

**Az „Elektronikus levéltár” projekt  
keretében a hosszú távú levéltári  
megőrzéshez szükséges szabályozási  
feltételek kidolgozása**

**A papír alapú dokumentumok  
megbízható elektronikus  
másolatával szemben  
támasztott technikai  
követelmények és gyakorlatban  
alkalmazható eljárásrend  
a Magyar Országos Levéltár részére**



**Változat száma:** 2.0

**Fájlnév:** FLX\_MOLSZM\_Elektronikus\_masolat\_v2.0\_1109  
13.doc

**A dokumentumot készítette:** FLEXUS Consulting Tanácsadó Kft.

**Oldalak száma:** 76

**Mellékletek száma:** 7

**Minőségbiztosítás:**



**Dokumentum jellemzők:**

Projekt megnevezése:	Az „Elektronikus levéltár” projekt keretében a hosszú távú levéltári megőrzéshez szükséges szabályozási feltételek kidolgozása
Dokumentum címe:	A papír alapú dokumentumok megbízható elektronikus másolatával szemben támasztott technikai követelmények és gyakorlatban alkalmazható eljárásrend
Verziószám:	V2.0
Minősítés (állapot):	munkaközi
Vonatkozó időszak:	-
Kiadás dátuma:	2011.09.13.
Készítette:	Dr. Cseh Gergő Bendegúz
Fájlnév:	FLX_MOLSZM_Elektronikus_masolat_v2.0_110913.doc

**Verziók jegyzéke:**

Verzió	Szerző	Dátum	Változtatás rövid leírása
0.1	Lits Bálint	2011.06.06	Termék annotált tartalomjegyzéke
0.2	Lits Bálint	2011. 06. 13.	Tartalomjegyzék elfogadott verziója
1.0	Cseh Gergő B.	2011. 09. 05.	Kiadott termék
2.0	Cseh Gergő B.	2011. 09. 13.	Termék véglegesítése

**Ellenőrizte:**

Ellenőrzést végezte	Ellenőrzés dátuma	Aláírás

**Jóváhagyta:**

Név	Szervezeti egység/Beosztás	Dátum	Aláírás
			-

**Kapják:**

#	Cégnév	Név	Szervezet	Beosztás
1				
2				
3				
4				

# Tartalomjegyzék

<b>1. VEZETŐI ÖSSZEFOGLALÓ</b> .....	<b>6</b>
<b>2. JOGI HELYZET ÉS NÉHÁNY MEGVALÓSULT DIGITALIZÁLÁSI PROJEKT</b> .....	<b>7</b>
2.1 HAZAI JOGI KÖRNYEZET .....	7
2.2 JAVASLATOK A HAZAI JOGI SZABÁLYOZÁS MÓDOSÍTÁSÁRA .....	15
2.3 MAGYARORSZÁGI DIGITALIZÁLÁSI PROGRAMOK .....	17
2.3.1 Digitalizálási program a MÁV Zrt. Központi Irattárában .....	17
2.3.2 Digitalizálás az ÁPV Zrt iratanyagában .....	19
2.3.3 Digitalizálás a Nemzeti Hírközlési Hatóság (NMHH) iratanyagában .....	20
<b>3. NEMZETKÖZI KEZDEMÉNYEZÉSEK</b> .....	<b>20</b>
3.1 A LEVÉLTÁRI LEÍRÁS ÉS ARCHIVÁLÁS NEMZETKÖZI SZABVÁNYAI .....	20
3.2 MOREQ, MOREQ2 ÉS MOREQ2010.....	22
3.3 A METAADATOK NEMZETKÖZI SZABVÁNYA, A DUBLIN CORE .....	24
3.4 PÉLDÁK A NEMZETKÖZI GYAKORLATBÓL.....	25
3.4.1 Az ENSZ digitalizálási szabványa .....	25
3.4.2 Az EU Domec projektje.....	26
3.4.3 A DOMEA-koncepció és a SAGA-ajánlás Németországban.....	28
3.4.4 A brit példa: PRONOM és DROID .....	29
3.4.5 A Svájci Szövetségi Levéltár gyakorlata.....	31
3.4.6 Egyéb nemzetközi példák.....	31
<b>4. AZ ELEKTRONIKUS MÁSOLATKÉSZÍTÉSI FOLYAMATOK ÁLTALÁNOS     ELŐFELTÉTELEI</b> .....	<b>33</b>
<b>5. FORMÁTUM-KÖVETELMÉNYEK</b> .....	<b>34</b>
5.1 A HOSSZÚ TÁVÚ MEGŐRZÉSRE ALKALMAS FÁJLFORMÁTUMOK KIVÁLASZTÁSÁNAK SZEMPONTJAI .....	35
5.2 AZ ELEKTRONIKUS MÁSOLATOK ARCHIVÁLÁSRA ALKALMAS FORMÁTUMAI .....	37
5.2.1 Adobe PDF/A-1.....	37
5.2.2. Tagged Image File Format .....	38
5.2.3 JPEG formátum .....	39
5.2.4 JPEG2000 formátum .....	40
5.2.5 TXT formátum.....	40

<b>6. JOGI KÖVETELMÉNYEK.....</b>	<b>41</b>
6.1 AZ ELEKTRONIKUS MÁSOLATOK HITELESSÉGE .....	41
6.2 A HITELESSÉG BIZTOSÍTÁSÁNAK ESZKÖZEI .....	42
<b>7. TECHNIKAI FELTÉTELEK.....</b>	<b>44</b>
7.1 A MÁSOLATTAL SZEMBEN TÁMASZTOTT KÖVETELMÉNYEK.....	44
7.2 A MÁSOLATKÉSZÍTŐ RENDSZERREL SZEMBEN TÁMASZTOTT KÖVETELMÉNYEK .....	46
<b>8. A DIGITÁLIS MÁSOLATKÉSZÍTÉS FOLYAMATA.....</b>	<b>48</b>
8.1 ÁLTALÁNOS SZEMPONTOK.....	48
8.1.1 A szkennelés időpontja.....	49
8.1.2 A szkennelés helye.....	50
8.1.3 A szkennelés módja.....	50
8.1.4 A digitális másolatkészítés eszközei .....	51
8.1.5 Dokumentációs követelmények .....	51
8.1.6 Személyi feltételek.....	51
8.2 A MÁSOLATKÉSZÍTÉS FOLYAMATA.....	52
8.2.1 A digitalizálandó anyagok kiválasztása .....	52
8.2.2 Az iratok előkészítése a digitális másolásra .....	53
8.2.3 Másolás és indexelés .....	54
8.2.4 A másolatkészítés során rögzítendő metaadatok .....	56
8.2.5. A digitális képfájlok elnevezésével kapcsolatos megoldások.....	57
<b>9. A DIGITÁLIS MÁSOLAT ELKÉSZÜLTE UTÁNI MUNKAFOLYAMATOK.....</b>	<b>58</b>
9.1 AZ ELKÉSZÜLT KÉPFÁJLOK OCR-ES FELDOLGOZÁSA.....	58
9.2 MINŐSÉGI ELLENŐRZÉS .....	59
9.3 A DIGITÁLIS KÉPFÁJLOK ÉS SZÖVEGES ÁLLOMÁNYOK TOVÁBBI KEZELÉSE .....	60
9.4 AZ EREDETI IRATOK TOVÁBBI KEZELÉSE .....	60
<b>FÜGGELÉK .....</b>	<b>61</b>
1. SZ. FÜGGELÉK A MEGBÍZHATÓ ELEKTRONIKUS MÁSOLATTAL SZEMBEN TÁMASZTOTT KÖVETELMÉNYEK .....	61
2. SZ. FÜGGELÉK ÖSSZEFOGLALÓ TÁBLÁZAT A DIGITALIZÁLÁSI FORMÁTUMOK KÖVETELMÉNYEIRŐL .....	70
3. SZ FÜGGELÉK DIGITALIZÁLÁSI ELLENŐRZÉSI LISTA.....	71
4. SZ. FÜGGELÉK A PDF/A-1 FORMÁTUM SPECIFIKÁCIÓJA A PRONOM ALAPJÁN .....	73

---

5. SZ. FÜGGELÉK A TIFF FORMÁTUM SPECIFIKÁCIÓJA A PRONOM ALAPJÁN .....	74
6. SZ. FÜGGELÉK A JPEG FORMÁTUM SPECIFIKÁCIÓJA A PRONOM ALAPJÁN.....	75
7. SZ. FÜGGELÉK: A JPEG2000 FORMÁTUM SPECIFIKÁCIÓJA A PRONOM ALAPJÁN .....	76

## 1. Vezetői összefoglaló

A maradandó értékű levéltári anyagok hosszú távú megőrzése és használhatóságának biztosítása a levéltári szakma elsődleges feladata. A hagyományos, általában papír alapú dokumentumok kezelésének, tárolásának és használatának széles körben elfogadott és alkalmazott elméleti és módszertani gyakorlata alakult ki idehaza és a nemzetközi szakmai téren egyaránt. Ugyanez nem mondható el az elektronikus úton keletkező, vagy hagyományos adathordozóról készült elektronikus másolatok esetében. A szakmai követelmény ezen iratok megőrzése terén ugyanaz – olvashatóság, hitelesség, integritás -, mivel a vonatkozó jogszabályok (elsősorban a Levéltári törvény) adathordozótól függetlenül iratnak tekinti ezen objektumok összességét.

A másolatkészítési eljárások alapelveinek meghatározásánál nem tekinthetünk el attól a problémától, hogy a mai levéltárak megőrizték a hagyományos hiteleshelyek funkcióit abban a tekintetben, hogy az általuk őrzött iratok jogbiztosító erejét, jogi hitelességét is fenn kell tartaniuk. Ráadásul az iratok keletkezésének – illetve jelen esetben az azokról való elektronikus másolat készítésének – időpontjában nem határolható el azon iratok köre, melyek jogbiztosító erejére a későbbiekben igény, szükség keletkezhet. Jellemző példája a jogbiztosító funkció utólagos „keletkezésének” az 1990 utáni kárpótlások időszaka, melynek során a levéltárakban őrzött olyan dokumentumok, illetve az azokról kiadott hiteles másolatok léptek elő elsőrendű bizonyítékká, melyek a hagyományos levéltári iratértékelés szempontjai szerint többségükben maradandó, történeti értékű iratnak sem voltak tekinthetők.

A jelen dokumentum célja a hagyományos, papír alapú iratok megbízható elektronikus másolatának elkészítése során tekintetbe veendő hazai és a releváns nemzetközi gyakorlat áttekintése, a vonatkozó jogszabályi környezet bemutatása, valamint a hazai és nemzetközi szabványoknak megfelelő technikai követelmények és eljárásrend meghatározása.

A hagyományos adathordozón keletkezett iratok elektronikus másolata, digitalizálása a jelen projekt által vizsgált környezetben elvileg két fázisban történhet, illetve történik meg. Egyfelől az iratképző szervek ilyen formán teszik kompatibilissé elektronikus iratkezelési rendszereikkel a beérkező vagy intézményen belül keletkező, esetleg már ott irattározott hagyományos adathordozójú iratokat. Másfelől a levéltárak maguk is folytatnak elektronikus másolatkészítési gyakorlatot iratanyaguk kímélése és szélesebb körű felhasználhatósága, sokoldalú kommunikálhatósága érdekében. Mindkét folyamat azonos elvekre, formátumkövetelményekre és metaadat-struktúrára kell épüljön a hosszú távú megőrzés és használhatóság garantálása érdekében, bár a levéltári feldolgozás során olyan további metaadatok rögzítésére is sor kerülhet – ez minden bizonnyal elengedhetetlen is -, melyekre

az iratképző szervnél végzett másolás során lehetőség és szükség sincs. A jelen dokumentum fókuszpontja az iratképző szervnél történő másolatkészítés kérdésköre, ugyanezen feladatok levéltáron belül történő végzése külön vizsgálat tárgyát képezi.

A dokumentum címében szereplő „megbízható” kifejezés az elektronikus másolatok tekintetében többrétegű követelményt jelent.

Megbízhatónak egyfelől akkor tekinthető a digitális másolat, ha technikai értelemben hosszú távon is értelmezhető, olvasható, megtekinthető az elektronikus verzió.

Másfelől hasonlóan fontos a megbízhatóság szempontjából, hogy a digitális másolat egyértelműen hozzáköthető legyen ahhoz a hagyományos adathordozójú dokumentumhoz, melyről készült.

Nem kevésbé fontos megbízhatósági kritérium a fentiekén túl, hogy a másolatkészítés rendszere garantálja annak egyértelmű megállapíthatóságát, hogy a másolat elkészítése után bármilyen formában megváltoztatták-e annak tartalmát.

A megbízható másolatnak tehát hosszú távon is értelmezhetőnek, az eredetivel képileg és tartalmilag is megfeleltethetőnek és megváltoztathatatlanoknak kell lennie ahhoz, hogy levéltári értelemben hiteles, valamint jogbiztosító erejű maradjon.

## 2. Jogi helyzet és néhány megvalósult digitalizálási projekt

Papír alapú iratokról mind az iratképző szerveknél, mind pedig a levéltárakban folyamatosan és nagy mennyiségben készülnek elektronikus másolatok, egy megvalósuló e-levéltári projekt során azonban ezen folyamatokat minden tekintetben a levéltári szakmai követelmények legfontosabbikának, nevezetesen a maradandó értékű iratok hosszú távú fennmaradásának kell alárendelni. Mindezt a vonatkozó hazai jogszabályi környezet, az érvényes nemzetközi szabványok és ajánlások megszabta kereteken belül, a megvalósult hazai és nemzetközi programok tapasztalatai alapján kell érvényesíteni.

### 2.1 Hazai jogi környezet

A téma hazai jogszabályi környezetének áttekintése alapvetően abból a feltevésből kell kiinduljon, hogy az iratképző szervnél a papír alapú dokumentumról keletkezett elektronikus másolat egyértelműen iratnak, tehát a Levéltári törvény hatálya alá tartozó objektumnak tekintendő. A köziratokról, a közlevéltárakról és a magánlevéltári anyag védelméről szóló **1995. évi LXVI. tv.** (Ltv.) 3.§ c) pontja szerint ugyanis *„irat: valamely szerv működése vagy személy tevékenysége során keletkezett vagy hozzá érkezett, egy egységként kezelendő rögzített információ, adategyüttes, amely megjelenhet papíron, mikrofilmen, mágneses,*

*elektronikus vagy bármilyen más adathordozón; tartalma lehet szöveg, adat, grafikon, hang, kép, mozgóképek vagy bármely más formában lévő információ vagy ezek kombinációja*”. Ennek megfelelően a digitális másolat – a hagyományos, papír alapú másolathoz hasonlóan – iratnak tekintendő, s így a közlevéltár ezen a téren is jogosult és köteles is gyakorolni az Ltv. 13. § e) pontjában, valamint a közlevéltárak és a nyilvános magánlevéltárak tevékenységével összefüggő szakmai követelményekről szóló **10/2002 (IV.13.) NKÖM rendelet** 7-9. §-aiban megfogalmazott, a közfeladatot ellátó szervek, valamint a tartós állami tulajdonú társasági részesedéssel működő gazdasági társaságok iratkezelésének ellenőrzésére vonatkozó feladatait. Sajnálatos módon a fenti jogszabályok nem térnek ki részleteiben az elektronikus dokumentumkezelő rendszerek kialakítása és működtetése során jelentkező speciális levéltári feladatokra, ennek hiányában is értelmezhető ugyanakkor a közlevéltárak ellenőrzési jogköre az elektronikus iratok keletkeztetésének folyamatában.

