



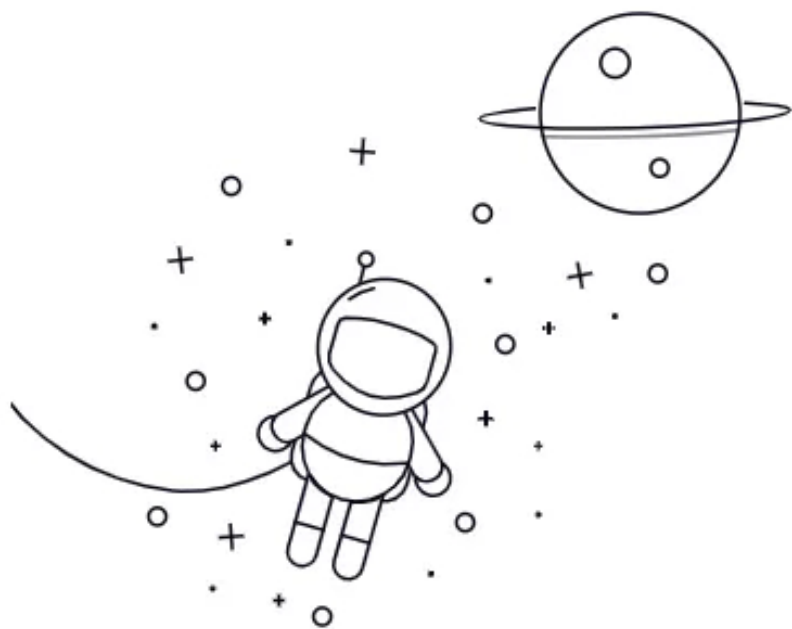
MREŽA ZNANJA

Ljubljana, 11. – 13. oktober 2022

Spletni arhivi: zdravilo za demenco svetovnega spleta?

Janko Klasinc

Narodna in univerzitetna knjižnica



404

UH OH! You're lost.

The page you are looking for does not exist. How you got here is a mystery. But you can click the button below to go back to the homepage.

HOME



„Arhiviranje spleta je postopek zbiranja delov svetovnega spleta in zagotavljanje njihove ohranitve za bodoče raziskovalce, zgodovinarje in splošno javnost.“

