

راهکارهای برنامه‌های برتر حفاظت وب در مقابله با چالش‌های فنی حفاظت

هاجر ستوده^۱، کیانوش رشیدی^{۲*}

تاریخ دریافت: ۹۲/۱۰/۲۰ تاریخ پذیرش: ۹۳/۱/۲۵

چکیده

هدف: هدف از پژوهش حاضر، بررسی برنامه‌های حفاظت وب به منظور شناسایی راهکارهای آنها در مواجهه با چالش فنی است.

روش: گردآوری داده‌ها به روش متن‌پژوهی انجام شد. به این ترتیب که برای شناسایی و بررسی راهکارهای چالش فنی، گستره‌ای از متون اعم از مقالات، دستنامه‌ها و مستندات و وب‌سایت‌های برنامه‌های برتر جست‌وجو و مطالعه شدند. همچنین برای تکمیل اطلاعات، با مسئولان برنامه‌ها تماس گرفته شد. همچنین، برنامه‌های برتر حفاظت وب از طریق جست‌وجو در گوگل، مستندات، مقالات و راهنماهای معتبر شناسایی شدند. پس از بررسی و حذف برنامه‌های به نسبت ضعیف‌تر، در نهایت ۶۱ برنامه برتر شناسایی و بررسی شدند.

یافته‌ها: دشواری‌های حفاظت از نظر فنی که در برنامه‌های برتر حفاظت وب بررسی قرار شدند عبارتند از: اجزای حفاظت‌شده، راهبرد حفاظت، روش حفاظت از نسخه‌ها، فراداده، نوع فراداده، روش ایجاد فراداده، خدمات جست‌وجو و مرور، انواع سبک‌های جست‌وجو و مرور، خدمات نمایه‌سازی، فهرست‌نویسی و رده‌بندی، نرم‌افزارهای دروگری، نرم‌افزارهای دسترسی، نرم‌افزارهای نمایه‌سازی و مدل حفاظتی. برنامه‌های برتر راهکارهای مختلفی را برای مواجهه با چالش‌های فنی در پیش گرفته‌اند. راهکارهای کنونی نیز از وضعیت آرمانی فاصله دارند و به دلیل داشتن نقاط قوت و ضعف خاص، راهکار واحد و استاندارد برای چالش‌ها وجود ندارد.

واژه‌های کلیدی: برنامه‌های حفاظت، چالش فنی، حفاظت وب.

مقدمه

فراهم‌آوری دسترسی به گستره‌ای از منابع پر محتوا و غنی و تضمین پایداری این دسترسی، از رسالت‌های کتابخانه‌هاست. در سراسر تاریخ، کتابخانه‌ها با حفاظت از منابع، نقش سرنوشت‌سازی در حفظ میراث معنوی ملت‌ها ایفا کرده‌اند. بخشی از میراث معنوی نسل‌های امروز در قالب منابع وبی منتشر می‌شود. اما منابع وبی به دلیل پویایی محتوا، کوتاهی عمر رسانه‌ها، منسوخ شدن سخت‌افزار و نرم‌افزار، دستبرد و رسوخ دزدان اطلاعات^۱، تهدیدهای ویروسی، میرایی وب‌سایت‌ها، دزدی اطلاعات و... در خطر نابودی هستند (هدسترام^۲، ۱۹۹۸). بسیاری بر این باورند که در گذر زمان تغییرات فنی اینترنت (به‌لحاظ معماری زیربنایی، برنامه‌های کاربردی، پروتکل‌ها و استانداردها) چنان سریع روی می‌دهد که چرخه زندگی استانداردهای رمزگذاری اطلاعات، از چرخه زندگی خود اطلاعات کوتاه‌تر شده است (اشلی و همکاران^۳، ۲۰۱۰). این تغییرات سبب می‌شود نه تنها نسخه‌های زنده بلکه دسترس‌پذیری منابع آرشیوی نیز دچار تهدید شود (تامپسون^۴، ۲۰۰۸).

از این‌رو، کتابداران و آرشیویست‌ها (به‌ویژه در کتابخانه‌های ملی) ناگزیرند اطلاعات دیجیتالی را نیز همانند دیگر منابع، حفاظت و دسترسی به آنها را برای نسل‌های آینده امکان‌پذیر کنند (پارخ^۵، ۲۰۰۹). اما تفاوت سرشتی منابع دیجیتالی و منابع سنتی و نیز فناوری‌های مورد استفاده موجب می‌شود حفاظت مواد دیجیتالی و به‌ویژه وبی ماهیت بسیار متفاوتی داشته باشد. بنابراین، نمی‌توان رویه‌ها، استانداردها و طرح‌های متداول در دنیای سنتی را (که برای گزینش، گردآوری، سازماندهی و حفاظت مواد به کار می‌روند) در دنیای وب نیز استفاده کرد (امانوئل و گیلداس^۶، ۲۰۰۹).

نگهداشت مواد دیجیتالی فرآیند چندبعدی و بسیار پیچیده‌ای است که در هر مرحله، از گزینش گرفته تا مدیریت دسترسی، با چالش‌های بسیاری روبه‌رو است. در میان چالش‌های مختلف حفاظت وبی که حوزه‌های مجموعه‌سازی، حقوقی، فنی، دسترسی، مالی و سازمانی را در بر می‌گیرد (رشیدی، ۱۳۹۰)، چالش‌های فنی از مهم‌ترین و بنیادی‌ترین چالش‌ها به‌شمار می‌آیند که شاید بیش از همه بر سرنوشت برنامه‌های حفاظت تأثیر بگذارند (اسمیت^۷، ۲۰۰۳). زیرا بزرگی اندازه وب، گوناگونی محتوا، گستردگی و تنوع قالب‌ها، و بی‌دوامی و بی‌ثباتی منابع دیجیتالی (خواه به لحاظ محتوا، نشانی، رسانه، قالب یا فناوری) به کارگیری فناوری‌های بسیار پیشرفته را ضروری می‌کند. این در حالی است که پیشرفت‌ها و دستاوردهای فناورانه در حوزه حفاظت هنوز نابالغ به‌شمار می‌آیند (هوج^۸، ۲۰۰۴؛ تامپسون، ۲۰۰۸). از این‌رو، تاکنون الگوی استاندارد برای حفاظت منابع وبی پدید نیامده است. تعداد فزاینده طرح‌های آرشیوگری وب در سراسر جهان،

گرچه نشان از آگاهی روزافزون از اهمیت راهبردی حفاظت وب برای نسل‌های آینده دارد، متضمن آن است که هیچ‌کدام از الگوهای آرشیوگری به‌تنهایی برای حفاظت از میراث پیوسته ملی و جهانی رضایت‌بخش نیستند (کتابخانه ملی استرالیا^۹، ۲۰۰۳). پس، این پرسش پیش می‌آید که کدام برنامه‌های حفاظت وب، چالش‌های متعدد پیش‌رو را مد نظر قرار داده‌اند و برای آنها راهکار اندیشیده‌اند. شناخت این برنامه‌ها، ضمن کمک به آشنایی با تجارب جهانی، زمینه را برای عملیاتی کردن حفاظت منابع وبی در ایران فراهم می‌آورد.

در این مقاله سعی شده است ضمن شناسایی برنامه‌های برتر حفاظت وب، راهکارهای این برنامه‌ها برای مقابله با چالش فنی بررسی شوند. چالش‌های فنی بسیار گسترده و گوناگونی فراراه حفاظت وب قرا گرفته‌اند. بررسی متون و ادبیات این حوزه نشان می‌دهد که از میان چالش‌های فنی، برخی مانند تعیین اجزای مورد حفاظت، راهبرد حفاظت، روش حفاظت از نسخه‌ها، فراداده، نوع فراداده، روش ایجاد فراداده، خدمات جست‌وجو و مرور، انواع سبک‌های جست‌وجو و مرور، خدمات نمایه‌سازی، فهرست‌نویسی و رده‌بندی، نرم‌افزارهای دروگری^{۱۰}، نرم‌افزارهای دسترسی و نرم‌افزارهای نمایه‌سازی بیش از همه بر زیست‌پذیری مواد وبی و آرشیوشده و در نتیجه سرنوشت برنامه حفاظت تأثیر گذارند^{۱۱}. از این‌رو، در مقاله حاضر تنها این مسائل فنی بررسی شده‌اند که در متون و ادبیات این حوزه بیش از همه بر آنها تأکید بوده است.