A maradandó értékű levéltári anyag védelme szempontjából meghatározó fontosságú, bár csak kisebb részben megvalósult az Ltv. 35/A § (2) bekezdésében megfogalmazott lehetőség, mely szerint *„felhatalmazást kap a köziratok kezelésének szakmai irányításáért felelős miniszter, hogy az e-közigazgatásért felelős miniszter, a közigazgatási informatika infrastrukturális megvalósíthatóságának biztosításáért felelős miniszter, valamint a kultúráért felelős miniszter egyetértésével kiadott rendeletben állapítsa meg a közfeladatot ellátó szerveknél alkalmazható iratkezelési szoftverekkel szemben és az elektronikus iratok levéltárba adásával, tárolásával kapcsolatban támasztott követelményeket*”.

Az Ltv-ben szereplő felhatalmazás első fele tekintetében megszületett a szabályozás, ez a **24/2006. (IV. 29.) BM-IHM-NKÖM együttes rendelet** a közfeladatot ellátó szerveknél alkalmazható iratkezelési szoftverekkel szemben támasztott követelményekről. A rendelet az iratkezelési szoftverek működésének minimális követelményeit fogalmazza meg azzal a kifejezett szándékkal, hogy a hagyományos iratkezelési fogalmakat az elektronikus iratkezelési környezetben is értelmezni lehessen. A rendelet konkrét technikai szabályokat, adatformátumokat, metaadat-szerkezetet vagy eljárásmodokat nem határoz meg, ezzel kapcsolatban az előírt formátumokra vonatkozó értelmező rendelkezésben utal a **12/2005 (X.27.) IHM rendelet** kötelezően betartandó előírásaira. Ez utóbbi jogszabály a szakirodalomban *formátumrendeletként* vált ismertté és széles körben hivatkozott dokumentummá, bár 2009 októbere óta már nincs hatályban. A 24/2006 IHM rendelet egyébként máshelyütt (4.§) nem az akkor még érvényben volt formátumrendelethez, hanem egy, a belügyminiszter, az informatikai és hírközlési miniszter, valamint a Miniszterelnöki Hivatal vezető miniszter által kiadandó külön szabályozáshoz köti az elektronikus adatformátumokra, metaadatokra vagy a használható elektronikus irattípusokra vonatkozó kérdéskör rendezését.



A fent említett, mára hatályát veszített formátumrendelet egyébként korántsem teljes körűen szabályozta az elektronikus adatformátumok kérdéskörét, hiszen mindössze a szöveges dokumentumok megjelenítésére alkalmas Microsoft Rich Text Format 1.7, a képi megjelenítésre alkalmas Portable Document Format 1.3, illetve az elektronikus levelek és az elektronikus aláírással kapcsolatos formátumok elfogadhatóságát és a vonatkozó szabványokat nevesítette.

A formátumrendeletet **az elektronikus közszolgáltatás működtetéséről szóló 222/2009. (X. 14.) kormányrendelet** helyezte hatályon kívül anélkül, hogy az egységes, hosszú távú levéltári megőrzésre alkalmas elektronikus iratformátumokra vonatkozó szabályokat életbe léptetett volna. Ehelyett a rendelet 18. §-a a következőképpen fogalmaz:

*(1) Az elektronikus közszolgáltatásban részt vevő informatikai rendszerek tervezése és megvalósítása során biztosítani kell a következő követelmények teljesítését:*

*a) képesnek kell lenniük az egymás közötti együttműködésre, az alrendszerek közötti kommunikáció, azonosítás, adatcsere, adatelérés, alkalmazás-integráció és azok biztonsága terén,*

*b) adatjelentéstani (szemantikai) szempontból egységes alapra kell épülniük, ami lehetővé teszi az átadott adatok, metaadatok közvetlen feldolgozását, az egységes fogalmi modellezést, tranzakció- és eseménykezelést,*

*c) ügyféloldalon a megkövetelt informatikai eszközök lehetséges minimumára kell építeni úgy, hogy azok minden, széles körben elterjedt operációs rendszerből funkcióvesztés nélkül elérhetők legyenek,*

*d) képesnek kell lenniük - megfelelő illesztéssel - az Európai Unió hasonló rendszerei számára történő adatszolgáltatásra, illetve az onnan átvett adatok értelmezésére.*

A fenti követelményeknek megfelelő szabványok, ajánlások és más műszaki előírások kidolgozását a Közigazgatási Informatikai Bizottság (KIB) feladatává teszi a 222/2009 kormányrendelet. A KIB ennek a kötelezettségének a [www.ekk.gov.hu](http://www.ekk.gov.hu) oldalon közzétett, **Az E-Közigazgatási Keretrendszer projekt eredményeként létrehozott Követelménytár c., 28. számú ajánlásával** tett eleget. A Követelménytár – egyebek között – tételesen megnevezi azokat az elektronikus adatcsere és adatmegjelenítési formátumokat, melyek használatát az elektronikus közigazgatás folyamatában ajánlottnak tekinti, s melyeknek érvényessége nyilvánvalóan kihat a papír alapú iratok megbízható elektronikus másolatának elkészítésére is.

A KIB 28. számú ajánlása a következő adatcsere- és adatmegjelenítési formátumokat tekinti elfogadhatónak a magyarországi e-közigazgatásban:

- Common Format and MIME Type for Comma-Separated Values (CSV) Files, RFC 4180
- Extensible Markup Language (XML) v1.1 Second Edition
- GZIP file format specification version 4.3, RFC 1952
- Internet Message Format, RFC 2822
- ZIP format
- Cascading Style Sheets Language Level 2, revision 1 CSS 2.1 Specification
- DOC Format
- Enhanced Compressed Wavelet (ECW)
- Extensible Hypertext Markup Language
- Graphics Interchange Format (GIF)
- Joint Photographic Experts Group Format (JPEG)
- Moving Picture Experts Group Format (MPG)
- MPEG-1 Layer 3
- Portable Document Format (PDF)
- Portable Network Graphics Format (PNG)
- Quicktime (Mov, qt)
- Open Document Format for Office Applications (OpenDocument)
- Rich Text Format (RTF)
- Tag image file format for image technology (TIFF/IT)
- UTF-8, a transformation format of ISO 10646, RFC 3629
- W3C Extensible Stylesheet Language (XSL) Version 1.1
- W3C HTML 4.01 Specification
- W3C XSL Transformations (XSLT) Version 1.0
- Waveform Audio File Format

A KIB ajánlása szerinti fenti adatformátumok közül a papír alapú dokumentumokról készülő elektronikus másolatok esetében nyilvánvalóan csak néhány – elsősorban képi és/vagy szöveges megjelenítést lehetővé tevő – formátum tekinthető relevánsnak. Ezek között is akadnak azonban olyanok, melyeket a hazai és nemzetközi levéltári szakirodalom nem tekint alkalmasnak elektronikus iratok hosszú távú megőrzésére.

Az elektronikus iratok keletkezésének és tárolásának levéltári szempontjai a hazai jogalkotásban továbbra sem érvényesülnek. **Az elektronikus közszolgáltatásról szóló 2009. évi LX. tv.** kimondja ugyan a központi iratkezelési rendszer keretei között működő elektronikus levéltár létrejöttét (26.§ (2) bekezdés), ennek részletes technikai szabályait azonban nem határozza meg. A felhatalmazással, melyben a törvény 31. §-a a technikai szolgáltatás részleteire vonatkozó szabályok kidolgozásával bízta meg, a kormányzat egyelőre nem élt. Sem **az elektronikus közszolgáltatásról és annak igénybevételéről szóló 225/2009 (X.14.) Korm. rendelet**, sem pedig **az elektronikus közszolgáltatás biztonságáról szóló 223/2009 (X.14.) Korm. rendelet** nem tér ki az elektronikusan keletkező, vagy hagyományos adathordozóról elektronikus formába másolt (digitalizált) iratok használatának és hosszú távú megőrzésének feltételeire. Utóbbi jogszabály foglalkozik ugyan a központi elektronikus szolgáltató rendszerben tárolt adatoknak az Ltv. szerint elfogadott irattári tervek (a törvényben hibásan „levéltári tervek”) alapján történő kezelésével, ennek mikéntjére azonban nem tér ki.

**A közfeladatot ellátó szervezetek iratkezelésének általános követelményeiről szóló 335/2005 (XII.29.) Korm. rendelet** előírja ugyan az elektronikus iratok használhatóságának fenntartását a megőrzési idő lejártáig (62. § (2) bekezdés), ennek technikai részleteivel kapcsolatban azonban csak annyit rögzít, hogy az elektronikus iratok levéltárba adását külön jogszabályban meghatározott formátum szerint kell elvégezni (65.§ (2) bekezdés).

A fentiekből egyértelműen kiderül, hogy a Ltv. 35/A § (2) bekezdésében adott felhatalmazás második, az elektronikus iratok levéltárba adásának követelményeire vonatkozó fele egyelőre kitöltetlen maradt, nem született meg az a jogi szabályozás, mely az elektronikus iratok levéltárba adásának feltételeit meghatározná.

Jelen dokumentum ugyan elsősorban az iratképző szervek által végzett digitalizálással foglalkozik, a levéltárba adható, illetőleg a vélhetően hosszú távon megőrizhető adatformátumok és metaadatok kérdésköre ugyanakkor nem választható el az elektronikus iratok létrejöttének körülményeitől.

Mint látható, a jogalkotónak sürgető feladata az Ltv-ben kapott felhatalmazás alapján az elektronikus iratok levéltárba adásának technikai-műszaki feltételeit meghatározó jogszabály megalkotása. Ennek hiányában jelen dokumentum csak az elfogadott nemzetközi szabványokra, valamint az elektronikus iratok levéltári archiválása terén komolyabb tapasztalatokkal bíró levéltári intézmények gyakorlatára támaszkodhat.

Az elektronikus iratkezelés egy részterületére, mely a jelen dokumentum elsődleges vizsgálódási területe, a fentiekkel ellentétben született jogi szabályozás, habár az így kialakult jogi helyzet korántsem egyértelmű.<sup>1</sup>

A **Polgári perrendtartásról szóló 1952. évi III. tv.** (Pp) 195 § (2) bekezdése a következőképpen fogalmaz:

*„Az eredeti papír alapú közokiratéval azonos bizonyító ereje van a közokiratról – általában műszaki vagy vegyi úton – készült felvételnek (fénykép-, film-, hang- stb.), úgyszintén az eredeti okiratról bármely adathordozó útján készült okiratnak, ha a felvételt, illetve az adathordozóról az okiratot bíróság, közjegyző vagy más hatóság, továbbá közigazgatási szerv vagy ezek ellenőrzése mellett más szervezet készítette. Ugyanilyen bizonyító ereje van az okirat megőrzésére hivatott szerv (pl. levéltár) által vagy ellenőrzése mellett készített felvételnek vagy okiratnak, továbbá a közokiratot kiállító vagy őrzésére hivatott szerv által a felvétel vagy adathordozó útján nyert adatok alapján készített okiratnak.”*

A Pp. tehát nem minden esetben teszi kötelezővé, a másolatnak az eredetivel megegyező bizonyító erejűvé válásához szükséges előfeltétellé az elektronikus aláírás és időbélyegző használatát. A másolatok bizonyító ereje, hitelessége kérdésében a Pp. konkrétan is megnevezi a levéltárakat, az itt végzett elektronikus másolatkészítés ezek szerint nem igényel elektronikus aláírást vagy időbélyegzőt. Más esetekben – különösen a magánokiratról a közokirat kiállítására jogosult által saját ügykörén belül végzett másolás esetén – viszont a Pp. kifejezetten előírja a hitelesítés említett elektronikus eszközeinek alkalmazását (196. § (5) bekezdés).

**Az elektronikus aláírásról szóló 2001. évi XXXV- törvény** 27. § (2) bekezdése felhatalmazta az informatikai területért felelős minisztert, hogy – az e-közigazgatásért felelős miniszterrel, a közigazgatási informatika infrastrukturális megvalósíthatóságának biztosításáért felelős miniszterrel, a köziratok kezelésének szakmai irányításáért felelős miniszterrel, az igazságügyért felelős miniszterrel és a kultúráért felelős miniszterrel egyetértésben – rendeletben szabályozza az elektronikus úton történő másolat készítésének kérdéskörét.

Ezen felhatalmazás alapján született meg a **13/2005 (X. 27.) IHM rendelet a papíralapú dokumentumokról elektronikus úton történő másolat készítésének szabályairól**. A rendelet hatálya a közokiratokról, a papíralapú teljes bizonyító erejű magánokiratokról, valamint a papíralapú számviteli bizonylatokról történő másolatok készítésére terjed ki. Ennek megfelelően a jogszabály nem fedi le a levéltárak illetékességébe tartozó iratképző szervek összességét, mindamelllett előírásai iránymutatóak az elektronikus másolatok készítésének gyakorlatát illetően.

---

<sup>1</sup> Erre vonatkozóan lásd: Az elektronikus iratarchiválás szabályozásának koncepciója. Készítette a Nemzeti Hírközlési és Informatikai Tanács megbízásából az Eötvös Károly Intézet 2008-ban.

A rendelet a másolatkészítés általános szabályai között kötelezően előírja, hogy az eredeti, papír alapú irat és annak elektronikus másolata közti képi vagy tartalmi megfelelést a másolatkészítőnek a metaadatok elhelyezése után hitelesítési záradékkal, valamint elektronikus aláírással kell igazolnia (4.§ (2) bekezdés). (Mint fentebb láttuk, a Pp. ezt nem minden esetben tekinti kötelezőnek.) Az ily módon hitelesített másolat elektronikus aláírásának megállapíthatóságáról az aláírónak az irat őrzési idejéig, nem selejtezhető irat esetében korlátlan ideig gondoskodnia kell. Papír alapú közokirat, illetőleg teljes bizonyító erejű magánokirat másolása esetén a közokirat kiállítására jogosult ezen felül még a kiállítási jogosultságát igazoló információkat, valamint időbélyegzőt is köteles elhelyezni az elektronikus másolaton. Mindez értelemszerűen érinti az elektronikus levéltár feladatait is, hiszen ebben az esetben a levéltárba kerülő iratok hitelességének megállapításához minden bizonnyal az Eat. 6.§ (1)-(2) bekezdései szerinti hitelesítés-szolgáltatást kénytelen igénybe venni, hiszen önmaga vélhetően nem lesz annak a teljeskörű informatikai kapacitásnak – beleértve annak technikai és tudásbeli aspektusait is – a birtokában, melyek segítségével az iratok ilyen értelmű hitelességét garantálhatná.