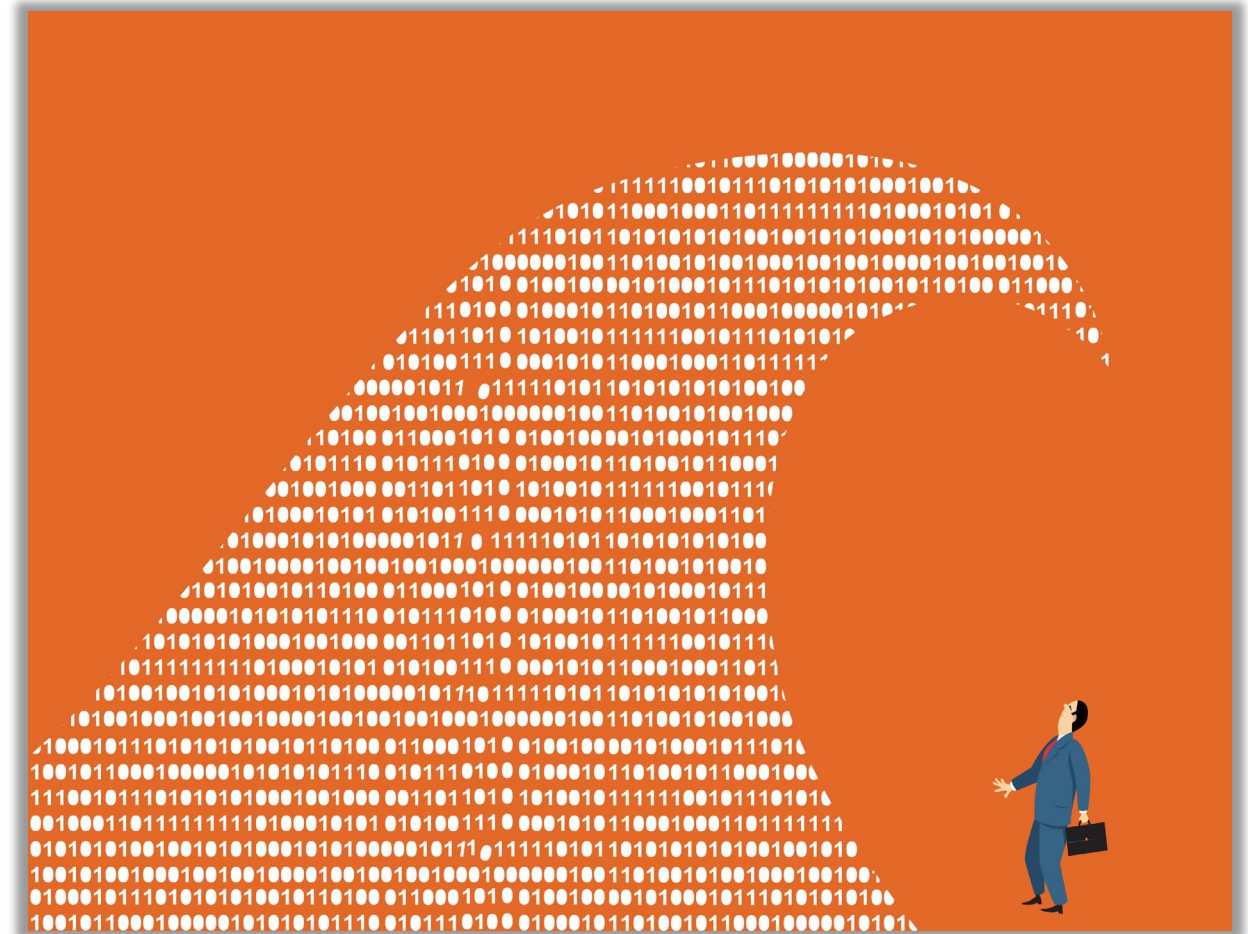
Vir: https://en.wikipedia.org/wiki/Web_archiving



Kako arhivirati svetovni splet?

„Zaradi velikega obsega spleta in količine informacij, ki jih vsebuje, spletni arhivisti za avtomatizirano shranjevanje običajno uporabljajo spletne robote (pajke).“

Vir: https://en.wikipedia.org/wiki/Web_archiving



Vir: <https://www.cfrichmond.org/Leadership-Impact/Blog/PostId/257/storytelling-with-data>



Wayback Machine

The screenshot shows the Wayback Machine homepage. At the top, there is a dark navigation bar with the Internet Archive logo and icons for various media types: WEB, BOOKS, VIDEO, AUDIO, SOFTWARE, and IMAGES. To the right of these icons are links for SIGN UP, LOG IN, and UPLOAD. Below this bar is a secondary navigation menu with links for ABOUT, BLOG, PROJECTS, HELP, DONATE (with a red heart icon), CONTACT, JOBS, VOLUNTEER, and PEOPLE.

The main content area features the Internet Archive logo and the text "Explore more than 725 billion web pages saved over time". Below this is a search bar with the text "http://" and a "BROWSE HISTORY" button. A "DONATE" button is also visible with the text "Find the Wayback Machine useful?".

A row of ten small thumbnail images shows various web pages that have been archived. Below this row are three sections:

- Tools**: A gear icon followed by the text "Tools". Below this are three links: "Wayback Machine Availability API", "WordPress Broken Link Checker", and "404 Handler for Webmasters".
- Subscription Service**: An "Archive-It" logo followed by the text "Subscription Service". Below this is a paragraph: "Archive-It enables you to capture, manage and search collections of digital content without any technical expertise or hosting facilities. Visit Archive-It to build and browse the collections."
- Save Page Now**: A "Save Page Now" logo followed by the text "Save Page Now". Below this is a search bar with the text "https://" and a "SAVE PAGE" button. A paragraph below reads: "Capture a web page as it appears now for use as a trusted citation in the future. Only available for sites that allow crawlers."

