

– ÍSLENSKT STAFRÆNT ÞJÓÐBÓKASAFN –
Stefna um stafræna endurgerð og varðveislu stafrænna gagna

Landsbókasafn Íslands – Háskólabókasafn er varðveislusafn samkvæmt lögum um skylduskil til safna og tekur til varðveislu gögn sem gefin eru út á Íslandi. Safnið stefnir að því að bjóða aukið aðgengi að íslenskum gögnum safnsins með því að:

- færa þau kerfisbundið í stafrænt form,
- safna og taka á móti útgefnu stafrænu efni,
- varðveita stafrænt efni til framtíðar
- gera stafrænt efni aðgengilegt á Vefnum.

Markmiðið með þróun stafræna þjóðbókasafnsins er þríþætt:

- Miðlun íslenskrar menningar** eins vel og kostur er með því að gera íslenskt efni aðgengilegt notendum utan safnsins. Aðgangur verður í fullu samráði við þá sem eiga höfundarrétt að efninu.
- Varðveisla til framtíðar** með því að varðveita upprunalegu gögnin, þ.e. þau sem berast safninu á hefðbundnu eða stafrænu formi, þau sem verða til við stafræna endurgerð og þeim sem safnað er af vefnum. Frumgögnum verður hlíft en áhersla lögð á notkun stafrænnar útgáfu í staðinn.
- Þjónusta við notendur** hvar sem þeir eru og hvenær sem er. Safnið leitast við að auðvelda notendum að finna stafræna efnið með því að tengja við skráningu þess og lýsigögn og gera það sýnilegt í upplýsingakerfum og leitarvélum. Til að auðvelda erlendum notendum afnot af efninu er stefnt að því að vefviðmót verði bæði á íslensku og ensku.

Aðferðir

Við stafræna endurgerð og miðlun stafræns efnis verður unnið eftir viðurkenndum aðferðum og stöðlum til að tryggja gæði og öryggi. Almennt er gert ráð fyrir að efnið verði aðgengilegt um vefi safnsins og að það verði öllum opið án endurgjalds. Þó má gera ráð fyrir að í sumum tilvikum verði einungis heimilt að veita aðgang að efni innan safnsins. Efnið er framleitt á mismunandi formi s.s. JPEG, Pdf, Pdf/A, HQ Pdf, og TXT.

Notaður er samskiptastaðallinn OAI/PMH (Open Archives Initiative / Protocol for Metadata Harvesting) í því skyni að unnt verði að bjóða upp á samleit í öllu stafrænu efni safnsins óháð skráningarformi og að safnið geti skipst á gögnum við samstarfsaðila svo sem Landskerfi bókasafna, Evrópubókasafnið TEL og Europeana. Til að leita í efninu er allur prentaður texti ljóslesinn (Optical Character Recognition – OCR) og þannig gerður tölvutækur. Einstök skjöl auðkennd með XML kóða þannig að unnt sé að þekkja þau aftur og leita í þeim. Notuð eru bókfræðileg gögn, lýsigögn og XML kóðun.

Allt efni í stafrænni endurgerð fer í gegnum vinnslulínu safnsins til að tryggja samræmda vinnslu. Því er síðan miðlað á mismunandi vefjum. Stefnt er að því að efnið verði aðgengilegt í einni samræmdri íslenskri leitargátt fyrir stafrænt menningarefni. Við varðveislu efnisins eru unnin tvö eintök. Annað eintakið verður vistað hjá viðurkenndum hýsingaraðila. Stuðlað verður að því að taka í notkun Creative Commons auðkenni þar sem það á við.

1. útgáfa samþykkt á framkvæmdaráðsfundi 26. júní 2006
2. útgáfa lögð fyrir stjórnarfund 11. júní 2008
3. útgáfa mars 2011.

Til þess að ná þessu markmiðum hefur safnið skipt verkefninu í nokkrar skilgreindar einingar. Framkvæmd fer eftir fjárveitingum á hverjum tíma en einnig eftirspurn, varðveislusjónarmiðum, hversu fágætt efnið er og notagildi.

Efni sem nú er til í stafrænu formi og aðgengilegt á vefnum:

- **Íslandskort (1997)**; 220 kort útgefin 1540-1900, sjá www.islandskort.is
- **Íslensk dagblöð og tímarit (2002)**; 3.5 milljón bls. af íslenskum dagblöðum og tímaritum frá upphafi til 1920, ásamt efni frá 20. öld, sjá www.timarit.is
- **Íslensk handrit (2009)**; 300.000 bls. handrita frá Lbs-Hbs og Árnastofnunum í Reykjavík og Kaupmannahöfn, ásamt efni frá Sagnaneti, sjá www.handrit.is
- **Íslenskar bækur (2010)**; 65.000 bls. íslenskra bóka frá upphafi og bækur sem tengjast Íslandi og Íslendingum, sjá www.baekur.is
- **Íslenski vefurinn (;** milljarður vefsíðna og annarra gagna sem birt eru eða gerð aðgengileg á þjóðarléninu .is frá 2004, sjá www.vefsafn.is
- **Efni frá háskólum**

Verkefni sem unnið er að og fyrirhuguð verkefni:

- **Íslandskort**; Herforingjaráðskort og Atlaskort til 1920.
- **Íslensk dagblöð og tímarit**; Stafræn endurgerð tímarita frá 20. öld í samstarfi við útgefendur og rétthafa. Greiniskráning ritrýndra tímarita og aðgangur að einstökum greinum.
- **Íslensk handrit**; Handritasafn Hins íslenska bókmenntafélags, Reykjavíkurdéildar (ÍBR) (290 handrit um 80.000 bls.) verður yfirfært á stafrænt form. Handritasafn Jóns Sigurðssonar (JS) (1000 handrit) og handritasafn Landsbókasafns (Lbs) (490 handrit).
- **Íslenskar bækur**; Stafræn endurgerð allra bóka útgefna fyrir 1844. (150.000 bls.)
- **Íslenski vefurinn**; Reglur og viðmið um vefsöfnun verði endurskoðaðar og unnið að bættu aðgengi.
- **Móttaka á útgefnu efni á pdf-formi**; Þróun Rafhlöðunnar og Skemmunnar, sem eru stafræn geymslusöfn og móttaka efnis á pdf formi.
- **Íslensk tónlist**; samstarfsverkefni með RÚV. Efni á hljómplötum og snældum.
- **Viðtöl**; Viðtöl og frásagnir sem safnað hefur verið á vegum Miðstöðvar munnlegrar sögu hafa verið sett í stafrænt form en er ekki aðgengilegt almenningi nema sem sýnishorn.

Framtíðarverkefni til lengri tíma

- Íslenskar bækur útgefna 1845-1900 (um 1.000 bækur, 250.000 bls.).
- Íslenskar bækur útgefna 1900-1949 (13.000 bækur, um 2.500.000 bls.).
- Önnur handritasöfn Landsbókasafns (13.500 handrit, um 3.600.000 bls.).
- Landsmálablöð og tímarit 1920-2000.
- Bækur útgefna 1950-2000 (58.000 bækur, um 11.6 millj. bls.).
- Kort útgefin 1920-2000.
- Myndefni og veggspjöld.
- Tónlistar- og hljóðefni.

1. útgáfa samþykkt á framkvæmdaráðsfundi 26. júní 2006

2. útgáfa lögð fyrir stjórnarfund 11. júní 2008

3. útgáfa mars 2011.