Mindezt a problémát jelentős mértékben árnyalja az elektronikus aláírás használatának elterjedésével kapcsolatos, informatikusi és közigazgatási szakmai körökben mindinkább terjedő kétely<sup>2</sup>, mely elsősorban a technika bonyolultságát okolja a módszer áttörésének késlekedéséért vagy elmaradásáért. Hangsúlyozandó, hogy mindezek a problémák még fokozottabban érvényesülnek a levéltári megőrzés szempontjából vizsgálva, hiszen az önmagában sem egyszerű, ráadásul változó technikai eljárások hosszú távú archiválása beláthatatlan feladatokat generálna, és aránytalan erőforrás-ráfordítást igényelne a közgyűjtemények részéről. Az elektronikus iratok archiválása terén egyébként is alapvető szempont a széles körben elterjedt, egységes, nyílt, stabil és időtálló formátumok és módszerek alkalmazása, ezen feltételeknek viszont az elektronikus aláírás alkalmazásának jelenlegi hazai gyakorlata nem felel meg. Amíg nem áll rendelkezésre a közigazgatásban egy hosszú távú archiválásra alkalmas, egységesen alkalmazott elektronikus aláírást leíró szabvány, addig a közlevéltárak csak aránytalanul nagy ráfordítással lesznek képesek az ily módon hitelesített dokumentumok integritásának garantálására.

Érdeemes ugyanakkor ismételten felhívni a figyelmet arra a különbségre, mely az iratok jogi értelemben vett *hitelesítése*, illetve levéltári, történeti értelemben vett *hitelessége* között fennáll. A levéltárak a hagyományos, papír alapú iratok átvétele során sem ellenőrzik az egykorú iratok formai jegyeinek (aláírás, pecsét) valódiságát, hiszen erre sem módjuk, sem kapacitásuk nincs. A levéltári értelemben vett hitelességet az iratképző szerv iratkezelése fölötti ellenőrzési jog gyakorlása, illetve az iratok átvételének és későbbi megőrzésének

---

<sup>2</sup> Az elektronikus aláírás használatának helyzetére, illetőleg „nem használatának” helyzetére vonatkozóan lásd a Budapest Főváros Levéltára megbízásából 2010-ben készült adatvagyon-felmérés eredményeit!  
[http://bfl.archivportal.hu/id-1863-adatvagyon\\_felmeres.html](http://bfl.archivportal.hu/id-1863-adatvagyon_felmeres.html)

zárt, ellenőrizhető és dokumentált eljárásrendje garانتálja. Ezt az elvet az elektronikus iratok – beleértve a papír alapú dokumentumokról készült digitális másolatokat is - kezelése során is érvényesíthetőnek, érvényesítendőnek tekintjük.

Amennyiben a jogalkotó a 13/2005 IHM rendeletben meghatározott hitelesítési kötelezettségeket fenn kívánja tartani az elektronikus másolatok készítése és megőrzése során, akkor mindez valószínűleg csak minősített archiválási szolgáltató közreműködésével teheti majd meg.

A rendelet által kötelezően előírt, az elektronikus másolathoz rendelendő metaadatok a következők:

- a papíralapú dokumentum megnevezése;
- a papíralapú dokumentum fizikai méretei;
- a másolatkészítő szervezet megnevezése és a másolatkészítő személy neve;
- a másolatkészítő rendszer, illetve a másolatkészítési szabályzat pontos megnevezése és verziószáma;
- a másolatkészítés ideje;
- érvényességi idő, vagy annak jelzése, hogy az érvényességi idő nem meghatározott.

Megjegyzendő, hogy a Pp. előírásait tekintve ezen metaadatok rögzítése sem előfeltétele az eredetivel azonos bizonyító erejű másolatok készítésének, mindazonáltal kétségkívül elősegíti az eredeti papír alapú dokumentum és annak elektronikus másolata közti megfelelés ellenőrzését.

A területet érintő jelenlegi hazai jogi szabályozás a fentiek értelmében részben hiányzik, részben pedig ellentmondásos. Egyelőre nem született meg az elektronikus iratok levéltárba adására vonatkozó követelményrendszer, pedig ez értelemszerűen ki kellene hasson az elektronikus úton történő másolatkészítés gyakorlatára is.

A jelen dokumentum tárgyában született rendelet egyfelől nincs egyértelmű összhangban a Pp-vel, másrészt olyan feltételeket fogalmaz meg az elektronikus formában készült másolatok tekintetében, melyek levéltári szempontból – a mai tudásunk szerint – hosszú távon aligha alkalmazhatóak.

Korántsem elhanyagolható szempont ugyanakkor, hogy a közlevéltárak illetékességi köre nem csak a Pp-ben, illetőleg a 13/2005. IHM sz. rendeletben meghatározott *közokiratokra*, hanem ennél jóval szélesebb körre, a *köziratokra* terjed ki. Ez utóbbi fogalmat a Levéltári tv. a következőképpen definiálja: *“közirat: a keletkezés idejétől és az őrzés helyétől függetlenül minden olyan irat, amely a közfeladatot ellátó szerv irattári anyagába tartozik vagy tartozott”*. A köziratok ezek szerint jóval szélesebb köre esetében a jogszabályok nem írnak elő

elektronikus aláírással vagy időbélyegzővel történő hitelesítési kötelezettséget (habár mint fentebb láttuk, a Pp. alapján ez még a közokiratok esetében sem egyértelmű). A levéltári iratok hitelességét, megbízhatóságát – ami soha nem az abban foglalt adatok valóságát, pusztán az iratkeletkeztető által átvett iratok változatlanlanságának tényét jelenti – a későbbiekben az elektronikus formába konvertált köziratok esetében is jogszabályi szinten biztosítani kell. Meg kell tehát teremteni azt a jogszabályi szintű garanciát, mely alapján az elektronikus levéltár a papír alapú iratokhoz hasonlóan a megfelelően dokumentált iratátvétellel vállalhatja a felelősséget a kezelésébe került elektronikus iratok megbízhatóságáért és hitelességéért.

Az elektronikus másolatkészítés során (is) nagy mennyiségben kerülnek személyes adatok a dokumentumkezelő rendszerbe. Ezen adatok megfelelő védelméről a személyes adatok védelmével kapcsolatos jogszabályok – elsősorban is az információs önrendelkezési jogról és az információszabadságról szóló **2011. évi CXII. tv.**, valamint az egyének védelméről a személyes adatok gépi feldolgozása során, Strasbourghban, 1981. január 28. napján kelt Egyezmény kihirdetéséről szóló **1998. évi VI. tv.** előírásai szerint kell gondoskodni. Az adatvédelmi jogszabályok érvényesülését az intézményi adatvédelmi és adatbiztonsági szabályzatoknak kell garantálnia.

## 2.2 Javaslatok a hazai jogi szabályozás módosítására

- A Magyar Országos Levéltár részére 2011-ben készített Szabályozási Modell<sup>3</sup> hangsúlyosan említi, hogy jelenleg a Pp., valamint a másolatkészítés szabályaira vonatkozó rendelet nincs összhangban, sőt a Pp. egyes paragrafusai önmagukban sem egyértelműek. A későbbiekben a jogalkotónak a fent részletezett, illetve a Szabályozási Modellben is említett<sup>4</sup> ellentmondásokat fel kell oldania annak érdekében, hogy a külön jogszabályban meghatározott szigorú feltételek mellett készült elektronikus másolat ugyanolyan joghatás kiváltására legyen alkalmas, mint az eredeti papír alapú dokumentum. Garantálni szükséges a Pp-ben (191. §), hogy senkit nem érhet joghátrány, ha a jogszabályoknak megfelelő hiteles másolatot nyújt be bizonyítékként egy esetleges per során. Mindez megnyithatná az utat az elektronikus másolatok mind szélesebb körben történő alkalmazása, valamint bizonyos esetekben (pl. nem maradandó értékű iratok kezelésekor) a papír alapú eredeti megsemmisíthetősége felé.

<sup>3</sup> Az „Elektronikus levéltár” projekt keretében a hosszú távú levéltári megőrzéshez szükséges szabályozási feltételek kidolgozása. Szabályozási Modell.

<sup>4</sup> uo. 123. old.

- Nem kevésbé sürgető jogszabályalkotási feladat az Ltv-ben kapott felhatalmazás alapján az elektronikus iratok levéltárba adásának technikai-műszaki feltételeit meghatározó jogszabály megalkotása.
- A fent említett Szabályozási Modell 5.15.1 pontjával összhangban szabályozni szükséges az iratkezelési szoftverek technikai követelményeire, metaadat-követelményeire, illetve a levéltárba adáshoz szükséges funkcionális követelményekre vonatkozó elvárásokat.
- Szintén rendezendő jogi probléma a levéltárak felelőssége az elektronikus úton keletkező másolatok tekintetében. A jelenlegi jogi szabályozás alapján a levéltárak az eredeti iratokat veszik át, azok elektronikus vagy egyéb másolatai nem tartoznak illetékességükbe. A 335/2005 Korm. rendelet a másolatok kezelésével kapcsolatban a következőket írja elő:

*61. § (2) "A feleslegessé vált munkapéldányokat és másolatokat az ügyiratból ki kell emelni, és a selejtezési eljárás mellőzésével meg kell semmisíteni."*

E paragrafus értelmében tehát a levéltárnak nincs ellenőrzési joga a másolatok kezelése és esetleges selejtezése felett, amennyiben az iratképző azt feleslegesnek ítéli, selejtezési eljárás (tehát levéltári jóváhagyás) nélkül selejtezheti. A későbbiekben célszerű volna ezt a rendelkezést felülvizsgálni abban az értelemben, hogy a külön jogszabályban meghatározott feltételek mellett készített elektronikus másolat őrzési ideje általában az eredeti irat megőrzési határidejével megegyezzen, illetve csak az illetékes levéltár egyedi döntése alapján legyen selejtezhető.

- Általában is célszerű erősíteni a levéltárak felügyeleti és ellenőrzési jogkörét az elektronikus iratok teljes életciklusa felett. Ennek érdekében a Levéltári tv. 11. §-a (A köziratok kezelésének közlevéltári ellenőrzése) abban az értelemben módosítandó, hogy a közlevéltár már az elektronikus iratot – akár így keletkező, akár hagyományos adathordozóról másolt formában – létrehozó informatikai rendszerek tervezése és kialakítása során érvényesíthesse az elektronikus iratokkal szemben támasztott levéltári követelményeket.
- A fenti követelményeket a Magyar Országos Levéltár az Ltv. 17. § (2) bekezdésében, valamint a 10/2002 NKÖM rendelet 5. § (1) bekezdésében meghatározott Ajánlás közzétételével szabhatja meg.
- A közlevéltári iratkezelés felügyeletének egyik kulcsa az irattári tervek illetékes levéltár általi jóváhagyása. A későbbiekben célszerű volna az irattári tervekben meghatározni az



elektronikus másolatok levéltárba adásának lehetséges módjait (lásd még Szabályozási Modell 5.10.1).

- Ezen is túlmenően levéltári szakmai szempontból a digitális másolatkészítés szabályait a jövőben magániratokra is ki kell terjeszteni, hiszen az Ltv. a magánlevéltári anyag védelmével is foglalkozik.

## **2.3 Magyarországi digitalizálási programok**

A hazai iratképző szervek és közgyűjtemények évek óta folytatják a papír alapú iratok elektronikus másolatának készítését, digitalizálását. Ezek a programok tervezése és kivitelezése során nem alakult ki egységesen elfogadott követelményrendszer, mely a fent vázlatosan ismertetett nemzetközi szabványok és ajánlások alapján épülne fel. Mint láttuk, a hazai jogszabályi környezet korántsem nyújt kellő támpontot a digitalizálási folyamatok egységes alapokon történő lebonyolításához. Az alábbiakban a teljesség igénye nélkül néhány megvalósult projekt tapasztalatait foglaljuk össze.

Megjegyzendő, hogy nem csak az iratképző szerveknél, de a közgyűjteményeknél is komoly digitalizálási programok folynak évek óta (említhető itt például a Magyar Országos Levéltár, Budapest Főváros Levéltára vagy az Állambiztonsági Szolgálatok Történeti Levéltára), a levéltári digitalizálás azonban részben más feltételekkel és célokkal zajlik, ezért nem tartozik jelen tanulmány témakörébe.

### **2.3.1 Digitalizálási program a MÁV Zrt. Központi Irattárában**

A MÁV ZRt. Központi Irattára a MÁV Jogi Iroda szervezetében működik.<sup>5</sup> A Központi Irattár 2001 elején alakult meg, amikor az Irattár működtetésére létrehozták a Levéltári csoportot. 2001-2004 között az Irattár ideiglenes helyszínen, a Vezérigazgatóság épületében működött. 2005-ben adták át az Irattár új épületét, amelyben korszerű körülmények teszik lehetővé az iratok és tervrajzok hosszú távú megőrzését és az ügyfelek gyors és szakszerű kiszolgálását.

A MÁV Központi Irattárában tárolt dokumentumok három csoportra oszthatók: irattárakra, tervtárakra és gyűjteményekre. Ezek összesen 66 fondban vannak elkülönítve. Az irattárak legfontosabb részét a hat volt MÁV Igazgatóság és a Záhonyi Üzemigazgatóság nem selejtezhető iratai adják, de egyes főosztályok, osztályok (pl. gépészet, személyügy) és alacsonyabb szervezeti egységek iratai is megtalálhatók. A tervtárak közül a

---

<sup>5</sup> <http://www.mavintezet.hu/leveltar>

legjelentősebbek a MÁVTI volt tervtára, a magasépítmenyi tervtár és a hidász tervtár. Ezek egyenként is több százezer tervrajzot tartalmaznak. Külön említendő az archív tervtárak, melyek 1945 előtti időszak feltárásához nyújtanak információkat.

A MÁV Zrt. Központi Irattárában évek óta zajlik az iratanyagok tömeges és szisztematikus digitalizálása, jelenleg már mintegy 3 millió rekordot, ezen belül 60 millió A/4-es lapnak megfelelő szkennelt oldalt tartalmaz az ezt kezelő egyedi adatbázis. Az adatbázis motor SQL, a feldolgozás a GDL ZRt. által a MÁV számára készített egyedi szoftverben történik. Kuriózumnak számít az iratanyag elektronikus feldolgozottsági szintje is, hiszen jelenleg már a teljes irattári anyag 85 %-a rendelkezésre áll elektronikus formában is. Mennyiségi tekintetben feltétlenül az egyik legtekintélyesebb hazai digitalizálási programnak tekinthető a MÁV-anyagok elektronikus feldolgozása.