At the bottom of the page, there are links for "FAQ", "Contact Us", and "Terms of Service (Dec 31, 2014)".

70 PB podatkov

727 milijard spletnih strani



Wayback Machine

NEW COOL RANDOM YAHOO! EXTRA HEAD LINES YAHOO INFO ADD URL

CLICK HERE TO VISIT THE STARS YAHOO! LOS ANGELES Weekly Picks

Search Options

[Yellow Pages](#) - [People Search](#) - [City Maps](#) -- [News Headlines](#) - [Stock Quotes](#) - [Sports Scores](#)

- [Arts](#) -- [Humanities](#), [Photography](#), [Architecture](#), ...
- [Business and Economy \[Xtra!\]](#) -- [Directory](#), [Investments](#), [Classifieds](#), ...
- [Computers and Internet \[Xtra!\]](#) -- [Internet](#), [WWW](#), [Software](#), [Multimedia](#), ...
- [Education](#) -- [Universities](#), [K-12](#), [Courses](#), ...
- [Entertainment \[Xtra!\]](#) -- [TV](#), [Movies](#), [Music](#), [Magazines](#), ...
- [Government](#) -- [Politics \[Xtra!\]](#), [Agencies](#), [Law](#), [Military](#), ...
- [Health \[Xtra!\]](#) -- [Medicine](#), [Drugs](#), [Diseases](#), [Fitness](#), ...
- [News \[Xtra!\]](#) -- [World \[Xtra!\]](#), [Daily](#), [Current Events](#), ...
- [Recreation and Sports \[Xtra!\]](#) -- [Sports](#), [Games](#), [Travel](#), [Autos](#), [Outdoors](#), ...
- [Reference](#) -- [Libraries](#), [Dictionaries](#), [Phone Numbers](#), ...
- [Regional](#) -- [Countries](#), [Regions](#), [U.S. States](#), ...
- [Science](#) -- [CS](#), [Biology](#), [Astronomy](#), [Engineering](#), ...
- [Social Science](#) -- [Anthropology](#), [Sociology](#), [Economics](#), ...
- [Society and Culture](#) -- [People](#), [Environment](#), [Religion](#), ...

[Yahoo! New York](#) - [Yahoo! Shop](#) - [Yahooligans!](#)

[Yahoo! Japan](#) - [Yahoo! Internet Life](#) - [Yahoo! San Francisco](#)

YAHOO

17. 10. 1996

Primer NUK

Zakon o obveznem izvodu publikacij (Ur. list RS, št. [69/06](#) in [86/09](#))

Periodičen zajem slovenike na spletu:

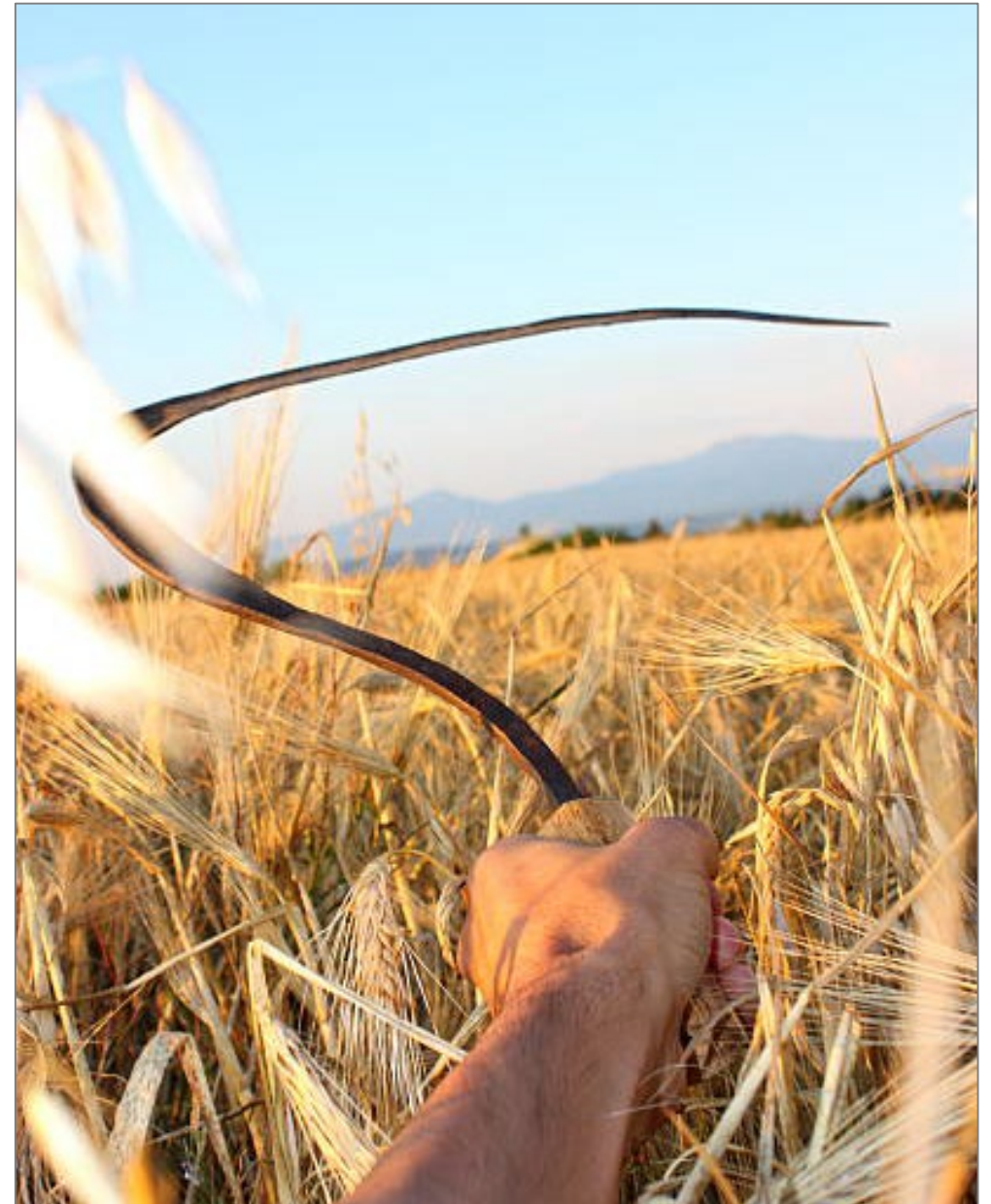
- vsebine slovenskih avtorjev
- vsebine v slovenskem jeziku
- vsebine o Sloveniji



Primer NUK

Selektivni pristop (od **2008**)

- okoli **1.800** spletnih domen
- frekvenca – od 1x tedensko do 1x letno
- globina – do 20 klikov
- 12 vsebinskih področij
(znanost, kultura, izobraževanje, vlada
in politika, gospodarstvo, prosti čas, itd.)





Primer NUK

Zajem celotne domene **.si**:

- vsaki dve leti od **2014**
- ~100.000 domen
- plitko zajemanje

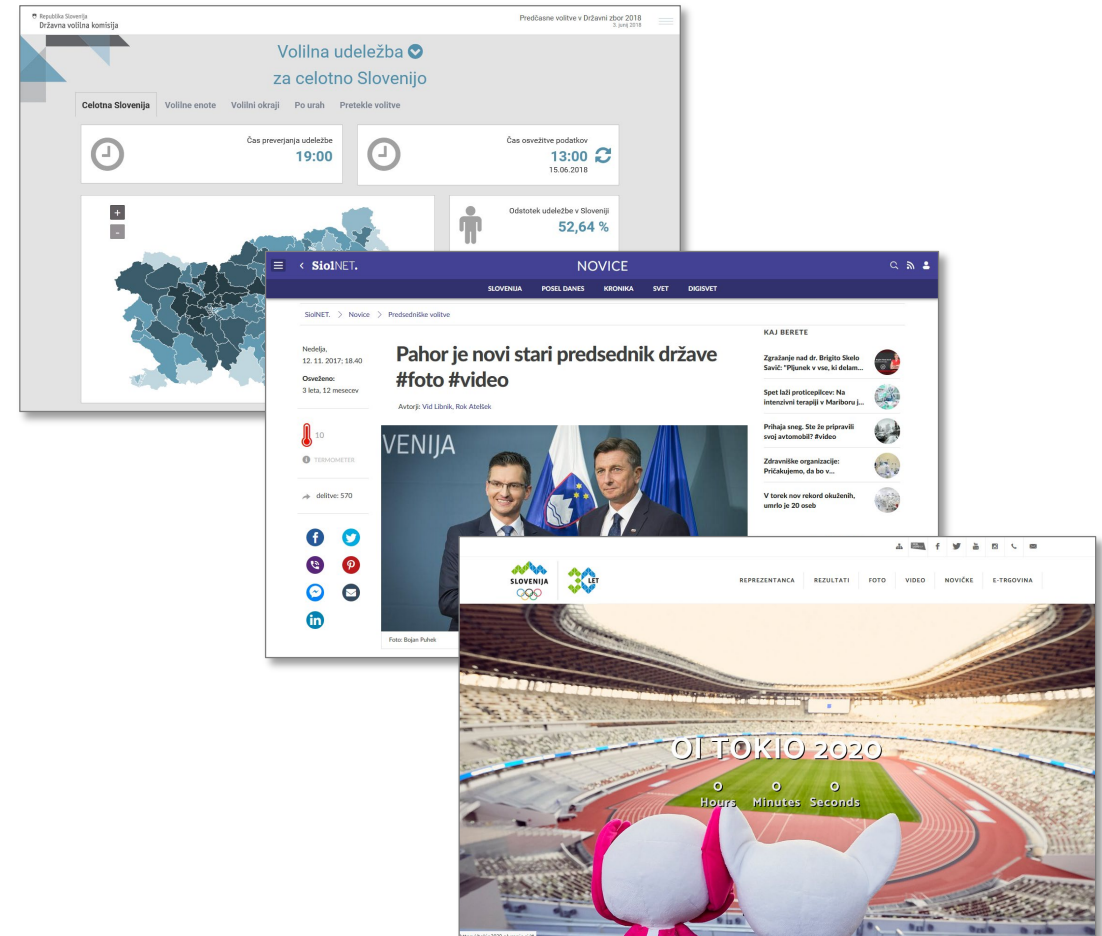




Primer NUK

Tematski zajemi:

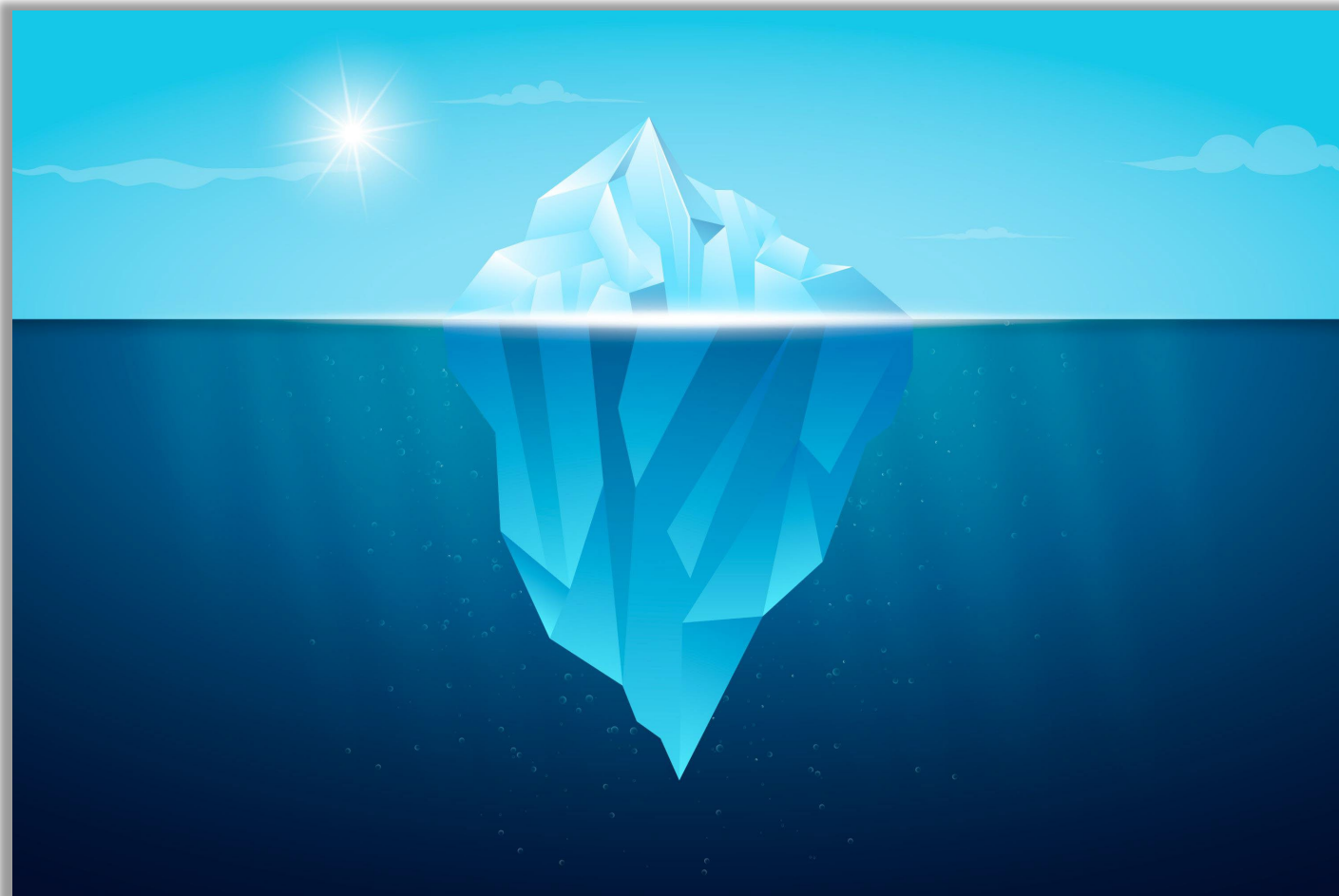
- volitve v državni zbor
- predsedniške volitve
- drugi pomembnejši politični dogodki
- olimpijske igre in drugi pomembni športni dogodki





Tehnične omejitve arhiviranja spleta

Globoki splet





Tehnične omejitve arhiviranja spleta

Pretočne vsebine

The screenshot shows the live website interface. At the top, there is a navigation bar with the text "POKRAJINSKA IN ŠTUDIJSKA KNJIŽNICA MURSKA SOBOTA" and a search bar. Below the navigation bar, there are several menu items: "Knjižnica", "Ponudba", "Izposoja gradiva", "Digitalna knjižnica" (highlighted), "Dogodki in obvestila", and "Moja knjižnica". The main content area is titled "Digitalna knjižnica / Videoposnetki" and features two video thumbnails. The first video is titled "Pravljični večer za odrasle 2015" and shows a man speaking at a podium. The second video is titled "Predstavitve knjige Vlado Kreslin, Slovenia in..." and shows a man in a hat speaking. On the right side, there is a sidebar menu with the following items: "Strokovni članki", "Fotogalerija", "Videoposnetki" (highlighted), "Digitalizirane knjige", "Digitalizirani časopisi", "Katalog informacij javnega značaja", "Oddaljen dostop do podatkovnih zbirk", and "Časopisi na spletu".

