta walidacji WARC, rozporządzenie o ochronie danych

agencjom – w kwestiach tworzenia i przekazywania rejestrów rządowych w formie zarchiwizowanych, publicznych stron internetowych. Jednym z przykładów tego typu wsparcia było wydanie w 2018 r. dobrze przyjętych wytycznych dotyczących archiwizacji stron internetowych. Zostały one także wykorzystane jako część wymagań europejskiego przetargu publicznego w 2021 r., dotyczącego utworzenia centralnej platformy mającej pozyskać około 1500 publicznych stron internetowych rządu centralnego Holandii. W artykule zostaną również przedstawione nasze doświadczenia i spostrzeżenia dotyczące integracji procesów pozyskiwania, przechowywania, zarządzania, zabezpieczania i zapewnienia dostępu do zarchiwizowanych publicznych stron internetowych holenderskiego rządu centralnego z istniejącą infrastrukturą oraz zarządzaniem danymi w repozytorium cyfrowym Archiwum Narodowego Holandii (w skrócie w-depot).

osobowych, prawo do publikacji, archiwizacja przez projektowanie, wytyczne archiwizacji stron internetowych, implementacja pywb (Python Wayback), ingest, przechowywanie i udostępnianie plików WARC, Centralna Platforma archiwizująca strony internetowe rządu Holandii

Het Nationaal Archief (eng. The National Archives of the Netherlands), as a permanent government agency under the Ministry of Education, Science and Culture, functions as the official archive for the Central Government (ministries and their agencies) with the legal duty, as laid down in the Public Records Act, to secure the future of the government record.

This article describes how we secure the future of one very specific form of government records, namely archived public government websites.

One part of performing this public task was to make sure we are able to receive, manage and present archived public websites that ministries and their agencies will eventually transfer to our trusted digital repository (e-depot in short).

The National Archives in the Netherlands also functions as a knowledge centre/consultancy on archival management and information management for the Dutch government, both on central and local level. A set of guidelines on archiving government websites that we drafted is an example of a product we delivered in performing this role.

This article will discuss in some detail the background of how we implemented processes related to the ingestion, storage, management and preservation of and providing access to archived public websites of the Dutch Central Government within the infrastructure and workflows of our trusted digital repository.

At the same time, it will explain how our guidance on archiving government websites has played a central role in this implementation. The article also illustrates the role that several aspects of those guidelines played when applied as requirements in a public European tender (2021). The objective of the tender

was to implement a central harvesting platform to harvest approximately 1500 public websites of the Dutch Central Government.

Why should we archive websites of government organizations?

Government websites should be treated as any other form of government information, with the same care, as such information is important for accountability, law & evidence, research and history. It should also be handled in accordance with the same principles that apply to archiving all digital government information: authenticity, completeness and accessibility.

Take for instance the website of a capital like Amsterdam. Nowadays, more often than not, this is the place where all the relevant and official information about public services (e-services) of the local government is offered to citizens. The website is the primary channel for this information, as well as a channel for citizens to apply for those services. Same is the case for most Dutch local governments, as well as most other countries.

Citizens can also obtain information about their rights and responsibilities from governmental websites. For example, on the municipal website of Amsterdam, on pages concerning child care, all the terms and conditions, including money amounts, can be found¹. This information might be updated – changed – daily. Citizens base their decisions and actions on this information. Therefore it is important that when the information is changed, its previous versions remain accessible so that if needed, they can be used as proof by both parties.

Governmental websites also contain information that for different reasons may remain relevant for a much longer period. One particularly striking and emotionally charged example of this can be found on the Netherlands central government website presenting the Dutch government's official viewpoint on the MH17 crisis². It is doubtlessly an example of historic significance: MH17 was an airplane with mostly Dutch citizens onboard that was shot down with

¹ Gemeente Amsterdam. Vergoeding Aanvragen Voor de Eigen Bijdrage Kinderopvang, <https://www.amsterdam.nl/veelgevraagd/?productid=%7BE63FD3B0-5875-4A03-9082-7DA97B0AFC2C%7D>, accessed 3 October 2022.