پیشینه پژوهش

بخشی از پژوهش‌ها در زمینه مسائل فنی حفاظت، بر مشکلات ناشی از قالب‌ها و رسانه‌ها متمرکز شده‌اند. برای نمونه، کانی^{۱۲} (۱۹۹۷)، به تشریح برخی چالش‌های فنی نظیر تنوع رسانه‌ها و قالب‌ها می‌پردازد. کریستنس^{۱۳} (۲۰۰۴)، قالب‌های فایلی را بررسی و چالش مربوط به تنوع قالب‌ها در آرشیوهای وبی را تجزیه و تحلیل کرده است. بارو^{۱۴} (۲۰۰۷)، بر این باور است که قالب‌های فایلی نقش مهمی در حفاظت دیجیتال بازی می‌کنند. کاپلان^{۱۵} (۲۰۰۸)، به معرفی پروژه تی.آی. پی. آر^{۱۶} می‌پردازد که با هدف شناسایی مسائل مربوط به انتقال قالب‌ها و بسته‌های اطلاعاتی در میان سپاره‌ها طراحی شده است.

بخش دیگری، به راهبردهای حفاظت پرداخته‌اند. از این جمله است پژوهش ملر، ویتلی، سرجنت^{۱۷} (۲۰۰۲) که اشکالات بسیاری را بر راهبردهای کنونی حفاظت وارد می‌دانند. این پژوهش تکنیک‌های پیشرفته مهاجرت داده‌ها را توصیف می‌کند و راهبرد مهاجرت هنگام درخواست را توضیح می‌دهد. کلی^{۱۸} (۲۰۰۲)، راهبردهای گوناگون حفاظت از وب‌سایت‌ها را

بیان می‌کند و بعضی از چالش‌های فنی را به تصویر می‌کشد. کاستاتوپوس، دور و پتراسکی^{۱۹} (۲۰۰۵)، الگویی را برای حفاظت دیجیتال معرفی و راهبردهای حفاظت دیجیتال را تجزیه و تحلیل می‌کنند. میکائیل، گرین، ماسلاو، فیلیپس و لگت^{۲۰} (۲۰۰۷)، با به تصویر کشیدن نتایج یک مورد کاوی، مسائل و مشکلات حفاظت در کتابخانه دیجیتال را بررسی کردند. آنان به معرفی دو راهبرد مهاجرت^{۲۱} و تقلیدگری^{۲۲} می‌پردازند.

دسته دیگری از آثار، به بررسی فراداده‌ها و اهمیت آنها در حفاظت وب پرداخته‌اند. برای نمونه دی (۱۹۹۸)، درباره نقش فراداده‌ها در حفاظت دیجیتال بحث کرده است و راهبردهای حفاظت دیجیتال را بر می‌شمرد و فراداده‌ها را در مدل ا.ی. آی. اس^{۲۳} توضیح می‌دهد. گروه کاری اسی. ال. سی/آر. ال. جی (۲۰۰۱)، فراداده‌های حفاظتی برای اشیای دیجیتال را به تصویر می‌کشد و چارچوب فراداده حفاظتی را تعریف و مدل ا.ی. آی. اس را بررسی قرار می‌کند. دی (۲۰۰۶)، برخی استانداردهای فراداده‌ای را توصیف می‌کند و بر دیدگاه برخی توسعه‌دهندگان، درباره اهمیت نگهداری و بازاستفاده از فراداده‌های موجود تأکید می‌کند.

دسته دیگری، از پژوهش‌ها به بررسی مسائل فنی می‌پردازند که در جریان طراحی، راه‌اندازی و اجرای برنامه‌ها یا پروژه‌های حفاظت وبی روی داده‌اند. برای نمونه، وتلی^{۲۴} (۲۰۰۴)، پروژه کامپیون^{۲۵} را بررسی می‌کند که در سال ۱۹۹۹ در دانشگاه لیدز^{۲۶} تأسیس شد و بعدها به بی. بی. سی دمسدی^{۲۷} تغییر نام داد و چالش‌های فنی فراروی آن را به تصویر می‌کشد و در ادامه پیشرفت‌هایی را بیان می‌کند که در راهبردهای فنی آن حاصل شده است. مانیاتیس، گروه تحقیقاتی اینتل، راسپولوس، دانشگاه هاروارد، گیلی، دانشگاه استنفورد، رزنتال، کتابخانه‌های دانشگاه استنفورد، بیکر^{۲۸} (۲۰۰۵)، سامانه حفاظت دیجیتال «لاکس»^{۲۹} را بررسی کرده است و اصول طراحی و مشکلات فنی این پروژه را شرح می‌دهد. زییکا و ماتجکی^{۳۰} (۲۰۰۷)، ضمن معرفی پروژه آرشیوگری وب ملی چک و اسلواکی، ویژگی فنی آرشیو را بررسی کرده‌اند و به چالش بزرگی اندازه و تعدد قالب‌های فایلی پرداخته‌اند. رابر (۲۰۰۸)، ضمن بررسی برخی چالش‌های تحقیقاتی در حفاظت دیجیتال و ضرورت تحقیقات زیرساختی و فناوری برای دستیابی به راهکارها، به اختصار فعالیت‌های بخش فناوری نرم‌افزار و سیستم دانشگاه وین را در این رابطه شرح می‌دهد.

جامعه آماری و روش پژوهش

برنامه‌های حفاظت وب برتر دنیا، جامعه این پژوهش را تشکیل دادند. منظور از برنامه‌های برتر دنیا، برنامه‌هایی است که در فهرست نتایج گوگل در صدر واقع شده‌اند یا در مستندات، مقالات و

راهکارهای برنامه‌های برتر حفاظت وب در مقابله با چالش‌های فنی حفاظت

راهنماهای معتبر به‌عنوان برنامه‌ای مهم معرفی شده‌اند. نمونه‌گیری این پژوهش از نوع هدفمند به‌شمار می‌آید.

اسامی این برنامه‌ها در جدول ۱ بیان شده است و در مواردی که به ذکر نمونه نیاز باشد، به دلیل رعایت اختصار از بیان نام آنها خودداری می‌شود و به ذکر شماره ردیف آنها در جدول ۱ بسنده خواهد شد.

جدول ۱. برنامه‌های برتر حفاظت منابع وبی

نام برنامه	ردیف	نام برنامه	ردیف
National Library of Iceland- Electronic Legal Deposit	۲۴	Anarchism web archive	۱
Netarkivet. dk (Royal library and the state and university library)	۲۵	AOLA (Austrian online archive)	۲
New Zealand Domain harvest	۲۶	Archipol	۳
OASIS (online archiving & searching internet sources)	۲۷	Archiv (Archive of the Czech web)	۴
Our Digital Island	۲۸	Archived Harvard's Website	۵
PADICAT (patrimoni digital de catalunya)	۲۹	BNF (national library of France)	۶
PANDORA Australia's Web Archive	۳۰	British Slave Trade Legacies	۷
Portuguese Web Archive	۳۱	Citeseer	۸
Program e-Helvetic of the Swiss National Library	۳۲	DACHS (digital archive for Chinese studies)	۹
Slovenia	۳۳	Dept. of labor web archiving project	۱۰
STORS	۳۴	Diva	۱۱
The CyberCemetery	۳۵	European archive	۱۲
The National Library of Wales Web Archiving Programme	۳۶	Finnish Web Archive	۱۳
UK Government Web Archive	۳۷	GCWA (Government of Canada Web Archive)	۱۴
UK Web Archive	۳۸	Ina (institut national de l'audio visuel)	۱۵
University of Michigan Web Archives	۳۹	Internet Archive	۱۶
VERS (Victorian electronic records strategy)	۴۰	k-12 web archiving	۱۷
WAS (The Web Archive Singapore)	۴۱	Kb web archiving	۱۸
Web Archive Norway	۴۲	Kulturarw	۱۹
Web Archiving @ the University of Melbourne	۴۳	LAC (Electronic collection of library and archives canada)	۲۰
web archiving project	۴۴	lawap (latin american web archiving project)	۲۱
Web harvesting project of the German National Library	۴۵	Mellon Project on Web Resource Collection Program Development	۲۲
WebCite	۴۶	Minerva (Library of Congress Web Archives)	۲۳

برای شناسایی راهکارهایی که برای چالش مسائل فنی وجود دارند، مستندات شناسایی شده و متون به دست آمده، درباره برنامه‌ها مطالعه شد. به منظور تکمیل اطلاعات، کوشش شد خود برنامه‌ها نیز بررسی شوند. در بسیاری از موارد اطلاعات به طور کامل حاصل نشد. از این رو برای تکمیل اطلاعات، طی تماسی با مسئولان برنامه‌ها از آنها درخواست شد که فهرست واریسی محقق ساخته را کامل کنند. ۲۵ برنامه، این فهرست را تکمیل کردند.