A tömeges szkennelés során nagy teljesítményű, automatikus papírtovábbítású szkennert, illetőleg speciális A/0-s tervszkennereket, valamint A/3-as síkgyas szkennereket vettek és vesznek igénybe. A tömeges feldolgozás fekete-fehér színben folyik, szükség esetén használnak (pl. színes fotók esetén) színes szkennert. A digitalizálás 300 ppi felbontásban történik, a rendszer TIFF fájlokat hoz létre Group4 tömörítéssel. Utóbbi minőségromlás nélküli méretcsökkenést eredményez.

Az automatikusan rögzülő technikai adatok mellett az adatrögzítők a következő metaadatokat rögzítik az iratok digitális konvertálása során:

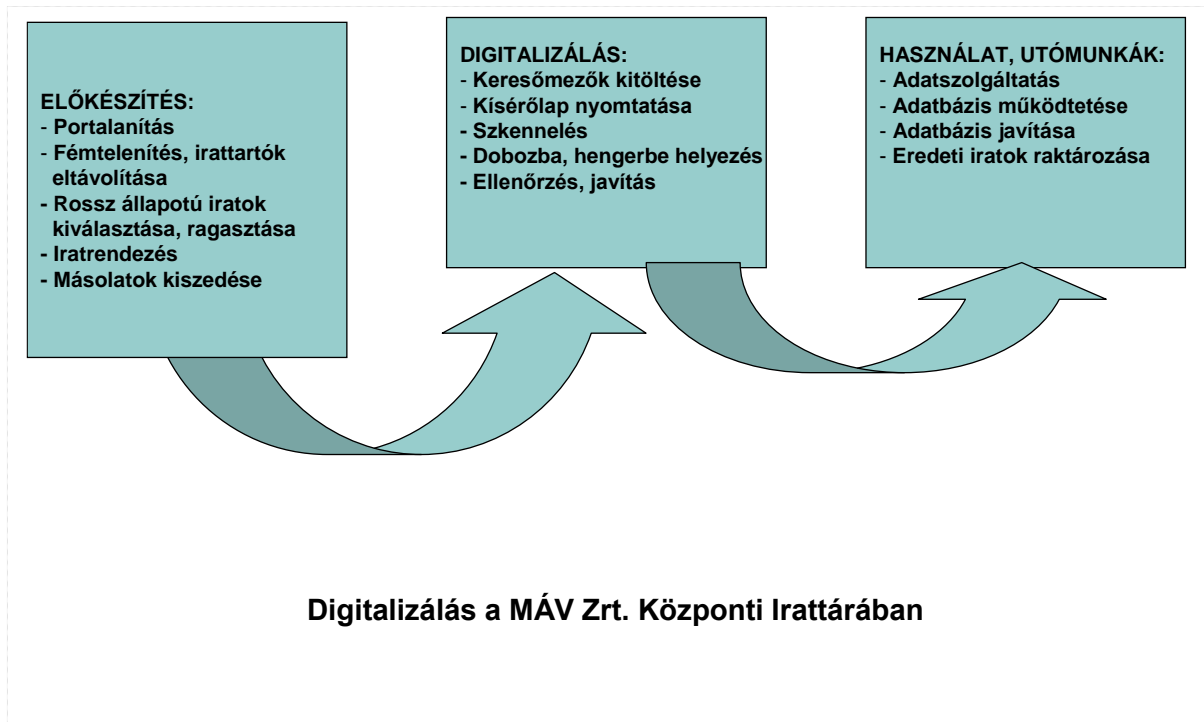
Ügyiratoknál:

Szerv, Szervezeti egység, Tárgy, Iktatószám (Elő- utóirat száma is), Irattári tételszám, Iktatás éve, Raktári hely (vonalkódolvasóval), Egyedi azonosító (vonalkódolvasóval), MÁV dobozkód (vonalkódolvasóval)

Tervrajzoknál:

Tervszám, Rajzszám, Terv tárgya, Tervezés éve, Tervező vállalat, Raktári hely (vonalkódolvasóval), Egyedi azonosító (vonalkódolvasóval), MÁV dobozkód (vonalkódolvasóval)

A MÁV-iratok digitalizálása a következő folyamatábra szerint történik:



### 2.3.2 Digitalizálás az ÁPV Zrt iratanyagában

Nagyságrendileg a MÁV Zrt. iratanyagának digitalizálásához mérhető a Magyar Nemzeti Vagyonkezelő Zrt. és jogelődeinek (Állami Vagyonügynökség, Kincstári Vagyonigazgatóság, Állami Privatizációs és Vagyonkezelő Zrt., Nemzeti Földalapkezelő Szervezet) iratainak feldolgozása.

A PrivDat Dokumentum Archiváló és Tároló Kft.<sup>6</sup> által kezelt iratanyagból mára összesen kb. 19 millió oldalnyit digitalizáltak és archiváltak elektronikus formában.

A munka során kétféleképpen (fekete-fehér) képeket állítanak elő 300 ppi felbontásban, az elkészült digitális másolatokat TIFF fájlformátumban, CCITT Group4 veszteségmentes tömörítéssel tárolják.

Az említett 19 millió oldalból már kb. 4 millió oldalnyit OCR-eztek és komoly energiákat fordítanak a maradék iratanyag karakterfelismerő programmal történő feldolgozására is.

A karakterfelismerésen „átesett” anyagokból kinyert szöveges állományokat TXT, illetve PDF/A formátumban tárolják.

Az iratanyag metaadatokkal való ellátása során a legfontosabb nyilvántartási és a visszakereshetőséget biztosító információk mellett a következő metaadatokat rögzítették:

- a digitalizálást végző személy neve,

<sup>6</sup> <http://www.privdat.hu/>

- az indexáló személy neve,
- az ellenőrző személy neve.

Az archiváló szoftver ezen kívül folyamatosan naplózza a digitalizálás időpontját és az elektronikus iraton történt esetleges változtatásokat is.

### 2.3.3 Digitalizálás a Nemzeti Hírközlési Hatóság (NMHH) iratanyagában

A Nemzeti Hírközlési Hatóság iratanyagának digitalizálása egy központi szkennelő rendszer működtetése révén valósul meg, mely a hatóság Számítógépes Ügyviteli Rendszeréhez (SZÜR) kapcsolódik.

A szkennelést az NMHH munkatársai saját erőből, központilag végzik. Kimenatként 300 DPI-s, fekete-fehér TIFF fájl jön létre, amely akár több bejövő dokumentumot is tartalmazhat. Minden keletkezett képfájl tartalmazza a szkennelést végző felhasználó nevét, valamint a szkennelés pontos dátumát. A rendszer az irat vonalkódja alapján az elektronikus példányt automatikusan becsatolja a megfelelő ügyiratba, megfelelő hozzáférési rendszerrel támogatva.

A karakterfelismerő szoftver a bemenetén kapott 300 DPI-s fekete-fehér TIFF állományból 200 DPI-s, kétrétegű PDF-et készít. Ez azt jelenti, hogy a kimeneti PDF dokumentum külső kinézete a szkennelt dokumentumot (és nem a felismert szöveget) mutatja, de szükség esetén a szöveg egyszerű kimásolás (például ügyiratba másolás) után szerkeszthetővé válik.

## 3. Nemzetközi kezdeményezések

### 3.1 A levéltári leírás és archiválás nemzetközi szabványai

A papír alapú iratok megbízható elektronikus másolatának elkészítése, illetőleg az elkészült másolatok metaadatokkal való szakszerű ellátása, leírása érdekében a folyamat során feltétlenül érvényesítendőnek tartjuk az általánosan elfogadott nemzetközi levéltári leírási szabványok alkalmazását.

A legfontosabb, e területet érintő levéltári szabványok a következők:

- **ISAD(G), Az Általános Levéltári Leírás Nemzetközi Szabványa**<sup>7</sup>

A Nemzetközi Levéltári Tanács által 1999-ben elfogadott dokumentum. A levéltári anyag kezelésének minden fázisainál (mint például keletkezés, értékelés,

---

<sup>7</sup> <http://bfl.archivportal.hu/id-1036-leveltari-anyagleiras-nemzetkozi.html>

gyarapodás, rendezés, állományvédelem) a leírás bizonyos elemei rögzítésre kerülnek, egyrészt az anyag további biztonságos megőrzése és kezelése, másrészt – megfelelő időben, megfelelő jogosultsággal rendelkező személyek részére – hozzáférhetővé tétele érdekében. A levéltári leírás részének tekintendő minden olyan információ, amely a levéltári anyag kezelésének bármely fázisa során születik és rögzítődik. Ezek az információk dinamikusak, bármikor módosíthatók és kiegészíthetők a tartalom és a kontextus alapján; különösen a számítógépen tárolt leírások könnyen helyesbíthetők, módosíthatók vagy integrálhatók. A leírás szabályai elsősorban a levéltári anyag életciklusának attól a ponttól kezdődő szakaszára koncentrálnak, amikor annak tartós megőrzésre való kiválasztása megtörténik, de alkalmazhatók korábbi fázisokban is.

- **EAD - Encoded Archival Description<sup>8</sup>**

A levéltári anyag leírását támogató adatszerkezeti megközelítésű metaadatszabvány. Fejlesztése 1993-ban kezdődött a University of California, Berkeley könyvtárában azzal a céllal, hogy a levéltári anyag segédleteinek, jegyzékeinek, nyilvántartásainak, indexeinek nyílt szabványként szolgáljon. Az EAD 145 adatmezőre bontja a segédletről magáról és a segédlettel ismertetett levéltári anyagról szóló tudnivalókat, amelyek a jelölőnyelv használata miatt könnyen értelmezhetők különféle alkalmazások számára.

- **OAIS - Open Archival Information System, Nyitott Archiválási Információs Rendszer<sup>9</sup>**

A NASA közreműködésével, az ISO fennhatósága alatt működő Consultative Committee for Space Data System (CCSDS) nevű bizottság készítette el az úgynevezett Nyitott Archiválási Információs Rendszer modelljét. Az OAIS referenciamodellt 2001-ben véglegesítették és a Nemzetközi Szabványügyi Szervezet ISO 14721:2003 számú szabványként jegyezte be. A Nyitott Archiválási Információs Rendszer megfogalmazza a hosszú távú megőrzést biztosítani képes archívum minimális követelményeit, fogalmi és koncepcionális keretet állít fel a levéltárak és egyéb gyűjtemények részére.

- **PREMIS (Preservation Metadata: Implementation Strategies)**

A hosszú távú digitális megőrzés fontosságára mind több intézmény figyel fel. Ezek nemzetközi egységesítését hivatott a PREMIS megvalósítani. Az OCLC (Online Computer Library Center) és az RLG (Research Libraries Group) közös vállalkozása 2000-ben indult. Célja, hogy megoldást dolgozzon ki digitális anyagok hosszú távú

---

<sup>8</sup> <http://www.loc.gov/ead/index.html>

<sup>9</sup> <http://public.ccsds.org/publications/archive/650x0b1.PDF>

megőrzésére. 2003-ban hozták létre a PREMIS (Preservation Metadata: Implementation Strategies) munkacsoportot, amely összefogja a további feladatokat. Legfontosabb céljai a metaadat szerkezetének és feladatának meghatározása a megőrzési folyamatban, a meglévő metaadat- sémák szintetizálása, és széles körben elfogadott és alkalmazható rendszer létrehozása. A PREMIS 2003 júniusában kezdte meg a munkáját több mint 30 szakértő részvételével, 237 oldalas ajánlásukat 2005-ben hozták nyilvánosságra *PREMIS Data Dictionary for Preservation Metadata* címmel. A PREMIS gondozását a Library of Congress vállalta.<sup>10</sup>

- **METS – Metadata Encoding and Transmission Standard<sup>11</sup>**

A METS séma a digitális objektumok leíró, adminisztratív és strukturális metaadatainak rögzítésére szolgáló séma. Az amerikai Kongresszusi Könyvtár Network Development and MARC Standards Office gondoskodik a fenntartásáról. A METS magában foglalja a leíró, adminisztratív, viselkedési, megőrzési metaadatokat, lehetővé teszi, hogy ezeknek az elemeknek a leírására a csomag készítője más szabványokat (Dublin Core, EAD, PREMIS stb.) használjon fel.

- **MODS (Metadata Object Description Schema)**

A MODS bibliográfiai adatelemkészletet a Kongresszusi Könyvtár Hálózatfejlesztő és MARC21 Szabványhivatala (Library of Congress Network Development and MARC21 Standards Office) fejlesztette ki azzal a szándékkal, hogy „alkalmas legyen már létező MARC21 rekordok válogatott adatainak továbbítására csakúgy, mint eredeti információforrásokról új bibliográfiai rekordok készítésére.

### 3.2 Moreq, Moreq2<sup>12</sup> és Moreq2010

Az Európai Unió szakosított szervezeteiben a '90-es évektől keresték annak módját, hogy a mind szélesebb körben terjedő elektronikus iratkezelési rendszerek egységes alapelvek és követelményrendszer szerint épüljenek fel, ezáltal biztosítva köztük az interoperabilitást és az adatok átjárhatóságát. Az elektronikus iratkezelésre vonatkozó követelmények átfogó előírásai összeállításának igényét először az 1996. évi DLM Fórumon fogalmazták meg, a megbeszélés során felmerült 10 intézkedési pont egyikeként. Ezt követően az Európai Bizottság IDA (Interchange of Data between Administrations – Kormányzatok közötti adatcsere) programja megbízást adott az elektronikus iratokat kezelő rendszerek (EIKR) mintaszabványjának kidolgozására. Ennek eredményeként 2001-re elkészült a MoReq (Model

<sup>10</sup> <http://www.loc.gov/standards/premis/>

<sup>11</sup> <http://www.loc.gov/standards/mets/>

<sup>12</sup> <http://www.moreq2.eu/moreq2>

Requirements for the management of electronic records), az elektronikus iratok kezelésére vonatkozó mintakövetelmények leírása.

2001-es megjelenését követően a MoReq első, eredeti változatát széles körben vették használatba az európai kontinensen és azon túl. Az Európai Unióban hamar felismerték a MoReq-hez hasonló mintakövetelményekből származó előnyöket, melyek stabil alapot biztosítanak az Elektronikus Dokumentumkezelő Rendszerek követelményeinek kidolgozásához. Megjelenése óta számos szoftverfejlesztő cégben a fejlesztési folyamat részévé vált a MoReq-ben foglalt követelmények feldolgozása.

Az információtechnológia azonban számos változáson ment keresztül 2001 óta. Radikális átalakuláson estek át az elektronikus iratok létrehozását, iktatását és kezelését érintő technológiák. A MoReq 2008-as kiadása, a MoReq2 a technológiai változások figyelembevételével készült. Kiemelt hangsúlyt kapnak az időközben kiadott újabb szabványok és a bevált gyakorlatok. Ennek megfelelően a MoReq2 az eredeti MoReq radikális frissítéseként és kiegészítéseként íródott.