² Ministerie van Algemene Zaken. MH17: Nederland En Australië Stellen Rusland Aansprakelijk – Nieuwsbericht – Rijksoverheid.nl., <https://www.rijksoverheid.nl/onderwerpen/neerhalen-vlucht-mh17/nieuws/2018/05/25/mh17-nederland-en-australie-stellen-rusland-aansprakelijk>, accessed 3 October 2022.

a rocket above Ukraine in 2014. All the people on board were killed. The event led to a political crisis between the Netherlands and Russia.

So the information on this webpage is relevant both politically and historically, as it is the primary publication of the government's position. Integrity and authenticity of this information is important. A researcher, both now and fifty years from now, wants to be sure that this information was in fact published by the Dutch government and has not been altered since, either illegally or by accident.

After the retention period has ended this information should be and will be transferred to the trusted digital repository (e-depot in short).

In following paragraphs, we offer a bit of history and context about actions we had to undertake as a national archive in order to be able to ingest, store, manage, preserve and provide access to this kind of archived government websites. Some insight will also be provided into our different roles as an archival institution: supplier, consultant and facilitator, knowledge institute and of course official archive for government records.

Legal context of the National Archives of the Netherlands

Within the same legal context of the Public Records Act, in which the National Archives of the Netherlands is designated as the official archive for the central government (ministries and their agencies), these governmental agencies are responsible for keeping the information they produce in so called "good, accessible state". That applies from the moment the information is created, until the moment that same information can either be destroyed or transferred to an archival repository for permanent preservation. The transfer takes place after 20 years. That means that governmental agencies have to keep at least some of those records in "good accessible state" for quite a long time. What "good and accessible" means is not specified in the Public Records Act, therefore it can be interpreted in many ways. To no surprise, this is also true of websites which are to be considered government records.

One example of our role as a knowledge centre for government organizations is to provide all governmental agencies with advice and guidelines on how to execute their responsibility for "good and accessible state". However, our advice/

interpretation is not mandatory for them to use. They can make other choices about how to specify and implement the obligations of the Public Records Act.

Because of the legal context in which we operate, we're also the supplier of a trusted long term digital repository for permanent retention of government information. Accordingly, we don't actively form a collection of archived websites through selecting and harvesting these ourselves. This is a key difference between us and other national archives, national libraries and other (inter-) national heritage institutions. Most of these organizations archive websites based on an organization, domain or theme specific selection policy, and some have outsourced the harvesting process to a third-party company or execute this entirely themselves.

Guidelines on archiving government websites

Since our functions include that of a knowledge institute and consultant for government agencies, we've put a fair amount of effort into advising the producers of information (records generators; note to editors: "producer" is one of the key terms in the OAIS-framework which is used globally by national archives and heritage institutions to model their infrastructure and work processes) – ministries and their agencies – on how they should create and eventually transfer archived public websites, which are a specific form of government records. This a clear example of the new role for archival institutions in the digital age – advising, by design, at the start of the process, rather than at the end (by then it is too late).

One example of this type of support that we've offered was issuing a very well received set of guidelines on archiving websites (2018).

Goals of the guidelines

Our guidelines³ fit the abovementioned new role perfectly. Its aim is to maximize standardization and uniformity of the output of the websites archiving process, carried out by governmental agencies. By influencing the output, we can

³ Nationaal Archief. Richtlijn Archiveren Overheidswebsites, <https://www.nationaalarchief.nl/archiveren/kennisbank/Richtlijn-Archiveren-Overheidswebsites>, accessed 3 October 2022.

influence the quality of the archived material that will be transferred to us for long term retention.

We had several goals when developing the guidelines; one of them was already described. Websites are a primary source of communication for governmental agencies and therefore they should be treated as any other form of governmental information, with the same care, due to importance of the information.

The second goal logically following from this first, is to ensure governmental websites are correctly archived by the corresponding governmental agencies; such organizations more often than not lack the knowledge needed for website archiving.

Political assignment

Besides these goals, we also have a political assignment as a permanent government agency under the Ministry of Education, Science and Culture.

In 2016 the Dutch Heritage Inspectorate sounded the alarm in a report which concluded that barely 5% of all Dutch government websites were properly archived. This was partly because Dutch government organizations did not have the knowledge and experience needed to archive their own websites in accordance with the Archives Act. The National Archives were then commissioned by the Minister of Education, Culture and Science to draw up guidelines and the Chief Information Officer of the government was tasked with implementing them in all national government organizations.