یافته‌ها

اجزای حفاظت شده

اهمیت محتوا، بافتار، ساختار و ویژگی دیداری منابع وبی برای برنامه‌های مختلف متفاوت است. بیشینه برنامه‌هایی که اطلاعاتی را در مورد اجزای حفاظت شده در اختیار قرار داده‌اند، هر سه مورد محتوا، ساختار و بافتار را گردآوری می‌کنند. این برنامه‌ها عبارتند از برنامه‌های ۱۲، ۱۳، ۱۴، ۲۲، ۲۳، ۲۴، ۲۸، ۲۹، ۳۰، ۳۳، ۳۵، ۳۶، ۳۷، ۴۲ و ۴۵. برنامه‌های ۳۴ و ۴۴ تنها از محتوای منابع وبی حفاظت می‌کنند. برنامه‌های ۳۱ و ۳۸ از محتوا و ساختار منابع، و برنامه ۱۱ هم از محتوا و هم از بافتار حفاظت می‌کند. برنامه‌های ۱۸ و ۳۲ علاوه بر محتوا، بافتار و ساختار منابع، از نمای دیداری هم حفاظت می‌کنند.

راهبرد حفاظت

برنامه‌های حفاظت منابع وبی از روش‌های حفاظت جریان بیتی‌ها، تقلیدگری، مهاجرت، تازه‌سازی^{۳۱} و قرینه‌سازی^{۳۲} استفاده می‌کنند. انتخاب راهبرد حفاظت یکی از بزرگ‌ترین دغدغه‌های طراحان و برنامه‌ریزهای حفاظت وب است. بیشینه برنامه‌ها (۲۷ مورد) اطلاعاتی را در این زمینه، خواه در مستندات خود و خواه در پاسخ به فهرست واریسی ارائه نکرده‌اند. با توجه به اهمیت این مسئله، مسکوت گذاردن آن به‌ویژه در مستندات مربوط چندان قابل انتظار نیست. چه بدون داشتن راهبردی برای حفاظت، کارکرد برنامه به یک فعالیت گردآوری صرف تقلیل می‌یابد. این احتمال نیز وجود دارد که برنامه‌ها کار را با گردآوری آغاز کنند و پس از توسعه مجموعه، راهبردهای آزمایش شده را بر آنها اعمال دارند. به هر حال، یافتن اطلاعات دقیق‌تر در این زمینه به تحقیقات بیشتر نیاز دارد.

دیگر برنامه‌ها از یک یا چند راهبرد حفاظت جریان بیتی، تقلیدگری، مهاجرت، تازه‌سازی و قرینه‌سازی بهره می‌گیرند. راهبرد مهاجرت از بیشترین فراوانی برخوردار است. ۷ برنامه به طور منفرد از این راهبرد استفاده کرده‌اند و شماری از برنامه‌ها نیز، تلفیقی از آن را با دیگر راهبردها

به کار گرفته‌اند. برنامه‌های ۲، ۳، ۴، ۱۳، ۱۹، ۲۶ و ۳۴ فقط از روش مهاجرت برای حفاظت استفاده می‌کنند. برنامه‌های ۳۰ و ۳۳ هم از روش تقلید‌گری و هم از روش مهاجرت بهره می‌برند. شایان ذکر است که راهبرد حفاظت برنامه پاندورا هنوز به‌طور قطع مشخص نشده است و احتمالاً هر دو راهبرد به کار گرفته می‌شود. برنامه ۱۱، از روش مهاجرت، تازه‌سازی و قرینه‌سازی و برنامه ۳۵ از روش مهاجرت، تقلید‌گری و تازه‌سازی استفاده می‌کنند. برنامه ۴۵ از روش‌های مهاجرت، تقلید‌گری، تازه‌سازی و قرینه‌سازی استفاده می‌کند. برنامه ۳۲ روش‌های مهاجرت و قرینه‌سازی را در دستور کار دارند، گرچه تاکنون در عمل از روش مهاجرت استفاده نکرده است. برنامه ۲۳ تنها از جریان بیتی حفاظت می‌کند، برنامه ۴۲ از روش تقلید‌گری و برنامه ۲۸ از روش‌های تقلید‌گری و قرینه‌سازی بهره می‌برد. برنامه‌های ۱۲، ۲۴ و ۳۱ از روش قرینه‌سازی استفاده می‌کنند. برنامه ۱۴ هنوز راهبردی قطعی برای حفاظت ندارد. در برنامه ۱۸ بیان شده که تاکنون به گردآوری منابع پرداخته و هنوز در عمل راهبرد حفاظتی خاصی برای منابع دروشده به کار نبرده است. با این حال، راهبردهای مهاجرت و تقلید‌گری را مفید و در دست بررسی اعلام کرده‌اند.

روش حفاظت از نسخه‌ها

در حفاظت از نسخه‌ها، می‌توان دو رویکرد افزایشی یا جایگزینی را در پیش گرفت. همه ۲۲ برنامه‌ای که به این پرسش پاسخ داده‌اند، اعلام کرده‌اند که از روش افزایشی استفاده می‌کنند، به این معنا که همه نسخه‌های قدیمی وب‌سایت‌ها را نگهداری می‌کنند و نسخه‌های جدید را جایگزین نسخه‌های قدیمی نمی‌کنند. این برنامه‌ها عبارتند از ۱۱، ۱۲، ۱۳، ۱۴، ۱۸، ۲۲، ۲۳، ۲۴، ۲۶، ۲۸، ۳۰، ۳۱، ۳۲، ۳۳، ۳۴، ۳۵، ۳۶، ۳۷، ۳۸، ۴۲، ۴۴ و ۴۵. احتمال می‌رود که در برنامه‌های دیگری که اطلاعاتی در این زمینه در مستندات خود ذکر نکرده‌اند، این امر بدیهی فرض شده و با توجه به هدف برنامه‌ها که حفاظت وب برای نسل‌های آینده است، روش افزایشی را در پیش گرفته باشند. به‌ویژه آنکه بسیاری از برنامه‌های پاسخ‌دهنده نیز اذعان کرده‌اند که دلایل آنان برای در پیش گرفتن این روش، آن بوده است که عمر این گونه برنامه‌ها هنوز چنان زیاد نیست که با مسائلی مانند تورم آرشیو یا کمبود فضا و مانند آن مواجه شوند و بخواهند نسخه‌های قدیمی را وجین کنند.

خدمات ارزش افزوده

دسترسی به آرشیوی غنی از مواد حفاظت‌شده بدون امکان دسترسی به محتوای منابع سودی ندارد.

از این رو، آرشیوها علاوه بر گردآوری و حفاظت از منابع، ناچارند به منظور تضمین بهره‌وری آنها، خدمات ارزش افزوده‌ای را نیز ارائه کنند. یکی از این خدمات تخصیص فراداده است که زمینه‌ساز فراهم‌آوری خدمات دیگر و از جمله نمایه‌سازی است. علاوه بر این شاید برنامه‌ها به روش‌های معمول دست به فهرست‌نویسی و رده‌بندی منابع بزنند. حاصل این خدمات، خدمات جست‌وجو و مرور است که در بخش‌های آینده بررسی خواهند شد.

فراداده

تولید یا استخراج و تخصیص فراداده‌ها بخش دیگری از پرسش‌های این پژوهش بود. برنامه‌ها طیف وسیعی از روش‌ها و از جمله استفاده از برجسب اچ. تی. ام. ال یا ایکس. ام. ال، فراداده‌های خودساخته و فراداده‌های استاندارد چون سرعنوان‌های موضوعی، مارک، هسته دابلین، متس^{۳۳} و مادز^{۳۴} استفاده کرده‌اند. جدول ۲ وضعیت برنامه‌ها را به لحاظ نوع فراداده‌های مورد استفاده، به تصویر می‌کشد.