Az elektronikus dokumentumkezelő rendszerek modellkövetelményeit tartalmazó ajánlás értelemszerűen foglalkozik a hagyományos, papír alapú iratok és az elektronikus dokumentumkezelő rendszerek kapcsolatával, ezen belül is az ezen rendszerekbe kerülő papír alapú dokumentumok digitalizálásával, elektronikus másolásával. Az ajánlás vonatkozó fejezete (6.5 Szkenelés és képfeldolgozás) részletezi a papír alapú iratok szkennelésének lehetséges körülményeit, a digitalizálás során alkalmazható eszközöket, fájlformátumokat, az optikai karakterfelismerő szoftverek (OCR) alkalmazásának követelményeit, vagy a szkennelési munkamenet naplózásának mikéntjét.

A MoReq2 ajánlásai általánosan elfogadottá váltak a hazai levéltári szakmai előírásokban, egyebek mellett A levéltárak középtávú informatikai stratégiája és feladatterve (2006-2010) c. dokumentum<sup>13</sup> is hangsúlyozza ennek fontosságát.

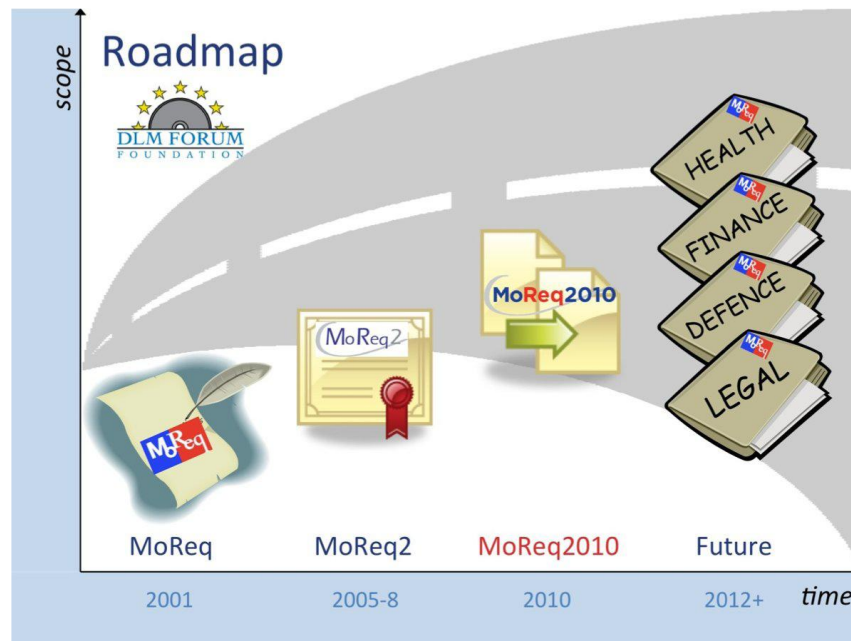
Hangsúlyozandó, hogy a MoReq2 rendkívüli fontosságú ajánlás az elektronikus dokumentumkezelő rendszerek kialakítása, illetve azok interoperabilitása szempontjából, elsődleges célja azonban nem az elektronikus iratok hosszú távú levéltári megőrzésének garantálása. Ráadásul az ajánlás minimum feltételeket fogalmaz meg, melynél az elektronikus levéltár kialakítása során rigorózusabb eljárásrendeket és követelményeket szükséges kialakítani. Ennek megfelelően a MoReq2 ajánlásait csak részben tekintjük jelen témánk szempontjából relevánsnak.

A MoReq2 után 2010-ben megjelent a MoReq2010<sup>14</sup>, mely már nem a „Model”, hanem a Modular szóból eredezteti a betűszót (Modular Requirements for Records

<sup>13</sup> [http://www.mol.gov.hu/letoltes.php?d\\_id=162](http://www.mol.gov.hu/letoltes.php?d_id=162)

Systems). Ez a – több mint félezer oldalas - verzió már nem általános követelményrendszert fogalmaz meg, hanem egyes területekre alkalmazható speciális elvárásokat (lásd a lenti ábrát!), illetőleg igyekszik alkalmazkodni az egyre heterogénebbé váló iratkezelési rendszerekhez. A fenti okokból a Moreq2010, kevésbé konkrét és témánk szempontjából kevés gyakorlati útmutatást jelent. (Várhatóan hasonló lesz a helyzet a 2012-ben várható újabb Moreq-verzióval is.)

Az alábbi ábra jelzi a Moreq továbbfejlesztésének az irányait.



### 3.3 A metaadatok nemzetközi szabványa, a Dublin Core

A metaadatok, vagyis „az adatot leíró adatok” meghatározó fontosságúak az elektronikus iratok kezelése során, hiszen azok értelmezésének, rendszerezhetőségének és visszakereshetőségének zálogát biztosítják. A 24/2006-os BM-IHM-NKÖM közös rendelet megfogalmazása szerint a metaadat: *„strukturált vagy félig strukturált információ, amely lehetővé teszi iratok létrehozását, kezelését és használatát hosszabb időn át azon tartományokon belül, amelyekben létrehozására sor kerül”*.

A közgyűjteményekben – könyvtárakban, levéltárakban – őrzött kulturális kincsek egységes leírásának nemzetközi és hazai szinten is általánosan elterjedt és használt metaadat-készlete az ún. Dublin Core.<sup>14</sup> A Dublin Core Metadata Initiative (DCMI = Dublin Core Metaadat Kezdeményezés) Igazgatósága az OCLC Kutatási és Speciális Projekt Irodája keretében működik, amely a szabályzat gondozásáért felelős. A

<sup>14</sup> <http://moreq2010.eu/>

<sup>15</sup>lásd: <http://www.dublincore.org/documents/2010/10/11/dces/#ISO15836>



szabvány előszava leírja annak eredetét, amely 1995-re nyúlik vissza, amikor is könyvtárosok, elektronikus források szakértői, tartalom-szolgáltatók gyűltek össze, hogy megállapodjanak egy könnyen érthető és alkalmazható forrás-leíró szabályzatban bármely tudományág területén. Az érdeklődés folyamatosan nőtt a DCMI iránt és mára több mint 20 nyelvre fordították le dokumentumait, amelyeket átvett a CEN/ISSS (Európai Szabványosítási Tanács/Információs Társadalom Szabványosítási Rendszere) és 2003-ban hivatalosan is ISO szabvánnyá vált.<sup>16</sup> A Dublin Core-nak egyszerre vannak erősségei és gyenge pontjai. A metaadatok létrehozásának egyszerűsége csökkenti a költségeket és elősegíti az együttműködést, de nem jelent megoldást minden jelentésbeli és funkcionális követelmény kielégítésére. Ugyanakkor mindezekkel kiegészíthető és összeegyeztethető, azaz kiinduló alapként szolgál a teljesebb körű leírások elkészítéséhez. A metaadat elemek beilleszthetők a MARC formátumú rekordokba. A DC HTML-ben és XML-ben egyaránt címkézhető 15 adatelemet tartalmaz. A 15 adatelem az egyes dokumentumok leírásának alapelemeit tartalmazza (Cím, Téma, Kiadó, Dátum, Típus, Formátum, Azonosító, Létrehozó, Forrás, Nyelv, Jogok, Tartalmi leírás, Közreműködő, Kapcsolat, Térbeli- és időbeli vonatkozások).

### 3.4 Példák a nemzetközi gyakorlatból

Az alábbiakban a teljesség igénye nélkül néhány nemzetközi viszonylatban is jelentős digitalizálási programot tekintünk át.

#### 3.4.1 Az ENSZ digitalizálási szabványa

Az Egyesült Nemzetek Szervezetének levéltári és iratkezelési részlege (UN ARMS<sup>17</sup>) 2009 áprilisában hagyta jóvá és 2011 áprilisában vizsgálta felül a digitalizálás iratkezelési követelményeit tartalmazó szabványát. A Record-keeping Requirements for Digitization<sup>18</sup> c. tömör, lényegretörő (27 oldalas) dokumentum négy nagyobb fejezetben tárgyalja a papír alapú vagy filmen tárolt iratanyagok szkennelésének lépéseit (A digitalizálás tervezése, A digitalizálást előfeltételei, A digitális konverzió, A digitalizálás utáni teendők), ezenkívül függelékben közli a digitalizálásra vonatkozó technikai specifikációt.

<sup>16</sup> ISO 15386, [http://www.niso.org/publications/isq/free/FE\\_DCMI\\_Harper\\_isqv22no1.pdf](http://www.niso.org/publications/isq/free/FE_DCMI_Harper_isqv22no1.pdf)

<sup>17</sup> United Nations Archives and Records Management Section), <http://archives.un.org/ARMS/>

<sup>18</sup> [http://archives.un.org/ARMS/sites/ARMS/uploads/files/Standard\\_RKreqfor%20digitisation3\(1\).pdf](http://archives.un.org/ARMS/sites/ARMS/uploads/files/Standard_RKreqfor%20digitisation3(1).pdf)

A „szabvány” titulust viselő dokumentum<sup>19</sup> tartalmából két dolgot feltétlenül érdemes kiemelni.

A dokumentum a függelékben közölt táblázatban hét különböző irattípus (fekete-fehér szöveg, vízjelet vagy szürke árnyalatokat tartalmazó dokumentum, egyszerű betű- vagy grafikonszíneket tartalmazó dokumentum, fekete-fehér fénykép, színes fénykép, fekete-fehér negatív film, színes negatív film) esetében tételesen meghatározza az alkalmazható felbontást, színmélységet, fájlformátumokat és a tömörítési lehetőséget. Általánosságban elmondható, hogy a szabvány a nagy felbontású, veszteségmentes tömörítést alkalmazó fájlformátumokat támogatja (PDF, ill. PDF/A, TIFF), a JPEG formátumú képek használatát viszont csak adatcsere és web-es megjelenítés céljából javasolja.

A szabvány másik említésre méltó eleme, hogy a metaadat-struktúrában alapvetően a Dublin Core elemkészletét használja fel, ugyanakkor nem javasol a digitalizálásra vonatkozó speciális metaadat-mezőket, melyek a digitális másolás körülményeit tartalmazzák. Igaz ugyan, hogy nyomatékosan hangsúlyozza a digitális másolat készítése során készítendő dokumentációk fontosságát, melyek révén a digitális másolat hitelessége garantálható, ezen dokumentálás mikéntjére azonban nem tesz javaslatot. (Hasonló a helyzet egyébként az ARMS-nek az iratkezelés során alkalmazandó metaadatok tárgyában kiadott útmutatójával is<sup>20</sup>.)

### 3.4.2 Az EU Domec projektje

Az e-DOMEAC (*Electronic Archiving and Document Management in the European Commission*) az Európai Bizottság elektronikus archiválási és dokumentumkezelési rendszere.<sup>21</sup>

A 2002-ben elindított e-DOMEAC projekt a Bizottság elektronikus dokumentumarchiválás és -kezelés terén lefektetett új elveit valósítja meg. A több éves szabályozási munka során a Bizottság két határozatában, valamint ezek végrehajtási szabályaiban fogalmazta meg az elektronikus iratok archiválásának és kezelésének alapelveit és gyakorlati teendőit. A rendszer jogi szabályozásáért a Főtitkárság felel. Ugyancsak a Főtitkárság irányítása alá tartozik számos olyan projekt, amely a dokumentumkezelést elősegítő informatikai alkalmazásokra irányul.

---

<sup>19</sup> Megjegyzendő, hogy az ARMS honlapján ugyanez a dokumentum máshelyütt nem szabványként (standard), hanem csak irányelvként (guidelines) szerepel.  
<http://archives.un.org/unarms/en/unrecordsmgmt/unrecordsresources/guideline%20on%20records%20digitisation.htm>

<sup>20</sup> Standard on Recordkeeping Metadata,  
[http://archives.un.org/unarms/doc/ARMS\\_Standard\\_Recordkeeping\\_Metadata\\_Final.pdf](http://archives.un.org/unarms/doc/ARMS_Standard_Recordkeeping_Metadata_Final.pdf)

<sup>21</sup> Document management in the European Commission — Collected decisions and implementing rules.  
Luxembourg: Publications Office of the European Union.

A fent említett dokumentumok a következők:

- A Bizottság 2002. január 23-i, 2002/47/EC ECSC, Euratom sz. határozata, melyben eljárási szabályait kiegészítette az iratkezelésre vonatkozó előírásokkal,
- A Bizottság 2004. július 7-i, 2004/563/EC Euratom sz. határozata, melyben eljárási szabályait az elektronikus és digitalizált iratok kezelésére vonatkozó szabályzattal egészítette ki,
- A Bizottság főtitkára által kiadott, a fenti határozatok végrehajtási utasítását tartalmazó, 2009. november 30-án keletkezett dokumentum (SEC(2009)1643, 30.11.2009).

A Domec projekt keretében kiadott fenti végrehajtási utasítás III.2.2.6.3 pontja foglalkozik a digitalizálás kérdésével. Az utasítás alapelveként leszögezi, hogy a Bizottság annak érdekében, hogy teljes körű elektronikus ügyintézés tegyen lehetővé és megteremtse a kizárólagosan elektronikus ügyintézésre való fokozatos áttérés feltételeit, szisztematikusan digitalizálja az egyéb formában, elsősorban papíron keletkező vagy így érkező iratokat.

Az alkalmazandó formátumok tekintetében az olvashatóságot, az állandóságot és a hosszú távú hozzáférhetőséget nevezi meg feltételnek a dokumentum. Ennek érdekében – nem zárva ki a későbbiekben esetlegesen általánosan elfogadottá váló formátumok bevezetését – a papír alapú dokumentumok digitalizálása során a PDF illetve PDF/A, valamint a min. 300 ppi felbontású TIFF formátumú képek előállítását írja elő. Mindkét esetben megengedi a szöveges keresést lehetővé tevő optikai karakterfelismerés (OCR) alkalmazását, de csak az eredeti PDF ill. TIFF állományok érintetlenül hagyásával. Megjegyzendő, hogy a fenti formátumokat az utasítás csak az iratkezelési rendszerben tartja elfogadhatónak, az elektronikus iratok hosszú távú megőrzése, levéltári archiválása során minden képi állományt az ISO 19005-ös szabvány szerint PDF/A formátumúvá konvertálva kell megőrizni.

A főtitkári utasítás nagy hangsúlyt fektet a digitalizálás minőségi kontrolljára, aminek érdekében – elővigyázatosságból – az eredeti irat szabályos őrzési idejétől függetlenül – hat hónapos időtartamra kötelezően előírja az eredeti irat megőrzését az esetleges hibák korrigálása érdekében.