Support base and adoption

Those countrywide guidelines are not, in themselves, mandatory for government organizations, since the organizations have the freedom to choose how they fulfil the 'good, orderly and accessible state' concept as set forth in Dutch archival law. It was therefore very important to closely involve the future users of the guidelines in their development. We shared issues and dilemmas. We also organized a public review, where everyone could comment

on a first draft of the guidelines⁴. We received no less than 600 comments from 30 different organisations, which enabled us to improve the guidelines and immediately bring them to the attention of potential future users.

Scope of the guidelines

The scope of the guidelines covers archiving an entire website and its daily changes. We stress the importance of documenting missing (dynamic) elements⁵ from archived websites in order to be able to account for them to our end users. We now opt for basic means of delivering this metadata when transferring archived websites: an .ods file, in which our records producers register, what content wasn't harvested properly. This as a means for us, as an archival institution, to inform our users why certain content isn't presented to them. This adds to the context information regarding this specific type of government records.

Since harvesting websites can be a very complex task, we advise to leave it to the experts, in other words to outsource the harvesting process. Nevertheless, as an archival institution one should certainly keep an eye on it and have insight into metadata, quality control, integrity and authenticity as early in the harvesting process as possible.

Another important and mandatory requirement in the guidelines is the use of WARC-files (web archive files), as it is the international ISO-standard (28500:2017)⁶, to structure, manage and store billions of records collected from

⁴ Kennisnetwerk Informatie en Archief. Voortgangsbericht Richtlijn Archiveren Overheidswebsites, <https://kia.pleio.nl/groups/view/1997dd74-cb58-420c-9056-85d1194729b9/kennisplatform-webarchivering/discussion/view/f63dd32d-8a3d-436f-9329-6aa3b8d2b8d8/voortgangsbericht-richtlijn-archiveren-overheidswebsites>, accessed 3 October 2022.

⁵ Archive-It Help Center. Known Web Archiving Challenges, <https://support.archive-it.org/hc/en-us/articles/209637043-Known-Web-Archiving-Challenges>, accessed 3 October 2022; Digital Preservation Coalition. New Report: Archiving the Dynamic Web Benefits from Server-Side Archiving, <https://www.dpconline.org/news/server-side-archiving>, accessed 3 October 2022.

⁶ Bibliothèque Nationale de France. The WARC File Format (ISO 28500) – Information, Maintenance, Drafts, <http://bibnum.bnf.fr/WARC/>, accessed 3 October 2022; iipc.github.io. The WARC Format 1.1., <https://iipc.github.io/warc-specifications/specifications/warc-format/warc-1.1/>, accessed 3 October 2022.

the web and elsewhere. It is used to build applications for harvesting, managing, accessing, and exchanging content.

Harvesting 1500 public websites of the Central Government

As already mentioned, our guidelines were also used as part of the requirements in a public European tender (2021). The objective of the tender was implementation of a central harvesting platform to harvest approximately 1500 public websites of the Central Government.

Implementation of the guidelines, the European tender and implementation of this central harvesting platform was carried out under the responsibility of the Chief Information Officer of the Central Government. Other important players in archiving the websites of Dutch central government organizations include the Public Information and Communications Service of the Office of the Prime Minister and the National Program for Durable Digital Information Management.

2019 – start of our Project implementation

Such is the context in which in 2019 we started implementing the project that would enable us to receive the archived materials, which in the long run would be transferred from the abovementioned central harvesting platform, as well as from other sources. Our aim was to formulate and implement requirements for various aspects of the OAIS-model with regard to archived websites. In other words, we had to integrate processes related to ingestion, storage, management and preservation of and providing access to archived public websites of the Dutch Central Government into the existing infrastructure and workflows of our trusted digital repository (e-depot in short).

We chose an ongoing project approach, starting in 2019, combining a Prince2 and Agile method. This was carried out by a multidisciplinary project team consisting of an information analyst, a record keeping advisor, an information architect, a preservation officer, a collection manager plus the product owner of our organization's website.

Desk research and surveys

We conducted interviews with a peer group consisting of several users in order to get a clear idea of the expectations our end users have with regard to using the archived websites. We will deliver the archived materials to our more regular users through electronic search aids, provided on our website. We are still figuring out how to offer this material in a fashion that researchers would find most suitable for approaching and using the archived online resources.