جدول ۲. فراداده‌های استفاده‌شده در برنامه‌ها

فراداده	فراوانی	درصد
برجسب اچ. تی. ام. ال / ایکس. ام. ال	۲	۳/۴
فراداده‌های خودساخته	۵	۱۰/۸
فراداده‌های استاندارد	۱۱	۲۳/۹
بدون فراداده	۸	۱۷/۴
نامعلوم	۲۰	۴۳/۵
جمع	۴۶	۱۰۰

چنانکه مشاهده می‌شود اطلاعات مربوط به فراداده‌ها در مورد حدود نیمی از برنامه‌ها به دست نیامد. پاسخگویان نیز در رفع کمبود اطلاعات در این زمینه کمکی نکردند. با این حال، بررسی آن دسته از برنامه‌هایی که اطلاعاتی را در این زمینه فراهم کرده بودند، نشان داد که شمار اندکی از آنها از هیچ‌گونه فراداده‌ای استفاده نکرده‌اند. اینها شامل هشت برنامه ۱۳، ۱۴، ۱۸، ۲۴، ۲۶، ۳۱، ۴۲ و ۴۴ هستند. ۱۷ برنامه برای دسترسی و استفاده از منابع آرشیو شده، از فراداده استفاده می‌کنند. برنامه‌های ۴، ۱۲، ۱۹، ۲۲، ۲۳، ۲۷، ۲۸، ۳۰، ۳۲، ۳۴ و ۳۵ از فراداده‌های استاندارد استفاده کرده‌بودند. از این میان، برنامه ۲۲ از فراداده مارک و هسته دابلین، برنامه ۱۹، از فراداده هسته

دابلین، برنامه ۴، از مارک ۲۱ و هسته دابلین، برنامه ۳۴ از مارک آمریکا، برنامه ۲۷، از هسته دابلین و برنامه ۲۳ از مدز (ترکیبی از فراداده خودساخته و مشتق از آرشیو) استفاده کرده‌اند. برنامه ۳۲ از فراداده «ان. ال. ان. زد»^{۳۵} و «متس» استفاده می‌کند. برنامه ۲۸ از دابلین کور و «ای. سی. آی. اس»^{۳۶} بهره می‌برد. معدودی از برنامه‌ها (شامل برنامه ۳۰ و ۳۷) فراداده‌ها را از برجسب اچ. تی. ام. ال/ایکس. ام. ال به دست می‌آورند. برنامه‌های ۱۱، ۳۳، ۳۶، ۳۸ و ۴۵ از فراداده‌های خودساخته استفاده می‌کنند. شایان ذکر است که برنامه ۳۸، از سرعنوان‌های موضوعی خودساخته بهره می‌برد.

نوع فراداده

مستندات برنامه‌ها، هیچ‌گونه اطلاعاتی درباره نوع فراداده‌ها اعم از توصیفی، ساختاری، مدیریتی و حفاظتی در اختیار قرار ندادند. در عین حال، مسئولان تماس نیز این مورد را در فهرست وارسی بی‌پاسخ گذاشتند، این امر شاید به دو دلیل باشد، نخست، ضعف مستندات در پرداختن به جزئیات فنی و نیز عدم اشراف مسئولان تماس به این جزئیات فنی. این ضعف به‌ویژه با توجه به اهمیت بالایی که در متون برای انواع مختلف فراداده‌ها قائل شده‌اند، انتظار نمی‌رود. انتظار این است که برنامه‌ها به این مهم توجه داشته باشند و با توجه به اهداف آنها، به‌ویژه فراداده‌های حفاظت را به‌عنوان یکی از انواع اصلی فراداده‌های خود لحاظ کرده باشند. با توجه به حساسیت این امر، انجام دادن تحقیقات بیشتر با روش‌های گردآوری متفاوت، برای روشن شدن این بعد از مسئله ضروری است.

روش ایجاد فراداده

ایجاد یا استخراج فراداده شاید به‌صورت خودکار (به‌وسیله ماشین) یا دستی (توسط انسان) صورت گیرد. برنامه‌های ۴، ۸، ۱۲، ۱۳، ۱۹، ۲۴، ۲۹، ۳۰، ۳۲ و ۴۵ به‌صورت خودکار فراداده‌ها را ایجاد می‌کنند و برنامه‌های ۱۱، ۱۸، ۲۲، ۲۳، ۲۸، ۳۳، ۳۴، ۳۵، ۳۶، ۳۸ و ۴۴ به‌صورت دستی به ایجاد فراداده می‌پردازند. تنها در دروگری گزینشی در برنامه ۲۶، تخصیص فراداده، هم به‌صورت دستی و هم خودکار صورت می‌گیرد.

خدمات جست‌وجو و مرور

بررسی برنامه‌ها نشان داد که برای دسترسی به منابع وبی، خدمات جست‌وجو و مرور ارائه شده است. همان‌گونه که از جدول ۳ برمی‌آید، بیش از نیمی از این برنامه‌ها، هر دو نوع خدمات را ارائه کرده‌اند. پس از آن برنامه‌هایی هستند که خدمات جست‌وجو (۱۳/۰۴ درصد) را ارائه می‌دهند. به این ترتیب، مشاهده می‌شود که خدمات جست‌وجو در این برنامه‌ها، بیش از همه مورد توجه بوده است.

جدول ۳. خدمات ارائه شده

خدمات	فراوانی	درصد
جست‌وجو	۶	۱۳/۰۴
مرور	۳	۶/۵
هر دو	۲۷	۵۸/۷
هیچ	۱	۲/۲
نامعلوم	۹	۱۹/۶
جمع	۴۶	۱۰۰

کاربران برنامه‌های ۱۸، ۲۴ و ۲۸ تنها با مرور منابع می‌توانند به آنها دست یابند. در برنامه‌های ۹، ۱۷، ۲۰، ۲۵، ۲۶ و ۳۴ تنها با جست‌وجو می‌توان به منابع آرشو شده دست یافت. شایان ذکر است برنامه ۲۶ دسترسی به مجموعه‌های گزینشگری خود را هم از طریق جست‌وجو و هم مرور امکان‌پذیر می‌کند. همچنین برنامه‌های ۱، ۴، ۵، ۷، ۱۰، ۱۱، ۱۲، ۱۳، ۱۴، ۱۵، ۲۱، ۲۲، ۲۷، ۲۹، ۳۰، ۳۱، ۳۲، ۳۳، ۳۵، ۳۶، ۳۷، ۳۸، ۳۹، ۴۳، ۴۴ و ۴۵ هم مرور و هم جست‌وجو را در اختیار کاربران قرار می‌دهند. شایان توجه است که تنها برنامه ۴۲ اعلام کرده بود که برای در دسترس قرار دادن منابع حفاظت شده، هیچ تسهیلاتی را به کار نگرفته است.

انواع سبک‌های جست‌وجو و مرور

بررسی برنامه‌ها از نظر نوع خدمات جست‌وجو روشن کرد که انواع مختلفی از شیوه‌های جست‌وجوی فیلدی و آزاد، شامل جست‌وجوی نشانی منبع، جست‌وجوی موضوعی، جست‌وجوی تاریخ، جست‌وجوی متن، جست‌وجوی قالب، جست‌وجوی عنوان وب‌سایت، جست‌وجو در مجموعه موضوعی و مانند آن فراهم شده است.

جدول ۴ فراوانی برنامه‌ها را از جنبه سبک‌های جست‌وجو به تصویر می‌کشد. همان‌گونه که مشاهده می‌شود جست‌وجوی نشانی و جست‌وجوی آزاد از بیشترین فراوانی برخوردارند. پس از آن، جست‌وجو بر اساس تاریخ، موضوع و جست‌وجوی در مجموعه موضوعی قرار دارند. جست‌وجو بر اساس قالب و نوع فایل کمترین تعداد را دارند.

بررسی برنامه‌ها به لحاظ سبک جست‌وجو نشان می‌دهد که ۶ برنامه تنها دو نوع سبک جست‌وجو را فراهم می‌کنند که عبارتند از ۵، ۱۰، ۱۵، ۲۴، ۲۸، ۴۳. تعداد ۶ برنامه مشتمل بر برنامه‌های ۱۳، ۱۸، ۲۰، ۲۳، ۲۷ و ۳۷ سه نوع سبک جست‌وجو را ارائه می‌کنند و ۱۳ برنامه بیش از سه سبک جست‌وجو را در اختیار قرار می‌دهند، این برنامه‌ها عبارتند از ۱، ۱۱، ۱۲، ۱۴، ۲۲، ۲۹، ۳۰، ۳۱، ۳۲، ۳۳، ۳۶، ۳۸ و ۳۹.

تنها برنامه‌هایی که فقط یک نوع سبک جست‌وجو را در اختیار قرار می‌دهند، برنامه‌های ۴، ۲۵ و ۲۶ (جست‌وجوی نشانی) و برنامه‌های ۱۷، ۲۱ و ۳۵ (جست‌وجوی آزاد) هستند. همان‌گونه که مشاهده می‌شود، تمایل برنامه‌ها به ارائه سبک‌های هر چه متنوع‌تری از جست‌وجو است.