### 3.4.3 A DOMEA-koncepció és a SAGA-ajánlás Németországban

A németországi közigazgatás informatikai korszerűsítésért felelős állami hivatal<sup>22</sup> az 1990-es végétől kezdve fokozatosan alakította ki a német közigazgatásban alkalmazandó elektronikus iratkezelésre vonatkozó ajánlások rendszerét, mely a DOMEA-Konzept (*Dokumenten-Management und elektronische Archivierung*) nevet viseli.<sup>23</sup> A koncepció 2005 novemberében született 2.1-es verziója az elektronikus iratkezelés egész területét igyekszik lefedni az alapelvek lefektetésétől az egyes munkafolyamatokra vonatkozó gyakorlati megoldási modellekig. A koncepció „váza”<sup>24</sup> az elektronikus iratkezelésre vonatkozó alapelvek lefektetése mellett végigkíséri az irat teljes életciklusát annak keletkezésétől (beérkezésétől) az iratkezelés egyes fázisain keresztül egészen a levéltárba adásig illetve selejtezésig. A koncepció részét képezi egy igen részletes fogalommagyarázat és szakirodalmi tájékoztató is.

A DOMEA-Koncepció központi dokumentumához az egyes részterületeket lefedő, összesen 11 kiegészítő modul készült, melyek egyebek között foglalkoznak az elektronikus dokumentumkezelő rendszerek és az elektronikus hitelesítés problémájával, a digitalizálással, hivatali és hivatalközi kommunikációval, az elektronikus iratok selejtezésével és megőrzésével, illetve ezek technikai aspektusaival stb.

Jelen témánk szempontjából a DOMEA-Konzept kiegészítő moduljai közül a hagyományos adathordozón készülő vagy érkező dokumentumok digitalizálásával foglalkozó, Scan-Prozesse c. dokumentum érdemel külön figyelmet. A 61 oldalas dokumentum az általános fejezeteken túl részleteiben is tárgyalja a különböző digitalizálási módokat és stratégiákat illetve annak időbeli és térbeli vonatkozásait, a szkennelési folyamat kivitelezést, az elkészült elektronikus irat indexelését és kezelését, az optikai karakterfelismerés (OCR) illetve az intelligens karakterfelismerés (ICR)<sup>25</sup> alkalmazásának lehetőségeit, a digitális másolatok minőségi kontrolljára vonatkozó előírásokat, az esetlegesen alkalmazható képjavítási eljárásokat, illetve a szkennelés technikai aspektusait.

---

<sup>22</sup> Koordinierungs- und Beratungsstelle der Bundesregierung für Informationstechnik in der Bundesverwaltung im Bundesministerium des Innern, KBSt

<sup>23</sup> Lásd ezzel kapcsolatban: Haraszti Viktor: A DOMEA-Koncepció és hatása az elektronikus iratkezelésre Németországban. Levéltári Szemle 2007/2. sz. 36-46.; ill. [http://www.verwaltung-innovativ.de/cln\\_108/nn\\_684536/sid\\_73033C3F786FEC4638BB150AD644AA3/DE/Organisation/domea\\_\\_konzept/domea\\_\\_konzept\\_\\_inhalt.html?\\_\\_nnn=true](http://www.verwaltung-innovativ.de/cln_108/nn_684536/sid_73033C3F786FEC4638BB150AD644AA3/DE/Organisation/domea__konzept/domea__konzept__inhalt.html?__nnn=true)

<sup>24</sup> DOMEA\_Konzept, Organisationskonzept 2.1. Dokumentenmanagement und elektronische Archivierung im IT-gestützten. Geschäftsgang

<sup>25</sup> A géppel írt, illetve nyomtatott szövegek felismerésére és konvertálására szolgáló OCR technológiához képest az ICR betűszó a kézzel, nyomtatott betűkkel írt szövegek értelmezését és egyéb intelligens funkciókat tartalmazó (pl. a szöveg kontextusát is elemző) programokat jellemzi.

A DOMEA-Konzept kiegészítő moduljai közül az elektronikus iratok archiválásának technikai szempontjaival foglalkozó, Technische Aspekte der Archivierung elektronischer Akten<sup>26</sup> c. dokumentum tekinthető még az elektronikus másolatok készítése szempontjából relevánsnak. Az ajánlás részletesen foglalkozik az elektronikus archiválás stratégiai kérdéseivel, az emuláció vagy migráció problémájával, egy ajánlott archiválási rendszer komponenseivel, a tárolóeszközök technikai kérdéseivel, a formátumok hosszú távú értelmezhetőségével vagy az elektronikus adatok konzisztenciájának garanciáival is.

A DOMEA-konceptió nagymértékben épít a német közigazgatásban használatos ún. SAGA-ajánlásra<sup>27</sup>. Ez utóbbi dokumentum az elektronikus iratkezelési rendszerek használata során konkrétan is megnevezi a hosszú távú archiválás szempontjából elfogadhatónak tekinthető fájlformátumokat, melyeket az elektronikus másolatkészítés (digitalizálás) folyamatában is alkalmazni szükséges. Ezek a – későbbiekben tételesen is ismertetendő – TIFF 6.0 a képek rögzítésére, bizonyos megszorításokkal a JPEG formátum fényképek tárolására, illetve PDF/A-1 (ISO-19005-1:2005) a szöveges dokumentumok és képi megfelelőjük párhuzamos megőrzésére.

### 3.4.4 A brit példa: PRONOM és DROID

A brit nemzeti levéltár évek óta élen jár a nemzetközi levéltári szakmában az elektronikus iratok kezelése terén. A National Archives honlapján<sup>28</sup> előkelő helyet foglalnak el a digitális információk megőrzésével kapcsolatos útmutatók, szakirodalmak, ajánlások, valamint a közvetlenül és ingyenesen igénybe vehető szolgáltatások (Digital Continuity Service). A National Archives saját honlapján az egyes fázisokon lépésről-lépésre átvezető ajánlást kínál mindazon intézmények, szervezetek vagy magánszemélyek számára, akik az általuk keletkeztetett digitális adatok hosszú távú megőrzésével foglalkozni kívánnak. Ezen ajánlás-sorozat része sok egyéb mellett a digitális megőrzés akciótervére, követelményrendszerének megfogalmazására, kockázatelemzésére és kockázatkezelésére, veszteségkezelésére, illetőleg a digitális megőrzési politikának az általános IT-környezetbe való integrálására vonatkozó ajánlások is. A fentiek mellett a National Archives részletes útmutatót és tájékoztatót is kínál az egyes elektronikus fájlformátumok használhatóságáról, előnyeiről és hátrányairól, különös tekintettel azok várható

<sup>26</sup> Technische Aspekte der Archivierung elektronischer Akten, [http://www.verwaltung-innovativ.de/cln\\_110/nn\\_1007684/SharedDocs/Publikationen/DE/domea\\_\\_konzept\\_\\_technische\\_\\_aspekte\\_\\_der\\_\\_archivierung\\_\\_elektronischer\\_\\_akten,templateId=raw,property=publicationFile.pdf/domea\\_konzept\\_technische\\_aspekte\\_der\\_archivierung\\_elektronischer\\_akten.pdf](http://www.verwaltung-innovativ.de/cln_110/nn_1007684/SharedDocs/Publikationen/DE/domea__konzept__technische__aspekte__der__archivierung__elektronischer__akten,templateId=raw,property=publicationFile.pdf/domea_konzept_technische_aspekte_der_archivierung_elektronischer_akten.pdf)

<sup>27</sup> Standards und Architekturen für E-Government, [http://www.cio.bund.de/DE/Standards/SAGA/saga\\_node.html](http://www.cio.bund.de/DE/Standards/SAGA/saga_node.html)

<sup>28</sup> <http://www.nationalarchives.gov.uk>

élettartamára és hosszú távú fenntarthatóságuk (olvashatóságuk, értelmezhetőségük) szempontjaira.

Mindezen lehetőségek közül érdemes külön is figyelmet fordítani egy valóban eredeti és hasznos szolgáltatásra, a digitális fájlformátumok, az ezeket előállító szoftvertermékek és egyéb, a digitális megőrzéssel kapcsolatos technikai kérdések információs forrásául foglalkozó PRONOM<sup>29</sup> rendszerre. Az online igénybe vehető szolgáltatás részletes technikai információt szolgáltat mindazon elektronikus fájlformátumokról, szoftverekről és vállalkozásokról, melyek az elektronikus iratkezelés terén érintettek, és pontos eligazítást ad a termékek és szolgáltatások kompatibilitásáról, interoperabilitásáról és a hosszú távú megőrzés szempontjából való előnyeikről/hátrányaikról.

A PRONOM rendszerhez a nemzeti levéltár ingyenesen használható, nyílt forráskódú szoftvert is fejlesztett DROID (Digital Record Object Identification)<sup>30</sup> néven. A program célja, hogy megvizsgálja és tájékoztatást nyújtson a felhasználó gépén található digitális fájlok formátumának legfőbb jellemzőiről a PRONOM elsődlegesen levéltári, megőrzési szempontjai szerint.

A digitális másolatkészítés terén a brit példából két kérdést érdemes még kiemelni.

Egyéb levéltári intézményekhez hasonlóan a National Archives is a PDF formátumot tekinti elfogadhatónak (azon belül is a PDF/A 1-a verziót) írott dokumentumok képi és szöveges tárolására. Némileg eltérően a nemzetközi gyakorlattól a brit nemzeti levéltár a rastergrafikus képek tárolására a JPEG2000 formátum veszteségmentesen tömörített változatát preferálja, s csak másodlagosan támogatja az egyébként általánosabban elterjedt TIFF-formátumú képek használatát.

Érdemes megemlíteni, hogy a PRONOM-hoz hasonló kezdeményezés az Egyesült Államokban is született GDFR (Global Digital Format Registry)<sup>31</sup>, melynek célja a PRONOM-hoz hasonlóan a digitális formátumokra vonatkozó információk felkutatása, összegyűjtése és szolgáltatásként történő megosztása.

A PRONOM továbbfejlesztésére irányuló projekt 2009-ben indult UDFR (Unified Digital Format Registry) néven brit, kanadai, amerikai, német, új-zélandi és holland közgyűjtemények (könyvtárak és levéltárak) közreműködésével.

---

<sup>29</sup> <http://www.nationalarchives.gov.uk/PRONOM/Default.aspx>

<sup>30</sup> <http://droid.sourceforge.net/>

<sup>31</sup> <http://www.gdfr.info/docs.html>

### 3.4.5 A Svájci Szövetségi Levéltár gyakorlata

A Svájci Szövetségi Levéltár (Schweizerischen Bundesarchiv<sup>32</sup>) a digitális fájlformátumok közti eligazodást, illetőleg az egyes formátumok archiválhatóságát nagymértékben elősegítő, tömör és kategorikus követelményrendszert alakított ki 2007-ben. A digitális iratok archiválásának követelményrendszere<sup>33</sup> című, mindössze 14 oldalas dokumentum második oldala rövid táblázatban közli azt a hat fájlformátumot, melyet a levéltár a hosszú távú megőrzés szempontjából alkalmazhatónak tart, s ennek megfelelően elfogad az iratképző szervektől. (A dokumentum többi része lényegében már csak ezen formátumok részletes specifikációit tartalmazza.) A hat archiválható fájlformátum ezek szerint a következő: strukturálatlan TXT a szöveges állományok esetében, PDF/A hivatali dokumentumok archiválására, CSV formátum a táblázatok, SIARD RDB DATA a relációs adatbázisok, TIFF a raszteres képek, illetve WAVE a hangfájlok számára. Bár a fenti megkötések nyilvánvalóan a hosszú távú levéltári megőrzés szempontjait tükrözik - egy dokumentumkezelő rendszerben pedig átmenetileg érvényesíthetők egyéb szempontok is -, a levéltár álláspontja értelemszerűen kihat az iratképzők digitalizálási programjainak tervezésére, kialakítására is.

### 3.4.6 Egyéb nemzetközi példák

A fenti példák közös jellemzője, hogy egy adott szervezetben vagy országban meghatározott szerepet betöltő levéltári intézmény többé-kevésbé kategorikus követelményrendszert alakított ki a papír alapú dokumentumok digitalizálását, az elkészült elektronikus másolatok fájlformátumait és metaadatait illetően. Számos – nem kevésbé tekintélyes – külföldi levéltári intézmény ugyanakkor az informatikai változások gyorsaságát és kiszámíthatatlanságát figyelembe véve inkább követelményrendszert alakított ki a digitalizálási formátumok egyedi értékeléshez, melyet az egyes digitalizálási projektek megvalósítása során érvényesíteni ajánl.

A Library of Congress például hét pontban foglalta össze a fájlformátumok értékelésének szempontjait<sup>34</sup>: nyilvánosság, elterjedtség, átláthatóság, a digitális objektum belső dokumentáltsága (amilyen mértékben az tartalmazza az alapvető leíró, technikai és egyéb adminisztratív metaadatokat), hardver- és szoftverfüggőség, a szabványok okozta esetleges korlátozott felhasználhatóság, illetve az esetleges alkalmazott védelmi technológiák (pl. kódolás).

<sup>32</sup> <http://www.bar.admin.ch/>

<sup>33</sup> Archivable file formats. Standards for Archiving Digital Records. July 2007.

<sup>34</sup> [http://www.digitalpreservation.gov/formats/intro/format\\_eval\\_rel.shtml](http://www.digitalpreservation.gov/formats/intro/format_eval_rel.shtml)

Hasonló utat követ a dán Állami és Egyetemi Könyvtár és a dán Királyi Könyvár A fájlformátumok kezelése c. ajánlása<sup>35</sup>, mely az alkalmazhatóság kritériumait az egyes fájlformátumok nyitottságában, szoftver-, hardver- és intézményfüggetlenségében, elterjedtségében, minőségében (beleértve a kis helyigényt, megbízhatóságot, egyszerűséget, veszteségmentességet, metaadat-támogatást stb.), illetve az esteleges elavulás folyamatos nyomon követésében látja.

A Holland Nemzeti Könyvtár szintén az értékelési kritériumok megadását preferálja a konkrét fájlformátumokra vonatkozó ajánlással szemben<sup>36</sup>. Az egyes formátumokkal szemben megfogalmazott kritériumok nagymértékben hasonlóak a korábbiakhoz: nyitottság, elfogadottság, összetettség, védelmi mechanizmusok, ön-dokumentáltság, megbízhatóság, illetve külső függőségek. Megemlíthető, hogy az ajánlás pontrendszer is tartalmaz, mely alapján egy adott formátumot a fenti kritériumok alapján egy 100 pontos skálán el lehet helyezni. (A példa kedvéért függelékében közöl a dokumentum egy összevetést az Adobe PDF/A-1 és a Microsoft Word 97-2003-as formátuma között a fenti pontozás szerint: előbbi 89,01%-ot, utóbbi 21,83 %-ot „teljesített” ebben az összevetésben.)

Hollandián belül maradv a Nemzeti Levéltár a fényképek digitalizálásához kiadott útmutatójában viszont konkrét fájlformátumot ír elő (TIFF 6.0 a mesterpéldányoknak, 1:10-es tömörítési arányú JPEG képeket felhasználásra), a metaadat-szerkezetre ugyanakkor nem tesz ajánlást.