Our project started with a lot of hypotheses and rudimentary requirements. Some aspects required research, others had to be learned through experience of just starting and moving forward. For example, which search options will we be able to implement? Search only through metadata? Within separate websites? Full text search? Will our national metadata scheme, our e-depot's data model, and the EAD as used in our collection management system, suffice to provide adequate administrative, descriptive and technical metadata for archived websites? Will the off the shelf viewer, Open Wayback, as it is installed in our E-depot solution, do the trick? Do we need to install a separate viewer in our infrastructure?

Metadata

In our guidance there's an example of a hierarchical structure with descriptions of which metadata we want to receive for each level. This matches almost all the mandatory metadata fields in the current data model of our long term repository. It is also sufficient to manage the data and to create and present a digital finding aid on our website. Therefore, we didn't adopt for example the latest OCLC's recommendations about descriptive metadata for Web Archiving⁷.

⁷ OCLC. Web Archiving Metadata Working Group, <https://www.oclc.org/research/areas/research-collections/wam.html>, accessed 3 October 2022.

Pywb as a viewer for rendering archived websites

We found that the off the shelf viewer in our e-depot, the Open Wayback, wasn't capable of rendering increments of one and the same website. Moreover, our e-depot software doesn't support external viewers. Our software supplier indicated the off the shelf version, used in our e-depot, isn't a complete version of the OpenWayback software and therefore isn't able to render harvested websites in increments, or provide a calendar function, a requirement we considered one of the key elements of our minimum viable product (note to editors: an mvp, minimum viable product is a one of the key terms originating from the domain of agile software development which we used in our project.) See for instance: https://en.wikipedia.org/wiki/Minimum_viable_product. This goes for the current and future software versions; apparently a viewer for rendering archived websites isn't a top priority for most of their customers.

There are justifiable restrictions on modifying the e-depot software code by ourselves in this aspect. Therefore, we would not be able to render harvested websites incrementally, either with our current deployment or with future ones.

Those experiences were one of the reasons we've decided to implement Pywb, which is a python based rendition of the globally used OpenWayback software. However, through desk research we found recent developments in the IIPC community (International Internet Preservation Consortium) concerning the transition from OpenWayback to pywb (Python Wayback). The IIPC community considers the latter the preferred replay tool for the future⁸. Although pywb offers a good solution, it also has certain issues that need to be considered; not least of them is the matter of double storage of WARC-files: in our long term repository and in pywb itself.

The problem is that the e-depot software doesn't support a linkage with this external viewer which forces us to copy all WARC-files into the viewer. This means a deviation from our current workflow; it also means we need twice as much disk space, so to speak.

Another problem is that our collection managers now have to exit the e-depot when performing a visual check during either a test or actual process of ingesting

⁸ International Internet Preservation Consortium. Support for Transitioning to Pywb, <https://netpreserve.org/projects/pywb>, accessed 3 October 2022; pywb.readthedocs.io. Webrecorder Pywb Documentation!, <https://pywb.readthedocs.io/en/latest/>, accessed 3 October 2022.

incrementally harvested websites. They have to open WARC-files through the external viewer instead.

Search functionality for archived websites

At this time, we only offer the option to search through the descriptive and (some) technical metadata which accompanies the WARC files we receive from the central harvesting platform.

Full text search will be developed along with development of generic full text search for all of our born digital records (note to editors: the usage of the term born digital is very common and widespread) See for instance: <https://dictionary.archivists.org/entry/born-digital.html>. The WACZ file format might be of help in this respect in the future. However, that is a recent development and we must carry out a more detailed analysis in order to fully understand its usage and possible advantages. WACZ stands for Web Archive Collection Zipped; it has several advantages compared to plain WARC-files. With regard to search functionality, one advantage of WACZ is that “the full-text data extracted from web pages is also included, ready to be ingested into search engines like Solr or loaded on-the-fly along with the replay”⁹.

WARC.gz characterisation and WARC validation

WARC.gz is a standard method of lossless compression of WARC-files for storage purposes. We found out that WARC-gz files were not properly identified by either the version of the file identification tool DROID that was deployed within our e-depot or the standalone version outside our e-depot. Identifying specific applications of Gzip containers is a challenge. A .gz-file will be identified, but a .warc.gz won't; same goes for other specific uses of the .gz container. The preservation community is aware of this, but as for now a file-specific signature hasn't been developed yet.