جدول ۴. فراوانی برنامه‌ها به لحاظ سبک‌های جست‌وجو^{۳۷}

سبک	فراوانی برنامه‌ها
نشانی	۲۸
جست‌وجوی آزاد	۲۶
تاریخ	۱۴
عنوان وب‌سایت	۱۲
موضوعی	۱۲
جست‌وجو در مجموعه موضوعی	۱۱
قالب	۷
نوع فایل	۲

خدمات مرور به هر دو شکل موضوعی و الفبایی در برنامه‌های ۱۵، ۲۷ و ۲۸ به کار گرفته شده است. برنامه‌هایی که تنها یک نوع مرور را فراهم می‌کردند برنامه ۴ (مرور موضوعی) و برنامه ۵ و ۷ (مرور الفبایی) بودند. در برنامه ۴۶ تنها با کلیک بر پیوندهای فراهم‌آمده توسط ناشران یا نویسندگان مورد استناد، می‌توان به منابع آرشیو شده دست یافت.

خدمات نمایه‌سازی، فهرست‌نویسی و رده‌بندی

متأسفانه ضعف مستندات و پاسخ‌ها در این مورد نیز سبب انعکاس تصویری نه چندان دقیق از وضعیت برنامه‌های برتر می‌شود. اندک اطلاعات موجود، نشان داد که ۲۶ برنامه یک یا چند مورد از خدمات نمایه‌سازی، فهرست‌نویسی و رده‌بندی را ارائه می‌دهند. از این میان، برنامه‌های ۵، ۸، ۱۲، ۱۳، ۱۴، ۱۶، ۲۸، ۳۱، ۳۵، ۳۸ و ۴۱ به نمایه‌سازی و برنامه‌های ۴، ۱۱، ۱۸، ۱۹، ۲۹ و ۴۵ به نمایه‌سازی و فهرست‌نویسی و برنامه‌های ۱۷، ۲۳، ۲۵ و ۳۴ به فهرست‌نویسی و برنامه‌های ۳۶ و ۴۴ به فهرست‌نویسی و رده‌بندی و برنامه‌های ۲۲، ۳۰، ۳۲ به هر سه مورد می‌پردازند.

نرم‌افزارها

آرشیوهای حفاظت به‌منظور انجام دادن فرآیندها و عملیات‌های خود به طیف وسیعی از

نرم‌افزارهای دروگری، نمایه‌سازی و دسترسی نیاز دارند. در این بخش وضعیت برنامه‌ها به لحاظ نرم‌افزارهای مورد استفاده بررسی می‌شود. جدول شماره ۶ فهرستی از نرم‌افزارهای مورد استفاده در برنامه‌های برتر را نشان می‌دهد.

نرم‌افزارهای دروگری

همان‌گونه که جدول ۵ نشان می‌دهد، شمار زیادی از نرم‌افزارهای دروگری در برنامه‌های بررسی شده استفاده می‌شوند. برنامه‌های ۱۲، ۱۳، ۱۴، ۱۶، ۲۲، ۲۳، ۲۴، ۲۶، ۲۹، ۳۱، ۳۲، ۳۳، ۳۵، ۳۷، ۴۱، ۴۳، ۴۴ و ۴۵ از نرم‌افزار هریتریکس، برنامه‌های ۳۳ و ۳۸ از نرم‌افزار هریتریکس و دبلیو. سی. تی.^{۳۸} استفاده می‌کنند. برنامه‌های ۱۸ و ۳۶ از نرم‌افزار دبلیو. سی. تی بهره می‌برند. برنامه ۵ از هریتریکس^{۳۹} و اچ. یو. ال - وکس^{۴۰} استفاده می‌کند. برنامه ۲۸ از اچ. تی. ترک و هریتریکس استفاده می‌کند. برنامه ۳۰ از اچ. تی. ترک و پانداز استفاده می‌کند. برنامه ۲، ۱۹ از دروگر «کمباین»^{۴۱} استفاده می‌کنند و برنامه ۶، از خزشگر «زیلم» بهره می‌برد. شایان ذکر است، بعضی از برنامه‌ها در طول زمان نرم‌افزار مورد استفاده خود را تغییر داده‌اند، مانند برنامه آرشو پیوسته اتریش که ابتدا از دروگر ندلیب استفاده کرده و بعد از دروگر ترکیبی بهره برده است. برنامه دروگری گزینشی نیوزلند قبل از سال ۲۰۰۷ از دروگر اچ. تی. ترک استفاده می‌کرده است، ولی بعد از آن از نرم‌افزار دبلیو. سی. تی استفاده می‌کند.

جدول ۵. نرم‌افزارهای مورد استفاده در برنامه‌های برتر

نمایه‌سازی	دسترسی	دروگری
		اچ. تی. تی. رک
ناچ وکس	تروو ^{۴۴}	اچ. یو. ال - وکس
	سی. ال. آی. ا ^{۴۵}	پانداز ^{۴۲}
لوسنس ^{۴۸}	فایرفاکس	دبلیو. سی. تی
سولر ^{۴۹}	ناچ وکس ^{۴۶}	دروگر کامباین
هادوپ ^{۵۰}	ورا ^{۴۷}	دروگر ندلیب ^{۴۳}
وی بک ماشین	وی بک ماشین	زیلم
		هریتریکس

این نرم‌افزارها از سه نوع تجاری، درون‌سازمانی و آزاد هستند. برای نمونه برنامه ۳۴، نرم‌افزار خود را به صورت درون‌سازمانی به وجود آورده است، ولی از آنجا که گزینه تجاری را برگزیده است، احتمال می‌رود آن را به شکل تجاری نیز عرضه کرده باشد. برنامه ۸ نرم‌افزار دروگری خود

را با نام «سایت سیرباکس»^{۵۱} و برنامه ۲۷ که نام نرم‌افزار خود را مشخصاً اعلام نکرده است، به صورت درون‌سازمانی به وجود آورده‌اند و نرم‌افزارهای دروگری هریتریکس، اچ. تی. تی. رک، اچ. یو. ال - وکس، پانداز، دروگر کامپاین و دبلیو. سی. تی که در برنامه‌های ۲، ۵، ۱۲، ۱۳، ۱۴، ۱۸، ۲۲، ۲۳، ۲۴، ۲۶، ۲۸، ۲۹، ۳۰، ۳۱، ۳۲، ۳۳، ۳۵، ۳۶، ۳۷، ۳۸، ۴۳، ۴۴، ۴۵ استفاده می‌شوند، به صورت دسترسی آزاد است. همان‌گونه که ملاحظه می‌شود، گرایش برنامه‌ها به استفاده از برنامه‌های آزاد بسیار بیش از طراحی درون‌سازمانی یا نرم‌افزارهای تجاری است.

نرم‌افزارهای دسترسی

به منظور فراهم‌آوری دسترسی به منابع آرشیو شده، برنامه‌های ۵، ۱۴، ۱۹، ۲۳، ۲۴، ۲۶ و ۲۸ از نرم‌افزار وی‌بک ماشین، برنامه ۱۶، ۲۹، ۳۳، ۴۳ و ۴۵ از نرم‌افزار وی‌بک ماشین و ورا، برنامه‌های ۱۳ و ۳۱ از نرم‌افزار وی‌بک ماشین و ناچ وکس، برنامه ۲۲ از نرم‌افزار وی‌بک ماشین و سی. ال. آی. استفاده می‌کنند. برنامه ۴ از ناچ وکس و ورا استفاده می‌کنند. بعضی از برنامه‌ها مانند برنامه ۱۲، ۳۰ و ۳۴ از نرم‌افزارهای درون‌سازمانی استفاده می‌کنند، در عین حال، به نام آنها اشاره نکرده‌اند. برنامه ۳۸ نیز از نرم‌افزار درون‌سازمانی به نام تروو استفاده می‌کند، برنامه ۴۲، از هریتریکس و دبلیو. سی. تی و برنامه ۱۸ از وی‌بک ماشین و فایرفاکس (که یک ابزار عمومی جست‌وجوی اینترنت است) بهره می‌برند.

