

VALSTYBĖS IR SAVIVALDYBIŲ INSTITUCIJŲ, ĮSTAIGŲ IR ĮMONIŲ INTERNETO SVETAINIŲ IŠSAUGOJIMO REKOMENDACIJOS

I SKYRIUS BENDROSIOS NUOSTATOS

1. Valstybės ir savivaldybių institucijų, įstaigų, įmonių interneto svetainių išsaugojimo rekomendacijos (toliau – Rekomendacijos) nustato valstybės ir savivaldybių institucijų, įstaigų ir įmonių, įrašytų į Lietuvos vyriausiojo archyvaro patvirtintą Valstybės ir savivaldybių institucijų, įstaigų ir įmonių, perduodančių dokumentus valstybės archyvams (toliau – įstaigos), sąrašą, interneto svetainių rekomenduojamą išsaugojimo tvarką.

2. Šiose Rekomendacijose vartojama sąvoka **internetu svetainės momentinė kopija** – tiksliai atgaminta interneto svetainės informacijos kopija.

Kitos Rekomendacijose vartojamos sąvokos atitinka Bendrųjų reikalavimų valstybės ir savivaldybių institucijų ir įstaigų interneto svetainėms apraše, patvirtintame Lietuvos Respublikos Vyriausybės 2003 m. balandžio 18 d. nutarimu Nr. 480, vartojamas sąvokas.

3. Šiomis Rekomendacijomis nesiekama propaguoti ar diskriminuoti vieną ar kitą gamintoją ir jo produktą.

II SKYRIUS INTERNETO SVETAINIŲ PRITAIKYMAS JAS IŠSAUGOTI

4. Interneto svetainės struktūrą rekomenduojama pateikti, interneto svetainės žemėlapyje įtraukiant nuorodas į visą interneto svetainės turinį ir tokiu būdu užtikrinant, kad bus išsaugota visa interneto svetainėje pateikiama informacija: įskaitant ir tą, kuriai nuskaityti reikalingas papildomas veiksmas (vartotojo prisijungimas, pasirinkimas iš sąrašo ir kt.), ir tą, kuri pateikiama tik kaip paieškos rezultatas.

5. Interneto svetainės turinys turėtų būti pasiekiamas iš pradinio interneto svetainės absoliučiojo universaliojo adreso.

6. Esant galimybei, vietoje dinamiškai generuojamų nuorodų turėtų būti naudojamos statinės nuorodos.

7. Jeigu interneto svetainėje naudojami skriptai (pvz., „JavaScript“), turėtų būti sudaryta galimybė svetainės informaciją pasiekti ir paprastu HTML formatu.

8. Jeigu interneto svetainėje yra pateikiamas srautinis vaizdo ar garso turinys (pvz., audiovizualinė medžiaga), turėtų būti sudaryta galimybė šią medžiagą persisiųsti ir kaip atskiras bylas.

9. Informacija, kurias interneto svetainės dalis leidžiama skaityti automatinio būdu, pateikiama „robots.txt“ faile. Informacija, kurią bus draudžiama pasiekti automatinio būdu, išsaugota nebus.

III SKYRIUS INTERNETO SVETAINIŲ MOMENTINIŲ KOPIJŲ SUDARYMAS

10. Rekomenduojama išsaugoti tą įstaigų interneto svetainėse skelbiamą informaciją, kuri atspindi įstaigos veiklos įrodymus (įstaigos veiklos srities naujienos, pranešimai spaudai, renginiai, svarbiausių įvykių suvestinės) ir yra tinkama tyrimams atlikti.

11. Kadangi srautinio turinio vaizdo ar garso informacijai saugoti reikalingi dideli atminties išteklių, rekomenduojama į tai atsižvelgti, vertinant šios informacijos išsaugojimo tikslingumą.

12. Atsižvelgdama į interneto svetainėje pateikiamos informacijos atnaujinimo dažnumą, įstaiga gali nuspręsti, kaip dažnai bus daromos interneto svetainės momentinės kopijos (pvz., 1, 2 kartus per metus ar kas 3 mėnesius).

13. Interneto svetainių momentinėms kopijoms rekomenduojama naudoti šias laisvai platinamas interneto svetainių archyvavimo programines priemones: Heritrix, Web Curator Tool ir Netarchievesuite.

14. Iš interneto svetainės momentinės kopijos turėtų būti galimybė atkurti pirminį interneto svetainės turinį ir struktūrą.

15. Jeigu interneto svetainėje naudojamas kitose interneto svetainėse paskelbtas turinys (pvz., www.youtube.com, www.flickr.com), jo gali nepavykti išsaugoti automatiškai būdu. Tokiu atveju šią informaciją rekomenduojama išsaugoti rankiniu būdu.

IV SKYRIUS

INTERNETO SVETAINIŲ MOMENTINIŲ KOPIJŲ SAUGOJIMAS

16. Interneto svetainės momentines kopijas rekomenduojama išsaugoti WARC formato rinkmenose, atitinkančiose LST ISO 28500:2014 „Informacija ir dokumentavimas. WARC failo formatas (tapatus ISO 28500:2009)“ standarto (toliau – LST SO 28500) reikalavimus (toliau – WARC rinkmena).

17. WARC rinkmenos pavadinime rekomenduojama nurodyti įstaigos pavadinimo trumpinį, interneto svetainės momentinės kopijos išsaugojimo laiką, eilės numerį ir svetainės absoliučiojo universaliojo adreso nuorodą (pvz., LVAT-20140120123000-00001-archyvai.warc).

18. WARC rinkmenų dydis neturėtų viršyti 1 GB. Jei interneto svetainės momentinės kopijos dydis viršija 1 GB, ją rekomenduojama suskaidyti į segmentus, saugomus atskirose WARC rinkmenose.

19. Siekiant sumažinti išteklių, reikalingų interneto svetainių momentinėms kopijoms saugoti, poreikį, WARC rinkmenos gali būti glaudinamos. Tam tikslui rekomenduojama naudoti GZIP formatą, kaip aprašyta RFC1950 „*ZLIB Compressed Data Format Specification version 3.3*“, RFC1951 „*DEFLATE Compressed Data Format Specification version 1.3*“ ir RFC1952 „*GZIP file format specification version 4.3*“ rekomendacijose.

20. Siekiant išsaugoti interneto svetainės momentines kopijas bei užtikrinti jų perskaitomumą, rekomenduojama atsižvelgti į technologijų senėjimą ir programinės įrangos kaitą. Tam tikslui rekomenduojama naudoti WARC rinkmenos įrašų konvertavimo galimybę, kaip aprašyta LST ISO 28500 standarto 6.8 skyriuje.

21. Interneto svetainės momentines kopijas rekomenduojama saugoti įstaigos tarnybinėse stotyse ar laikmenose, skirtose duomenų ilgalaikiam išsaugojimui (pvz., magnetinėse juostose). Laikmenų, kuriose saugomos interneto svetainės momentinės kopijos, informacija ne rečiau kaip kas 3 metus turi būti perrašoma į kitą laikmeną.

22. Rekomenduojama turėti interneto svetainės momentinių kopijų atsargines kopijas, saugomas skirtingose laikmenose.

23. Įstaigos turėtų užtikrinti, kad interneto svetainės momentinės kopijos ir jų atsarginės kopijos būtų saugomos patikimoje ir saugioje aplinkoje.