⁹ Webrecorder. Announcing WACZ Format 1.0., <https://webrecorder.net/2021/01/18/wacz-format-1-0.html>, accessed 3 October 2022.

During our ingest workflow, a .warc.gz file will be checked for a .warc file (based on the WARC PUID). If it contains a .warc file, it will be registered in the share of pywb (Python Wayback).

We also found, with regard to the international standard file format WARC, that our e-depot does not have a built-in tool to validate WARC and WARC.gz files. Thus, we realized we needed to collect more information on validity of WARC files during the pre-ingestion process, which led us to run and review several WARC validation tools. A factsheet was prepared, listing the output of these tools, the best way to apply them, and so on¹⁰. Most of those validation tools that check conformance with WARC standard ISO 28500, aren't very sophisticated. Some of them only check block and/or payload digests. Most of them support version 1.0 of the WARC standard (of 2009) and only few support version 1.1 (of 2017).

The most prominent finding of our research was that using a combination of tools will probably be the best strategy for now. That is, if it is a viable and useful strategy at all. We also found that most web archiving initiatives don't give much attention to this aspect of quality control.

GDPR, publication rights and archived websites

Recently, during the transfer of several archived websites of a financial government agency, we found resumés of board members with various personal data. We also came across portraits of board members, researchers, working group members and employees. Permission of these people was asked beforehand, but there wasn't any signed statement of consent. It also wasn't feasible to find all of the portrayed people and ask their consent.

In other words, we encountered some challenges related to GDPR and portrait rights.

Another legal issue we had to mitigate concerned copyrights to photos from the Dutch Press Agency (ANP) and other publishers like Shutterstock and the like. There is no information attached to the individual photos with regard to their respective owners; nor has this agency kept any form of a register of images that have been used on the website and conditions applicable to such usage.

¹⁰ Open Preservation Foundation. WARC Validation Tool Experiences, <https://openpreservation.org/blogs/warc-validation-tool-experiences/>, accessed 3 October 2022.

Together with ANP, we discussed these issues with the given government agency. The government agency and ANP signed an agreement stating that ANP would grant its permission to the agency so that the photos can be transferred to us as a national archive within the context of the archived websites. On top of that, in order to minimize the risk of improper use by third parties after the transfer, we will emphasize this in the PDF of the inventory describing this archive. We will also make this clear by attaching some information to each clickable inventory item that we share on our website.

The risk of improper use of this content or of displaying GDPR-sensitive data is further minimized by the fact that for the time being we don't offer full text search on our website. Moreover, the content of WARC files presented on our website is not being indexed by Google, which means that any personal data will not be systematically findable. Combined, these mitigating measures provided enough confidence among all parties for us to be able to present these archived websites on our website.

Another risk mitigating measure minimizing the risk of legal problems is one that can be used before an agency starts archiving their website; maybe even at the moment of creating the website or content. This could be regarded as a form of archiving by design. As part of the implementation process of archiving central government websites, a legal review tool was developed for agencies so they can check for themselves whether, and how much, GDPR-sensitive content is actually present on their websites. The tool also checks for content involving portrait rights, trademark law and database law¹¹.

What did we learn along the way?

Websites are a primary source of governmental information that need to be preserved for the purposes of accountability and future research.

¹¹ Informatiehuishouding.nl. Modeltoets AVG En Publicatierechten Voor Archivering Overheidswebsites Door ICT Recht, <https://www.informatiehuishouding.nl/Producten+%26+publicaties/instrumenten/2019/12/13/modeltoets-avg-en-publicatierechten-kopie>, accessed 3 October 2022; Ministerie van Onderwijs, Cultuur en Wetenschap. Websitearchivering – Rijksprogramma Voor Duurzaam Digitale Informatiehuishouding, <https://www.informatiehuishouding.nl/onderwerpen/websitearchivering>, accessed 3 October 2022.

In order to achieve that, they need to be archived according to clear archival requirements, so as to ensure their sustainable accessibility. Those requirements must be practical and usable, but also acceptable to users. Involving users, especially in the review, really helped in developing guidelines that would be used as intended.

One last important point is that the process didn't end with the publication of the guidelines. Their implementation is equally important. It should involve all the key players, including the National Archives, not only as the final archival repository but also as a consultant and facilitator.