در مورد نرم‌افزارهای دسترسی نیز تمایل به کاربرد نرم‌افزارهای آزاد بیشتر است. نرم‌افزارهای دسترسی در برنامه‌های ۵، ۱۳، ۱۴، ۱۸، ۱۹، ۲۳، ۲۴، ۲۸، ۲۹، ۳۱، ۳۳، ۳۶، ۴۲، ۴۳ و ۴۵ به صورت آزاد در اختیار هستند. اما برنامه‌های ۲۲ و ۳۴، از نرم‌افزارهای تجاری استفاده می‌کنند، نرم‌افزارهایی که برنامه‌های ۱۲، ۳۰، ۳۲، ۳۵، ۳۸ و ۴۴، برای در دسترس‌پذیر کردن منابع آرشیو شده خود استفاده می‌کنند، از نوع درون‌سازمانی هستند.

نرم‌افزارهای نمایه‌سازی

برنامه‌های ۴، ۵، ۱۴، ۲۹ و ۴۳ از نرم‌افزار ناچ وکس، برنامه‌های ۱۲، ۲۸ و ۳۲ از نرم‌افزار لوسنس، برای نمایه‌سازی استفاده می‌کنند. برنامه‌های ۱۳ و ۳۱ هم از ناچ وکس و هم وی‌بک ماشین، برنامه ۲۲ از ناچ وکس و از سولر - لوسنس، برنامه ۱۸، برنامه ۲۴ از وی‌بک ماشین، برنامه ۳۳ و ۴۴ از سولر، برنامه ۳۸ از «سولر» و «هادوپ» استفاده می‌کنند.

در این مورد نیز همانند دیگر انواع نرم‌افزاری، گرایش کلی برنامه‌هایی که اطلاعاتی را در این زمینه فراهم کرده‌اند، به کاربرد نرم‌افزارهای آزاد است، به نحوی که شمار زیادی از برنامه‌ها شامل

۵، ۱۲، ۱۳، ۱۴، ۱۸، ۲۲، ۲۳، ۲۴، ۲۹، ۳۰، ۳۱، ۳۲، ۳۳، ۳۸، ۴۳ و ۴۴ از نوع دسترسی آزاد استفاده می‌کنند. نرم‌افزار مورد استفاده برنامه‌های ۲۸، ۳۴ و ۳۷ از نوع تجاری و نرم‌افزارهای نمایه‌سازی برنامه ۳۵ از نوع درون‌سازمانی است.

مدل حفاظتی

به کارگیری الگوی استاندارد ضمن تضمین توسعه همه‌جانبه برنامه‌ها و پیش‌بینی نیازهای کوتاه‌مدت و بلندمدت آن، زمینه‌ساز همکاری بین آرشیوها و در نهایت ایجاد سامانه آرشیو مجازی جهانی است. اطلاعات گردآمده نشان از آن دارد که دو مدل ا.ا. آی. اس و تی. دی. آر در میان برنامه‌های مورد بررسی کاربرد دارد. برنامه‌های ۴، ۸، ۱۱، ۱۸، ۳۰، ۳۲، ۳۳، ۳۵، ۳۶، ۳۸، ۴۳ و ۴۵ از مدل ا.ا. آی. اس استفاده می‌کنند و برنامه ۴۱ هم از مدل ا.ا. آی. اس و هم از تی. دی. آر الگوبرداری می‌کند. برنامه ۱۴، اعلام کرده است که شاید در آینده از مدل تی. دی. آر استفاده کند و هنوز تصمیم قطعی برای آن در نظر نگرفته است. برنامه ۴ و ۴۳ مشخصاً اعلام کرده‌اند که از پروتکل دروی فراداده‌های منابع آزاد (ا.ا. آی. - پی. ام. اچ) استفاده می‌کند.

بحث و نتیجه‌گیری

در این مقاله به بررسی راهکارهای برنامه‌های برتر حفاظت وب در مواجهه با چالش‌های فنی (که بیشترین و بنیادی‌ترین مشکلات را در حفاظت پدید می‌آورند) پرداخته شد. یافته‌های پژوهش نشان داد که برنامه‌های مورد بررسی در مقابل این چالش‌ها، راهکارهای گوناگونی را در پیش گرفته‌اند. اهمیت محتوا، بافتار، ساختار و نمای دیداری منابع وبی برای برنامه‌های مختلف، متفاوت است. ولی برای بیشتر برنامه‌ها، از هر سه ویژگی محتوا، بافتار و ساختار منابع حفاظت می‌کنند. برنامه‌های حفاظت منابع وبی از روش‌های حفاظت از جریان بیت‌ها، تقلیدگری، مهاجرت، تازه‌سازی و قرینه‌سازی استفاده می‌کنند، ولی بیشتر برنامه‌ها، روش‌های مهاجرت و تقلیدگری را مفید دانسته‌اند.

با توجه به هدف حفاظت که نگهداری همه نسخه‌های پیشینه‌های آثار تولید شده است، تمام مسئولان بیان کرده‌اند که از نسخه‌های قدیمی حفاظت می‌کنند. هیچ‌کدام از برنامه‌ها در مستندات خود به این موضوع اشاره‌ای ندارند، به نظر می‌رسد این امر تنها ناشی از بدهت آن بوده است و نه عدم توجه به این معیار مهم و زیربنایی.

برای دسترسی به منابع آرشیوشده، تولید فراداده ضروری است. بسیاری از برنامه‌ها، به تولید فراداده می‌پردازند که شاید فراداده‌های استاندارد یا خودساخته باشند. در بیشتر موارد از

فرا داده‌های استاندارد استفاده شده است. انواعی از فراداده‌ها از جمله فراداده‌های حفاظتی، توصیفی، فنی و ساختاری وجود دارد، اما هیچ کدام از برنامه‌ها اطلاعاتی در مورد نوع فراداده در اختیار قرار نمی‌داند. شاید دلیلش این است که جزییات فنی را در مستندات ذکر نمی‌کنند یا مسئولان تماس از این امر آگاهی نداشتند. فراداده‌ها شاید به صورت دستی یا خودکار ایجاد شوند. با بررسی برنامه‌ها مشخص شد که تقریباً تعداد برنامه‌هایی که به صورت دستی و خودکار به تولید فراداده می‌پردازند، یکسان است. گرچه ایجاد فراداده به صورت دستی وقت گیر و هزینه بر است.

برای دسترسی به منابع آرشیو شده، علاوه بر فراداده به خدماتی نظیر نمایه‌سازی، فهرست‌نویسی، و رده‌بندی نیاز است. با توجه به دشواری فهرست‌نویسی و رده‌بندی، شمار بیشتری از برنامه‌ها تنها به نمایه‌سازی می‌پردازند، با این حال شمار اندکی نیز فهرست‌نویسی و رده‌بندی را اجرا می‌کنند. بیشتر برنامه‌ها، خدمات دسترسی به منابع آرشیو شده را در اختیار کاربران قرار می‌دهند، خواه از طریق مرور و خواه از طریق جست‌وجو در آرشیو. این موضوع از توجه به اهمیت دسترسی در کنار حفاظت نشان دارد.

تسهیلات جست‌وجو در برنامه‌های مختلف به سبک‌های جست‌وجوی نشانی، جست‌وجوی موضوعی، جست‌وجوی تاریخ، جست‌وجوی آزاد، جست‌وجوی فرمت، جست‌وجوی عنوان وب‌سایت، جست‌وجو در مجموعه‌های موضوعی و مرور به صورت موضوعی و فهرست‌الفبایی وب‌سایت‌هاست. جست‌وجوی نشانی و آزاد از بیشترین فراوانی جست‌وجو برخوردارند. پس از آن، جست‌وجو بر اساس تاریخ، موضوع و مجموعه موضوعی قرار دارند. جست‌وجو بر اساس قالب و نوع فایل کمترین فراوانی را در بر می‌گیرند. بررسی برنامه‌ها از نظر تنوع سبک‌های جست‌وجو، از گرایش گسترده به ارائه سبک‌های هر چه متنوع‌تر جست‌وجو نشان دارد. این امر نشانگر آن است که برنامه‌ها نه تنها به گردآوری و حفاظت، بلکه به بهره‌وری منابع نیز به‌طور جدی می‌اندیشند.