A fenti sorba illeszthető az Egyesült Államok Levéltári Szervezete, a NARA (National Archives and Records Administration) által felügyelt kezdeményezésre<sup>37</sup> 2009 novemberében született, Digitalizálási tevékenység<sup>38</sup> c. útmutató is. Ez a dokumentum szintén nem kíván állást foglalni egyes megoldások mellett vagy ellen, ehelyett módszereket kínál az egyedi döntések meghozatalához. A szerzők szerint az egyes digitalizálási projektek céljai különbözőek lehetnek (a széleskörű hozzáférés biztosítása, állományvédelem, publikáció, kutatási igény kiszolgálása stb.), de az eredeti adathordozó típusa, minősége, állapota, a megőrzés kockázati szintje is nagymértékben befolyásolhatja azon megoldások kiválasztását – beleértve a fájlformátumok vagy a metaadat-szerkezet kérdését is -, melyek a legalkalmasabbak a digitalizálás elvégzésére.

---

<sup>35</sup> Handling file formats. 2004.

[http://www.kb.nl/hrd/dd/dd\\_links\\_en\\_publicaties/publicaties/KB\\_file\\_format\\_evaluation\\_method\\_27022008.pdf](http://www.kb.nl/hrd/dd/dd_links_en_publicaties/publicaties/KB_file_format_evaluation_method_27022008.pdf)

<sup>36</sup> Evaluating File Formats for Long-term Preservation.

[http://www.kb.nl/hrd/dd/dd\\_links\\_en\\_publicaties/publicaties/KB\\_file\\_format\\_evaluation\\_method\\_27022008.pdf](http://www.kb.nl/hrd/dd/dd_links_en_publicaties/publicaties/KB_file_format_evaluation_method_27022008.pdf)

<sup>37</sup> Federal Agencies Digitization Guidelines Initiative

<sup>38</sup> Digitization Activities. Project Planning and Management Outline. November 2009.

<http://www.digitizationguidelines.gov/guidelines/DigActivities-FADGI-v1-20091104.pdf>



A fenti megoldások ideális esetben kiegészítik egymást, vagyis az általános kiválasztási szempontokon túl konkrét ajánlásokat is megfogalmaz az adott levéltári intézmény.

## 4. Az elektronikus másolatkészítési folyamatok általános előfeltételei

Papír alapú eredeti irattári/levéltári anyagról történő szisztematikus másolatkészítési folyamatok tervezése során elengedhetetlen bizonyos alapelvek előzetes rögzítése és azok szigorú betartása, betartatása, hiszen csak ezen az úton válik az elektronikus másolat hosszabb távon és széles körben is felhasználhatóvá és megőrizhetővé. Az egyes projektek megkezdése előtt feltétlenül tisztázni és rögzíteni kell az alábbi alapvető kérdéseket.

- Az adott másolatkészítési folyamat célja (a hivatali ügyintézés támogatása, hosszú távú megőrzés, állományvédelem, publikálás, széles körű megismerhetőség biztosítása stb.).
- Az ismert és a feltételezhető<sup>39</sup> eredeti dokumentumok típusai, tulajdonságai, méretei, állapota.
- A digitális másolatkészítés helye az iratkezelés folyamatában, az egyes munkafolyamatok kapcsolódása a hagyományos ügymenethez.
- Az elkészült elektronikus másolatok helye az iratkezelés rendszerében.
- Az eredeti, papír alapú dokumentumok további kezelésének lépései.
- A másolás során létrejövő digitális objektumok jellemzői (hosszabb távú megőrzésre és munkapéldánynak egyaránt):
  - o az alkalmazandó fájlformátumok és azok jellemzői;
  - o az eredeti („mester”) és az esetlegesen készítendő használati példányok paramétereinek elkülönítése;
  - o a másolatkészítés technikai körülményei;
  - o a másolásakor rögzítendő szükséges metaadatok köre.
- A másolatkészítési folyamat során alkalmazandó eljárások tekintetében:
  - o a másolatkészítés időbeli és térbeli meghatározása;

---

<sup>39</sup> Egy elektronikus dokumentumkezelő rendszer alkalmazása esetén például egy közigazgatási szerv nagy valószínűséggel prognosztizálhatja, hogy a hozzá hivatalból érkező és digitalizálandó papír alapú dokumentumok nagy többsége A/4-es méretű és nyomtatott szöveget, valamint aláírás(oka)t tartalmazó irat lesz.

- az elkészült másolatok azonosításához és használatához szükséges metaadatok körének meghatározása;
  - annak definiálása, hogy a munkafolyamatban ki és milyen metaadatokat köteles rögzíteni;
  - annak meghatározása, hogy a metaadatokat hol és mi módon kell tárolni (a digitális objektumba ágyazva, külön rendszerben vagy vegyesen);
  - az elkészült elektronikus másolat és az eredeti irat viszonyának meghatározása, összetartozásuk garantálásának módszerei;
  - a metaadatok formátumának és szerkezetének meghatározása;
  - az esetlegesen alkalmazott hitelesítési eljárások helyének meghatározása a másolatkészítés folyamatában.
- A digitális másolatok felhasználása terén:
- az elkészült másolatok minőségi kontrolljának meghatározása;
  - az elkészült másolatok kezeléséért és biztonságáért felelős személyek, szervezetek körének meghatározása;
  - a másolatok használatának szabályozása;
  - a másolatok hosszú távú megőrzési vagy selejtezési eljárásrendjének kialakítása.

## 5. Formátum-követelmények

A papír alapú iratokról készülő elektronikus másolatok hosszú távú megőrzésének egyik kulcskérdése a másolatkészítő rendszerek által előállított digitális fájlok megfelelő formátuma. A „megfelelőség” ebben az értelemben azt az alapvető levéltári szempontot kell jelentse, hogy a tárolt adatok minél hosszabb távon használhatóak, értelmezhetőek és megbízhatóak maradjanak. Ezen feltételek önmagukban is igen összetettek, ezért – mint azt a fenti nemzetközi kitekintés is megmutatta – a levéltári intézmények sem viszonyulnak egységesen a kérdéskörhöz. Egyes levéltárak konkrét és rigorózus elvárásokat alakítanak ki az iratképzőknél végzendő digitális fájlformátumok tekintetében, másutt inkább azokat a szempontokat igyekeznek rendszerezni és tudatosítani, melyek alapján az egyes digitalizálási programokban alkalmazandó fájlformátumokat ki kell választani. A hazai gyakorlat és néhány negatív példa ismeretében véleményünk szerint az egyértelműbb állásfoglalás – amennyiben ennek folyamatos felügyelete és szükség szerinti revíziója is biztosított – célravezetőbb az elektronikus iratok hosszú távú értelmezhetőségének fenntartása szempontjából.

Mint fentebb láthattuk, a hazai jogi szabályozás egyelőre nem jutott el oda, hogy az elektronikus iratok hosszú távú megőrzésének és levéltárba adásának feltételeit egyértelműen meghatározza, ezért ajánlásunkban az általánosan elfogadott nemzetközi szabványokra, valamint a hazai és külföldi gyakorlati tapasztalatokra tudunk építeni.

## 5.1 A hosszú távú megőrzésre alkalmas fájlformátumok kiválasztásának szempontjai

Egy elektronikus irat esetében használatos fájlformátum levéltári szempontú alkalmazhatósága számos feltételtől függ. Fontos kiemelni, hogy az archiválási célra alkalmazható fájlformátumok kiválasztása mindenképpen számos kompromisszum megkötését kell eredményezze, hiszen legtöbb esetben a hosszú távú megőrizhetőség és értelmezhetőség kedvéért egyébként rendkívül hasznos, az adatállomány funkcionalitásának lényegi elemeit kell feláldozni. (Az általánosan elterjedt gyakorlat, mely például szöveges állományok vagy táblázatok PDF-formátumban történő archiválását támogatja, a szövegszerkesztő és táblázatkezelő programok által keletkeztetett saját fájlok számos hasznos funkcionalitását zárja ki a megőrzés köréből.<sup>40</sup>)

Általánosságban megállapítható, hogy az egyes fájlformátumok hosszú távon és a lehető legszélesebb körben való alkalmazhatósága a következő feltételek függvénye.

- **Nyitottság**, azaz annak a mértéke, amennyire a digitális tartalmak előállítói és kezelői hozzáférhetnek a formátum teljes körű specifikációjához és eszközeihez, hogy annak technikai integritását megismerhessék. A nyílt szabványok általában szélesebb körű alkalmazhatóságot és technikai támogatást garantálnak egy adott formátumnak. Mindamellet egy szabadalmazott fájlformátum is válhat széles körben elfogadottá és alkalmazottá, bár ezek a korlátok általában akadályozzák annak elterjedését.
- **Elfogadottság**, azaz annak a mértéke, amennyire a digitális tartalmak előállítói és felhasználói körében elterjedt egy adott fájlformátum használata. Minél elterjedtebb egy formátum, annál kisebb az elavulás veszélye és annál valószínűbb, hogy az esetleges elavulással párhuzamosan megszületnek a konvertálás vagy emulálás eszközei, ami jelentős kiadásoktól mentesíti az elektronikus iratok hosszú távú megőrzéséért felelős közgyűjteményeket.
- **Egyszerűség**, azaz annak a szintje, amennyire az adott formátumban tárolt adatok felismerhetővé és értelmezhetővé tehetők egyszerű szoftveres megjelenítő eszközökkel. Az egyszerűség, átláthatóság a szöveges állományok esetében például akkor

<sup>40</sup> Ebből a megfontolásból ered, hogy bizonyos közgyűjtemények nem támogatják a PDF formátum használatát azzal az indokkal, hogy az így keletkezett állományok, például éppen a táblázatos dokumentumok túl sok használható információt veszítenek el egy ilyen konverzióval.  
[http://www.kb.nl/hrd/dd/dd\\_links\\_en\\_publicaties/publicaties/KB\\_file\\_format\\_evaluation\\_method\\_27022008.pdf](http://www.kb.nl/hrd/dd/dd_links_en_publicaties/publicaties/KB_file_format_evaluation_method_27022008.pdf)

érvényesül igazán, ha az adott fájl azokat sztenderd karakterkódolással tárolja (pl. UNICODE). Az egyszerűség követelménye ellen hat mindenféle titkosítás vagy tömörítési eljárás, ezért ezeket a levéltári alkalmazások során lehetőség szerint kerülni kell. Különösen nagy helyigényű digitális tartalmak (bár ezek általában hang- vagy mozgóképállományok esetén valószínűek, de érinthetik a raszteres képek tárolását is) bizonyos mértékű, veszteségmentes tömörítési eljárás elfogadható abban az esetben, ha azok széles körben alkalmazott, nyílt tömörítési módszeren alapulnak.

- **Belső dokumentáltság.** A digitális fájlformátumok közül azoknak a hosszú távú értelmezhetősége és használhatósága valószínűbb, melyek legfontosabb leíró, technikai és adminisztratív metaadataikat saját digitális állományukon belül tárolják, nem pedig attól elkülönítve. Az ilyen állományok későbbi értelmezése, visszafejtése sokkal egyszerűbb és nagymértékben csökkenti az adat használhatatlanná válásának kockázatát. Az az alapelv, hogy a legfontosabb metaadatoknak lehetőleg az elektronikus irat életciklusának a leghamarabbi időszakában, ideális esetben a keletkezéssel egyidőben automatikusan meg kell születnie és az adatállománnyal együtt tárolódnia, a kellő ön-dokumentáltságú fájlformátumok esetében érvényesül leginkább.
- **Szoftver- és hardverfüggetlenség,** vagyis annak a mértéke, amennyire egy konkrét fájlformátum használhatósága a használt hardvereszközöktől, operációs rendszerektől vagy kezelő szoftvektől függ.
- **A szabadalmak hatása.** A szabadalommal korlátozott formátumban keletkező tartalmak általában nehezebben, költségesebben fenntarthatók a levéltárakban még abban az esetben is, ha az adott esetben a licenstdíjak alacsonyak vagy minimálisak. A szabadalom ugyanis általában gátolja a nyílt szabványú alkalmazások fejlesztését és elterjedését, ezáltal gátjává válik a széles körű elterjedésnek. Másfelől viszont egy szabadalmazott formátum széles körű elterjedése és általános használata jele lehet annak, hogy az adott szabadalom kellően rugalmas és hosszabb távon is fenntartható digitális tartalmakat eredményez.
- **A technikai védelmi eszközök hatása.** A digitális tartalmak előállítói a szellemi javak védelme érdekében olyan védelmi eszközökkel láthatják el az adott fájlformátumot, melyek korlátozzák a hozzáférést vagy speciális követelményekhez kötik annak használatát (jelszavas védelem, adott hardvereszközhöz való kötöttség, másolásvédelem stb.). Minden ilyen megoldás akadályozza a hosszú távú értelmezhetőséget és használhatóságot, ezért archiválási célú használatuk feltétlenül kerülendő.
- **Megbízhatóság.** Ez a fogalom önmagában is sokrétű követelményrendszert takar, hiszen a megbízható fájlformátumnak egyszerre kell stabilnak, az adatsérülékenységgel szemben ellenállónak lennie, az egymást követő verziók között „lefelé” és „felfelé” is

nagyfokú kompatibilitással kell rendelkeznie. Mérlegelendő mindezek mellett, hogy az adott fájlformátum mennyire időtálló: milyen gyakran és milyen mértékű változásokkal követik egymást az új verziók stb., mivel ezek a tényezők mind nagymértékben befolyásolják a tartós megőrzés lehetőségét.

Mindezen szempontokat természetesen nem csak új elektronikus dokumentumok létrehozásakor, hanem a korábbi digitális állományok esetleges konvertálása esetén is szükséges érvényesíteni.

Hangsúlyozandó ugyanakkor, hogy a fenti szempontrendszer folyamatos érvényesítése az iratképzők és az iratőrzők állandó feladata, hiszen a technikai változások nyomán folyamatosan új formátumok kerülnek forgalomba, terjednek el, mások pedig elavulnak, támogatásuk megszűnik, használatuk visszaszorul.

## **5.2 Az elektronikus másolatok archiválásra alkalmas formátumai**

A papír alapú iratokról készülő elektronikus másolatok esetében alapelveként kell leszögezni, hogy a digitalizált állomány eredeti, képi formátumának megőrzése a legfontosabb archiválási feladat, mivel ez biztosítja leginkább az elektronikus iratok levéltári értelemben vett hitelességével kapcsolatos követelményeket. A dokumentumok szöveggé alakított formájának tárolása rendkívül hasznos a későbbi használat, a teljes szövegű keresés lehetőségének biztosítása szempontjából, de mindez csak az eredeti képi állomány megőrzésével párhuzamosan fogadható el.