Živa stran

The screenshot shows the archived website interface. The navigation bar and menu items are identical to the live page. However, the main content area is titled "Digitalna knjižnica / Videoposnetki" and displays two error messages instead of video thumbnails. The first error message reads: "Prišlo je do napake. Poskusite znova pozneje. (ID predvajanja: 0vXHMND0ZT0uc_oN) Več o tem". The second error message reads: "Prišlo je do napake. Poskusite znova pozneje. (ID predvajanja: 6LZBHRT6qBscKzg) Več o tem". The sidebar menu on the right is also identical to the live page, with "Videoposnetki" highlighted.

Arhivirana stran



Tehnične omejitve arhiviranja spleta

Družbena omrežja

The screenshot shows the live Twitter profile for 'Portal Kamra' (@PortalKamra). The profile includes a header with the website 'www.kamra.si', a profile picture, and a bio: 'Kamra ponuja edinstven dostop do različnega gradiva, ki ga hranijo in z vami delijo slovenske splošne knjižnice, muzeji in druge lokalne kulturne ustanove.' It also shows location (Slovenija), website (kamra.si), and join date (January 2011). The page features a 'Follow' button, a 'New to Twitter?' sign-up section with options for Google, Apple, and phone/email, and a 'You might like' section for 'Mestna knjižnica Ljub...'.

Živa stran

The screenshot shows the archived version of the Twitter profile for 'Portal Kamra'. The page is mostly blank, with a large blue Twitter bird icon in the center and the message: 'Something went wrong, but don't fret — let's give it another shot.' Below the message is a 'Try again' button.

Arhivirana stran



Tehnične omejitve arhiviranja spleta

REPUBLICA SLOVENIJA
MINISTRSTVO ZA KMETIJSTVO,
GOZDARSTVO IN PREHRANO

Domov

- Aktualno
- Javni razpisi PRP
- PRP 2014-2020 do 2022
- Izpostavljene vsebine
- SKP 2023-2027
- Portal znanja

Išči

Dostopnost

Novice

Preberite si aktualne novice

04.10.2022
Predlog Strateškega načrta SKP 2023-2027 potrjen na Vladi RS in posredovan v presojo Evropski komisiji

29.09.2022
Načrt izvajanja finančnih instrumentov v kmetijstvu za obdobje 2023-2027

29.09.2022
LEADER/CLLD projekti, ki ozaveščajo in prispevajo k manj odpadne hrane

29. september 2022
Mednarodni dan ozaveščanja izgubih hrane in odpadki hrane
PREPREČIMO IZGUBE HRANE IN NASTANJANJE ODPADNE HRANE. ZA LJUDI. ZA PLANET. #FUDG

Išči

Dostopnost

Živa stran

Domov

- Aktualno
- Javni razpisi PRP
- PRP 2014-2020 do 2022
- Izpostavljene vsebine
- SKP 2023-2027
- Portal znanja

Išči

Dostopnost

Novice

Preberite si aktualne novice

Vse novice

Arhivirana stran



Vsebinske omejitve arhiviranja spleta





Kaj je arhiviranje spleta?



Skupek kompleksnih, hibridnih in specifičnih metod ohranjanja, ki vključujejo programsko in strojno opremo, spletne robote, algoritme, politike, kuratorje in uporabnike, ki dokumentirajo in hkrati ustvarjajo zgodovino svetovnega spleta.

(Ben David, 2019)



Janko Klasinc
Narodna in univerzitetna knjižnica
01 / 2001 211

janko.klasinc@nuk.uni-lj.si

<http://arhiv.nuk.uni-lj.si/>