برنامه‌ها برای دروگری منابع وبی از نرم‌افزارهای خاصی استفاده می‌کنند. نرم‌افزاری که در بیشتر برنامه‌ها استفاده شده، هریتریکس است که از نوع دسترسی آزاد محسوب می‌شود. بیشتر برنامه‌ها از نرم‌افزار دسترسی آزاد برای دروگری استفاده می‌کنند. بیشتر برنامه‌ها از نرم‌افزار وی‌بک ماشین برای در دسترس قرار دادن منابع آرشیو شده استفاده می‌کنند. این نرم‌افزار از نوع دسترسی آزاد است. در کل بیشتر برنامه‌ها از نرم‌افزار دسترسی آزاد استفاده می‌کنند. نرم‌افزارهایی که بیشتر برای نمایه‌سازی به کار می‌روند، نرم‌افزار ناچ و کس و لوسنس هستند. این دو نوع نرم‌افزار از نوع دسترسی آزاد محسوب می‌شوند. بیشتر برنامه‌ها از نرم‌افزار دسترسی آزاد برای نمایه‌سازی

منابع استفاده می‌کنند. بیشتر برنامه‌ها از مدل ا. آی. آی. اس استفاده کرده‌اند، زیرا این مدل قابلیت انعطاف دارد و هر برنامه‌ای از آن استفاده می‌کند. به‌طور کلی، برنامه‌های برتر حفاظت وب در رویارویی با چالش‌های فنی وب، از خطوط کلی که در جدول ۶ نشان داده شده، پیروی می‌کنند:

جدول ۶. راهکارهای برنامه‌های برتر حفاظت وب در مقابله با چالش‌های فنی وب

چالش	راهکار
اجزای حفاظت‌شده	محتوا، بافتار و ساختار منابع
راهبرد حفاظت	روش مهاجرت و تقلیدگری
نوع فراداده‌ها	فراداده‌های استاندارد
روش ایجاد فراداده	دستی و خودکار
خدمات ارزش افزوده	نمایه‌سازی، فهرست‌نویسی، و رده‌بندی
نوع خدمات	خدمات دسترسی به منابع حفاظت‌شده
تسهیلات جست‌وجو	جست‌وجوی آزاد، جست‌وجوی نشانی، تاریخ و موضوع
نرم‌افزار دروگری	نرم‌افزار آزاد هریتريکس
نرم‌افزار دسترسی	نرم‌افزار آزاد وی‌بک ماشین
نرم‌افزار نمایه‌سازی	نرم‌افزارهای آزاد ناچ‌وکس و لوسنس
مدل حفاظتی	مدل ا. آی. آی. اس

چنانکه از نتایج پژوهش بر می‌آید حفاظت وب از بعد فنی فعالیتی بسیار پیچیده، گسترده و پویاست و پیاده‌سازی آن در عمل به دانش تخصصی و فنی بسیار گسترده و امکانات پیشرفته نیاز دارد. در ایران، تاکنون به‌طور جدی به مسئله حفاظت بلندمدت منابع سایبری (که رشد فزاینده‌ای داشته‌اند) پرداخته نشده است. تجربه «پروژه حرم» و کنسرسیوم محتوای ملی، از معدود تلاش‌های جدی است که از سوی کتابخانه ملی و کتابخانه‌های مختلف سراسر کشور با هدف حفاظت از میراث معنوی به‌شکل دیجیتالی آغاز شده است. با آنکه این تلاش‌ها بسیار ارزشمند و در خور تقدیر است، هیچ‌یک به حفاظت از وب عمومی یا وب «نمایان» توجهی نشان نداده است. از این‌رو، ضروری خواهد بود که نهادهای ملی متولی حفاظت از میراث معنوی یا نهادهای علاقه‌مند به حفاظت، هر چه سریع‌تر حفاظت از منابع وبی را در دستور کار قرار دهند.

گرچه ایجاد نظام پایداری برای حفظ میراث دیجیتالی ابعاد وسیع و پیچیده‌ای دارد و در برخی مراحل مستلزم سرمایه‌گذاری کلان است، احتمالاً هزینه‌های نادیده گرفتن آن، به‌ویژه برای نسل‌های آینده سنگین‌تر خواهد بود. بنابراین، مهم‌ترین اولویت، ایجاد درک مشترک از اهمیت

مواد وبی و گردآوری و حفاظت آنهاست. اعضای جامعه علمی و دانشگاهی به‌ویژه کتابخانه‌ها باید به درک مشترک و عمیقی از فرصت‌ها و تهدیدهای موجود دست یابند تا بتوانند از فرصت‌های موجود به بهترین وجه بهره ببرند و تهدیدها را هم به فرصت تبدیل کنند در این زمینه، با توجه به نبود الگوی استاندارد، ضروری است پیش از هر گونه اقدام عملی، با الهام از خط‌مشی‌ها و الگوهای برگرفته از برنامه‌های برتر حفاظت وب جهانی، راه‌تدوین طرحی عملیاتی برای برنامه ملی نگهداشت میراث دیجیتال منتشر شده در فضای سایبری ایران هموار شود.

پی‌نوشت‌ها

1. Hackers
2. Hedstrom
3. Ashley
4. Thompson
5. Parekh
6. Emmanuelle and Gildas
7. Smith
8. Hodge
9. National Library of Australia
10. Harvesting

۱۱. به‌منظور شناسایی این چالش‌ها، بیش از صد منبع پیرامون حفاظت منابع وبی و دیجیتالی بررسی شد و مهم‌ترین چالش‌هایی که در بخش اعظمی از این متون مورد تأکید قرار گرفته بود، استخراج و بررسی شد (رشیدی، ۱۳۹۰).

12. Kuny
13. Christensen
14. Barve
15. Caplan
16. TIPR (Towards Interoperable preservation repository)
17. Mellor, Wheatley, Sergeant
18. Kelly
19. Constantopoulos1&Doerr2& Petraki
20. Mikeal, Green, Maslov, Phillips, Leggett

۲۱. در روش مهاجرت منابع به‌طور دوره‌ای از یک محیط نرم‌افزاری و سخت‌افزاری به محیط دیگر، از پیکربندی یک سخت‌افزار یا نرم‌افزار به دیگری و از یک نسل فناوری رایانه‌ای به نسل‌های بعدی منتقل می‌شوند.

۲۲. راهبرد تقلید یا نمونه‌سازی، حفاظت از وضعیت ظاهری و ساختاری دقیق منابع اینترنتی را مد نظر قرار داد. این راهبرد، تنها روش معتبر بازآفرینی قابلیت‌های کارکردی مدارک و نمای دیداری آنهاست.

23. OAIS: Open Archival Information System Reference Model

24. Wheatley

25. CAMiLEON

26. Leeds

27. BBC Domesday

28. Maniatis , Intel Research, Roussopoulos, Harvard University, Giuli, Stanford University, Rosenthal, Stanford University Libraries, Baker

29. LOCKSS: Lots of Copies Keep Stuff Safe

30. Zabicka and Matejky

۳۱. روش تازه‌سازی از جریان بیتی اشیا حفاظت می‌کند و هنگامی که رسانه‌ها یا سامانه‌های اساسی منسوخ یا ناخوانا می‌شوند، ضروری است.

۳۲. در روش‌های نسخه‌برداری یا قرینه‌سازی، اطلاعات دروشده به صورت دسته‌ای کپی یا تصویر قرینه آن تهیه می‌شود.

33. METS

34. MODS

35. NLNZ

36. ACIS

۳۷. به دلیل همپوشانی سبک‌ها در برنامه‌ها، شمار کل، از شمار کل برنامه‌ها بیشتر است.