This entire process has underlined for us how important it is to set requirements for information management long before transfer; in other words – archiving by design. Without that, by the time of transfer, a website will have been changed lots of times, or even dismantled; its information would be lost forever.

Therefore, archival institutions need to be there at the very beginning, and engage in a dialogue with governmental agencies, be their partner along the entire way, from start to finish, at every stage.

Bibliography

- Archive-It Help Center. Known Web Archiving Challenges, <https://support.archive-it.org/hc/en-us/articles/209637043-Known-Web-Archiving-Challenges>, accessed 3 October 2022.
- Bibliothèque Nationale de France. “The WARC File Format (ISO 28500) – Information, Maintenance, Drafts”, <http://bibnum.bnf.fr/WARC/>, accessed 3 October 2022.
- Dictionary of Archives Terminology. Born Digital, <https://dictionary.archivists.org/entry/born-digital.html>, accessed 3 October 2022.
- Digital Preservation Coalition. New Report: Archiving the Dynamic Web Benefits from Server-Side Archiving, <https://www.dpconline.org/news/server-side-archiving>, accessed 3 October 2022.
- Geemente Amsterdam. Vergoeding Aanvragen Voor de Eigen Bijdrage Kinderopvang, <https://www.amsterdam.nl/veelgevraagd/?productid=%7BE63FD3B0-5875-4A03-9082-7DA97B0AFC-2C%7D>, accessed 3 October 2022.
- ipcc.github.io. The WARC Format 1.1, <https://ipcc.github.io/warc-specifications/specifications/warc-format/warc-1.1/>, accessed 3 October 2022.
- Informatiehuishouding.nl. Modeltoets AVG En Publicatierchten Voor Archivering Overheidswebsites Door ICT Recht, <https://www.informatiehuishouding.nl/Producten+%26+publica->

- ties/instrumenten/2019/12/13/modeltoets-avg-en-publicatierechten-kopie, accessed 3 October 2022.
- International Internet Preservation Constortium. Support for Transitioning to Pywb, <https://net-preserve.org/projects/pywb>, accessed 3 October 2022.
- Kennisnetwerk Informatie en Archief. Voortgangsbericht Richtlijn Archiveren Overheidswebsites, <https://kia.pleio.nl/groups/view/1997dd74-cb58-420c-9056-85d1194729b9/kennisplatform-webarchivering/discussion/view/f63dd32d-8a3d-436f-9329-6aa3b8d2b8d8/voortgangsbericht-richtlijn-archiveren-overheidswebsites>, accessed 3 October 2022.
- Ministerie van Algemene Zaken. MH17: Nederland En Australië Stellen Rusland Aansprakelijk – Nieuwsbericht – Rijksoverheid.nl, <https://www.rijksoverheid.nl/onderwerpen/neerhalen-vlucht-mh17/nieuws/2018/05/25/mh17-nederland-en-australie-stellen-rusland-aansprakelijk>, accessed 3 October 2022.
- Ministerie van Onderwijs, Cultuur en Wetenschap. Websitearchivering – Rijksprogramma Voor Duurzaam Digitale Informatiehuishouding, <https://www.informatiehuishouding.nl/onderwerpen/websitearchivering>, accessed 3 October 2022.
- Nationaal Archief. Richtlijn Archiveren Overheidswebsites, <https://www.nationaalarchief.nl/archiveren/kennisbank/Richtlijn-Archiveren-Overheidswebsites>, accessed 3 October 2022.
- OCLC. Web Archiving Metadata Working Group, <https://www.oclc.org/research/areas/research-collections/wam.html>, accessed 3 October 2022.
- Open Preservation Foundation. WARC Validation Tool Experiences, <https://openpreservation.org/blogs/warc-validation-tool-experiences/>, accessed 3 October 2022.
- Pywb.readthedocs.io. Webrecorder Pywb Documentation!, <https://pywb.readthedocs.io/en/latest/>, accessed 3 October 2022.
- Webrecorder, Announcing WACZ Format 1.0., <https://webrecorder.net/2021/01/18/wacz-format-1-0.html>, accessed 3 October 2022.
- Wikipedia. Minimum Viable Product, https://en.wikipedia.org/wiki/Minimum_viable_product, accessed 3 October 2022.