38. WCT: Web Curator Tool

39. Heritrix

40. Hul-wax

41. Combine harvester

42. Pandaz

43. Nedlib

44. Trove (NLA discovery service)

45. CLIO

46. NutchWAX

47. wera

48. Lucene

49. SOLR

50. HADOOP

51. Citeseerbox

منابع

۱. رشیدی، کیانوش (۱۳۹۰). بررسی برنامه‌های حفاظت دیجیتال مواد وبی و راهکارهای آنها در مواجهه با چالش‌های مربوطه. پایان‌نامه کارشناسی ارشد. دانشگاه شیراز، دانشکده علوم تربیتی و روانشناسی.
2. Ashley, K. , Davis, R. , Guy, M. , Kelly, B., Pinsent, Ed. & Farrell, S. (2010). A Guide to Web Preservation. Retrieved October 20, 2011, from <http://jiscpowr.jiscinvolve.org/wp/files/2010/06/Guide-2010-final.pdf>.
3. Barve, S. (2007). File Formats in Digital Preservation. Retrieved October 20, 2011, from http://drtc.isibang.ac.in:8080/xmlui/bitstream/handle/1849/312/027_p48_sunita-barve_ana-formatted.pdf?sequence=1.
4. Caplan, P. (2008). Repository to Repository Transfer of Enriched Archival Information Packages. *D-Lib Magazine*, 14(11/12).
5. Christensen , NH. (2004). Towards format repositories for web archives. Retrieved October 20, 2011, from <http://iwaw.europarchive.org/04/Christensen.pdf>.
6. Constantopoulos, P., Doerr, M. & Petraki , M. (2005) . Reliability modelling for long term digital preservation. Retrieved October 20, 2011, from <http://www.ics.forth.gr/isl/publications/paperlink/Reliability%20modelling.pdf>.
7. Day, M. (1998). Metadata for Digital Preservation. Retrieved October 20, 2011, from <http://www.ariadne.ac.uk/issue22/metadata/>.
8. Day, M. (2006). Long_Term Preservation of Web Content. Retrieved October 20, 2011, from www.ukoln.ac.uk/preservation/publications/2006/web.../md-final-draft.pdf.
9. Hedstrom, M. (1998). Digital Preservation: A Time Bomb for Digital Libraries. *Computers and the Humanities*, 31: 189–202.
10. Hodge, G. & Frangakis, E. (2004). Digital Preservation Access to Scientific Information: The State of the Practice. Retrieved October 20, 2011, from <http://www.dtic.mil/cgiin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf>
11. Kelly, B. (2002) . Approaches to the preservation of web sites. Retrieved October 20, 2011, from <http://www.ukoln.ac.uk/web-focus/events/conferences/online-information-2002/paper.pdf>.
12. Kuny, T. (1997). A Digital Dark Ages? Challenges in the Preservation of Electronic Information. In: IFLA Council and General Conference. Retrieved October 20, 2011, from <http://archive.ifla.org/IV/ifla63/63kuny1.pdf>.
13. Mellor, P., Wheatley , P., Sergeant, D. (2002). Migration on Request, a Practical Technique for Preservation. In: European Conference on Digital Libraries - ECDL
14. Mikeal, A., Green, C., Maslov, A., Phillips, S. & Leggett , J. (2007). Preserving the Scholarly Side of the Web. Retrieved October 20, 2011, from <http://repository.tamu.edu/bitstream/handle/1969.1/5680/LA%20WEB%202006%20Paper.pdf?sequence=1>.

15. National Library of Australia (2003). Guidelines for the Preservation Of Digital Heritage. Retrieved October 20, 2011, from <http://unesdoc.unesco.org/images/0013/001300/130071e.pdf>.
16. OCLC/RLG Working Group on Preservation Metadata (2001). Preservation Metadata for Digital Objects: A Review of the State of the Art. Retrieved October 20, 2011, from http://www.oclc.org/research/activities/past/orprojects/pmwg/presmeta_wp.pdf
17. Ross, S. (2003). Challenges to Digital Preservation and Building Digital Libraries. Available at: <http://archive.ifla.org/IV/ifla69/papers/209e-Ross.pdf>
18. Smith, A. (2003). New-model scholarship: how will it survive? Washington, D.C.: Council on Library and Information Resources. Retrieved October 20, 2011, from <http://www.clir.org/pubs/abstract/pub114abst.html>
19. Thompson, D. (2008). Digital Curation Manual. Retrieved October 20, 2011, from <http://www.era.lib.ed.ac.uk/bitstream/1842/3345/1/McHugh%20open-source.pdf>
20. Wheatley, P. (2004) . Digital Preservation and BBC Domesday. Retrieved October 20, 2011, from <http://cool.conservation-us.org/coolaic/sg/emg/library/pdf/wheatley/Wheatley-EMG2004.pdf>.
21. Zabicka , P. , & Matejky, L.(2007). Czech Web Archive Analysis. *New Review of Hypermedia and Multimedia*, 13(1): 27-37.
22. Ashley, K. , Davis, R. , Guy, M. , Kelly, B., Pinsent, Ed. & Farrell, S. (2010). A Guide to Web Preservation. Retrieved October 20, 2011, from <http://jiscpowr.jiscinvolve.org/wp/files/2010/06/Guide-2010-final.pdf>.
23. Barve, S. (2007). File Formats in Digital Preservation. Retrieved October 20, 2011, from http://drc.isibang.ac.in:8080/xmlui/bitstream/handle/1849/312/027_p48_sunita-barve_ana-formatted.pdf?sequence=1.
24. Caplan, P. (2008). Repository to Repository Transfer of Enriched Archival Information Packages. *D-Lib Magazine*, 14(11/12).
25. Christensen , NH. (2004).Towards format repositories for web archives. Retrieved October 20, 2011, from <http://iwaw.europarchive.org/04/Christensen.pdf>.
26. Constantopoulos, P., Doerr, M. & Petraki , M. (2005) . Reliability modelling for long term digital preservation. Retrieved October 20, 2011, from <http://www.ics.forth.gr/isl/publications/paperlink/Reliability%20modelling.pdf>.
27. Day, M. (1998). Metadata for Digital Preservation. Retrieved October 20, 2011, from <http://www.ariadne.ac.uk/issue22/metadata/>.
28. Day, M. (2006). Long_Term Preservation of Web Content. Retrieved October 20, 2011, from www.ukoln.ac.uk/preservation/publications/2006/web.../md-final-draft.pdf.
29. Emmanuelle, Bermes and Gildas, Illien (2009). “Metrics and Strategies for Web Heritage Management and Preservation”. In: 75TH IFLA General Conference and Council(23-27 August 2009, Milan, Italy).
30. Hedstrom, M. (1998). Digital Preservation: A Time Bomb for Digital Libraries. *Computers and the Humanities*, 31: 189–202.

31. Hodge, G. & Frangakis, E. (2004). Digital Preservation Access to Scientific Information: The State of the Practice. Retrieved October 20, 2011, from <http://www.dtic.mil/cgiin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf>
32. Kelly, B. (2002) . Approaches to the preservation of web sites. Retrieved October 20, 2011, from <http://www.ukoln.ac.uk/web-focus/events/conferences/online-information-2002/paper.pdf>.
33. Kuny, T. (1997). A Digital Dark Ages? Challenges in the Preservation of Electronic Information. In: IFLA Council and General Conference. Retrieved October 20, 2011, from <http://archive.ifla.org/IV/ifla63/63kuny1.pdf>.
34. Mellor, P., Wheatley, P., Sergeant, D. (2002). Migration on Request, a Practical Technique for Preservation. In: European Conference on Digital Libraries - ECDL.
35. Mikeal, A., Green, C., Maslov, A., Phillips, S. & Leggett, J. (2007). Preserving the Scholarly Side of the Web. Retrieved October 20, 2011, from <http://repository.tamu.edu/bitstream/handle/1969.1/5680/LA%20WEB%202006%20Paper.pdf?sequence=1>.
36. National Library of Australia (2003). Guidelines for the Preservation Of Digital Heritage. Retrieved October 20, 2011, from <http://unesdoc.unesco.org/images/0013/001300/130071e.pdf>.
37. OCLC/RLG Working Group on Preservation Metadata (2001). Preservation Metadata for Digital Objects: A Review of the State of the Art. Retrieved October 20, 2011, from http://www.oclc.org/research/activities/past/orprojects/pmwg/presmeta_wp.pdf
38. Parekh, r. and Parekh, P. (2009). "Planning for Digital Preservation of Special Collections in Gujarat University Library". Available at: www.inflibnet.ac.in/caliber2009/CaliberPDF/11.pdf.
39. Rauber, Andreas (2008). "IT Research Challenges in Digital Preservation". Available at: http://rcdl.ru/doc/2009/103_110_Invited-1.pdf.
40. Ross, S. (2003). Challenges to Digital Preservation and Building Digital Libraries. Available at: <http://archive.ifla.org/IV/ifla69/papers/209e-Ross.pdf>
41. Smith, A. (2003). New-model scholarship: how will it survive? Washington, D.C.: Council on Library and Information Resources. Retrieved October 20, 2011, from <http://www.clir.org/pubs/abstract/pub114abst.html>
42. Thompson, D. (2008). Digital Curation Manual. Retrieved October 20, 2011, from <http://www.era.lib.ed.ac.uk/bitstream/1842/3345/1/McHugh%20open-source.pdf>
43. Wheatley, P. (2004) . Digital Preservation and BBC Domesday. Retrieved October 20, 2011, from <http://cool.conservation-us.org/coolaic/sg/emg/library/pdf/wheatley/Wheatley-EMG2004.pdf>.
44. Zabicka, P. , & Matejky, L.(2007). Czech Web Archive Analysis. *New Review of Hypermedia and Multimedia*, 13(1): 27-37.