Az iratképző szervek által a papír alapról történő elektronikus másolatkészítés során alkalmazható fájlformátumok esetében különbséget kell tenni a hosszú távú megőrzésre szánt, illetőleg a napi használatban alkalmazható, de archiválásra nem alkalmas formátumok között. Az alábbiakban felsorolt formátumok mindegyikénél jelezzük, hogy ezen funkciók közül melyik betöltésére alkalmas és melyikre nem.

### **5.2.1 Adobe PDF/A-1<sup>41</sup>**

A Portable Document Format (PDF) az Adobe Systems által kifejlesztett bináris fájlformátum. Ebben a formátumban dokumentumok tárolhatók, amelyek különböző szoftverekkel, hardverekkel és operációs rendszerekkel olvashatók és létrehozhatók. A PDF alkalmas szöveget, ábrát és képeket tartalmazó dokumentum leírására eszközfüggetlen és felbontásfüggetlen formában.

---

<sup>41</sup> Lásd részletesebben: <http://www.adobe.com/enterprise/standards/pdfa/>

A PDF egy nyílt szabvány, emiatt bárki jogdíjmentesen írhat alkalmazásokat, amelyek olvasnak vagy írnak PDF formátumot.

A PDF/A-1 formátum a PDF 1.4 verziójának egy változata, melyet kifejezetten a dokumentumok hosszú távú megőrzése érdekében alakítottak ki. Az ISO-szabvánnyá<sup>42</sup> vált formátum hivatalos elnevezése is ezt a szándékot tükrözi: *Electronic document file format for long-term preservation*. A fejlesztő cég a hosszú távú megőrzés érdekében a PDF/A formátum kialakításakor minden olyan funkciót eltávolított az 1.4-es verzióból, melyek az elektronikus iratok hosszú távú biztonságos őrzését és kezelését veszélyeztethetik. Ennek megfelelően ez a formátum megfelel a hosszú távú megőrzés fent sorolt szempontjainak: az elterjedtségen kívül öndokumentáló, egyszerű, nagymértékben eszközfüggetlen is. Nem tartalmaz külső adatokra való hivatkozást (a fontokra, színekre vonatkozó információkat a fájlban belül tárolja), nem rendelkezik jelszavas védelmi lehetőséggel vagy titkosítási módszerrel, nincsenek benne végrehajtható programkódok, Java scriptek vagy multimédiás tartalmak. A fentiekből következően az ilyen típusú fájlok nagyobb helyigényűek a hagyományos pdf-dokumentumoknál, mivel számos, egyébként külső információt a fájlban belül tárolnak.

A PDF/A-1 formátumnak két változata elérhető: a PDF/A-1a, illetve a PDF/A-1b, melyek egymástól alapvetően a szöveges keresési lehetőségekben térnek el. Utóbbi verzió az adott dokumentum formahű megjelentetésének minimális lehetőségeit szolgáltatja, míg az előbbi ennek teljes szövegű tárolását és kereshetőségét is biztosítja. A hosszú távú megőrzés tekintetében mindkét verzió elfogadható levéltári szempontból, ugyanakkor a PDF/A-1a típusú dokumentumok kínálta előnyök a dokumentumok későbbi használatát kétségkívül nagymértékben elősegítik.

A fentiekből következően a digitális másolatkészítés során a dokumentumok tartalmi és formai megőrzésére az Adobe PDF/A-1a típusú fájlformátum ajánlott.

### 5.2.2. Tagged Image File Format<sup>43</sup>

Az egyik legelterjedtebb raszteres képformátum, melyet képek, fotók, és rajzok tárolására fejlesztett ki az Aldus Corporation. Az Adobe Systems 1994-ben megvásárolta az Aldust, így ez a cég szerezte meg és gyakorolja a formátum tulajdonjogát. Jelenlegi verziója a 6.0 számot viseli, mely 1992-ben jelent meg.

A formátum számos tömörítési lehetőséget kínál (RLE, CCITT Group3 és Group4, JPEG, JBIG, LZW stb.) és mindezt tetszőleges színpalettával. Ezeken felül a tárolás

<sup>42</sup> ISO 19005-1:2005, [http://www.iso.org/iso/iso\\_catalogue/catalogue\\_tc/catalogue\\_detail.htm?csnumber=38920](http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=38920)

<sup>43</sup> Részletes specifikációját lásd: <http://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf>

módja, az adatfelépítés és a színmélység is meghatározható (1-től 24-ig terjedően), illetve többcsatornás képek is létrehozhatóak, valamint a formátum lehetővé teszi egy fájlban belüli több kép tárolását is.

A hosszú távú megőrzés szempontjából ugyanakkor a TIFF legfontosabb tulajdonsága az információvesztés nélküli tárolás lehetősége, ami különösen alkalmassá teszi az eredeti dokumentumok hiteles képi formájának eltárolására.

A TIFF képek használata általánosan elterjedt, gyakorlatilag minden képszerkesztő és megjelenítő szoftverrel kompatibilis, 2004. óta nemzetközi ISO szabvány<sup>44</sup> is.

A képfomátum – különösen a tömörítés nélküli verzió - hátránya ugyanakkor a fájlok viszonylag nagy mérete, ami a tömörítetlen TIFF-képeket alkalmatlanná teszi a gyors adatcserére vagy internetes publikálásra. A TIFF-képek méretei korábban komoly digitalizálási programok megvalósítását is veszélyeztették<sup>45</sup>, a tárolókapacitások folyamatos növekedésével azonban ennek a problémának a jelentősége fokozatosan csökken.

A fentiekből adódóan az eredeti iratanyagok digitalizálása során a tömörítetlen, vagy veszteségmentesen tömörített TIFF formátumú, 6.0 verziójú képek használata a hosszú távú megőrzés szempontjából ajánlott.

### 5.2.3 JPEG formátum

A JPEG – a Joint Photographic Experts Group terméke - az egyik legáltalánosabban elterjedt, képek tárolására szolgáló fájlformátum, különös tekintettel az interneten való publikálás és információmegosztás területére. Népszerűségét elsősorban hatékony tömörítési eljárásának köszönheti, melynek révén a tömörítetlen, raszteres képek fájlméretéhez képest akár századrészére csökkenthetők a fájlméretek. A tömörítési eljárás igény szerint paraméterezhető, így kialakítható a szükséges kompromisszum a fájlméret, a minőség és a felhasználhatóság között. A formátum 24 bites színmélységet, 16,7 millió színt képes megjeleníteni. A JPEG formátum alkalmazásának egyes aspektusait több ISO szabvány is rögzítette<sup>46</sup>.

Az eljárás veszteségesen tömörít, ezért nem őrzi meg az eredeti képpontok, s így a tárolt információk összességét.

A papír alapú dokumentumokról készülő elektronikus másolatok készítése során a JPEG formátum használata eredeti (mester) példányok készítése során, illetve

<sup>44</sup> ISO 12639:2004 [http://www.iso.org/iso/iso\\_catalogue/catalogue\\_tc/catalogue\\_detail.htm?csnumber=34342](http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=34342)

<sup>45</sup> Lásd például: Biszak Sándor – Lakatos Andor – Vajk Ádám: Az egyházi anyakönyvek digitalizálásának lehetőségei – módszertani tanulmány. [http://www.leveltaros.hu/akvi\\_tanulmany.pdf](http://www.leveltaros.hu/akvi_tanulmany.pdf)

<sup>46</sup> ISO 10918/1-5, [http://www.iso.org/iso/iso\\_catalogue/catalogue\\_tc/catalogue\\_detail.htm?csnumber=18902](http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=18902)

hosszú távú levéltári megőrzésre a veszteséges tömörítés miatt nem alkalmazható. Az elkészült digitális képek használati példányaként vagy publikációs és adatcsere eszközként rugalmassága és elterjedtsége okán nagymértékben alkalmas.

#### 5.2.4 JPEG2000 formátum<sup>47</sup>

A JPEG2000 formátumot szintén a Joint Photographic Experts Group hozta létre 2000-ben azzal a céllal, hogy felváltsa a JPEG formátumot kiküszöbölve annak több hiányosságát. Fejlettebb kódolási technológiájának köszönhetően az új formátum alacsonyabb minőségromlás mellett is magasabb szinten képes megőrizni a digitalizált oldal részletgazdagságát. A JPEG2000 formátum már képes a fájlon belül tárolni a legfontosabb leíró metaadatokat vagy a szerzői jogokkal kapcsolatos információkat, ugyanakkor a nagyméretű képek gyorsabb megjelenítésére és kezelésére is képes. Igazi előnye jelen témánk szempontjából a JPEG2000 formátumnak abban rejlik, hogy egyaránt képes veszteséges és veszteségmentes tömörítésre is. A JPEG2000 formátum jelenleg már több mint egy évtizedes és elődjéhez hasonlóan 2004-ben ISO nemzetközi szabvánnyá is vált<sup>48</sup>, egyelőre a korábbi verzióhoz képest összehasonlíthatatlanul kevésbé terjedt el a felhasználók körében.

Habár a JPEG2000 formátumhoz sokan fűztek nagy reményeket és nagy jelentőségű levéltári intézmények – például az angol Nemzeti Levéltár - is preferálják a használatát a digitalizálási eljárások során, az elektronikus másolatok hosszú távú megőrzésére jelenleg nem tekinthető támogatott formátumnak. Egyéb területeken használata a JPEG formátumhoz hasonlóan praktikus és elfogadható.

#### 5.2.5 TXT formátum

Az egyszerű, formázás nélküli szövegek tárolására szolgáló formátum a digitális másolás során készült elektronikus iratok területén elsősorban a képként rögzített és a fentiek szerint (PDF/A-1 ill. TIFF) tárolt képek karakterfelismerő programokkal feldolgozott állományainak megőrzésére szolgálhat. Kiemelendő ugyanakkor, hogy a dokumentumok szkennelése során a fent ismertetett PDF-formátumú fájlok használata feleslegessé teszi a szöveges állományok külön elmentését és tárolását.

<sup>47</sup> Lásd ezzel kapcsolatban: <http://www.jpeg.org/jpeg2000/>

<sup>48</sup> ISO-15444-1, [http://www.iso.org/iso/iso\\_catalogue/catalogue\\_tc/catalogue\\_detail.htm?csnumber=37674](http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=37674), ill. ISO-15444-2, <http://www.jpeg.org/metadata/15444-2.PDF>



TXT állományok archiválása esetén csak általánosan elfogadott karakterkódolású (US-ASCII, ISO-8859-1, UNICODE Utf-8, Utf-16) szövegek felelnek meg a hosszú távú megőréssel szemben támasztott követelményeknek.

Ismételten hangsúlyozandó, hogy az elektronikus másolatok készítése során elsődleges jelentősége a képként való tárolásnak van, mivel ez biztosíthatja hosszú távon is az irat levéltári értelemben vett hitelességét. A szöveges állományok párhuzamos készítése és tárolása csak a használatot elősegítő kiegészítő funkcionalitásnak tekinthető, nem pedig a dokumentum levéltári értelemben vett megőrzésének.

## 6. Jogi követelmények

### 6.1 Az elektronikus másolatok hitelessége

Az elektronikus formába másolt (konvertált) iratok hosszú távú megőrzése és alkalmazása szempontjából kiemelt jelentősége van a dokumentum hitelessége fenntartásának. A levéltári, történeti értelemben vett hitelesség és a technikai hitelesítés közti különbségtétel fontosságát fentiekben külön is hangsúlyoztuk. (Lásd ezzel kapcsolatban a 2.1-es fejezetet!) Amint az a jogi áttekintésből is kiderült, az iratok – akár hagyományos, akár elektronikus formáról van szó – hitelességének egy lehetséges garanciáját a Pp. is elismeri azzal, hogy az iratanyagot hivatalból őrző intézmény „hiteléből” vezeti le az általa őrzött iratanyag megbízhatóságát.

A hagyományos iratanyagok esetében alkalmazott általános gyakorlat szerint a levéltár az általa őrzött iratanyagról szükség esetén hitelesített másolatot ad ki, amivel azt igazolja, hogy az iratképzőtől változatlan formában átvett iratanyag az őrizetében van. Természetesen nem vállal felelősséget a közgyűjtemény az iratok tartalmáért vagy az azon szereplő aláírások, pecsétetek vagy egyéb korabeli hitelesítési eszköz eredetiségéért. A hazai bírói gyakorlat a közlevéltárak által a fentiek szerint kiadott hitelesített másolatot bizonyító erejűnek, hitelesnek tekinteni. Ez a hitelesség a levéltár és az iratképző viszonyából, az iratkezelés jogszabályok szerinti levéltári felügyeletéből, az iratátadás részletes dokumentáltságából, valamint a levéltári őrizet kialakult és szabályozott rendjéből fakad.

Az elektronikus iratok a jelenlegi jogi szabályozás keretében ugyanúgy a iratnak tekintendők és a levéltári jogszabályok hatálya alá tartoznak, mint a középkori oklevelek vagy egy 20. századi közigazgatási szerv iratanyaga. Ebből következően a fenti elvek garantálhatják az elektronikus iratok levéltári értelemben vett hitelességét figyelembe véve ugyanakkor ezen iratfajták speciális tulajdonságát is.

Az elektronikus iratok sajátosságainak meghatározása nem tartozik jelen tanulmány körébe, néhány fontos jellemzőjüket ugyanakkor meg kell említeni a hitelesség fenntartásával kapcsolatban.

Az elektronikus iratok legfőbb sajátossága ebben a tekintetben az, hogy a hagyományos iratoktól eltérően megjelenítésükhöz és értelmezésükhöz bonyolult számítástechnikai eszközök szükségeltetnek, melyek ráadásul folytonos változásnak, az elavulás veszélyének vannak kitéve. A hagyományos irat hitelességének egyik garanciája, hogy az adathordozó és a benne foglalt információ szorosan összetartozik, így kölcsönösen hitelesítik egymást. Az elektronikus forma esetében ez a feltétel nem adott, az adathordozó semmilyen formában nem garantálja a hitelességet, hiszen ezen adatok korlátlanul és módosulásmentesen másolhatók egyik adathordozóról a másikra. (Sőt ez nem csak lehetőség, hanem szükségszerűség is, hiszen az adathordozók még az elektronikus iratoknál is jobban ki vannak téve az elavulás, előregedés vagy meghibásodás veszélyének.)

Az elektronikus irat esetében ráadásul az adattartalom fizikai helye is sokkal bonyolultabb, mint a hagyományos iratok esetében, ezért az iratok integritásának, mint a hitelesség egyik legfőbb garanciájának megőrzése különösen nagy hangsúlyt kell kapjon.

## 6.2 A hitelesség biztosításának eszközei

A papíralapú dokumentumokról elektronikus úton történő másolat készítésének szabályairól szóló 13/2005. (X. 27.) IHM rendelet a következőképpen határozza meg az elektronikus másolat megfelelőségének jogi kritériumát:

*„képi megfelelés: az elektronikus másolat azon tulajdonsága, amely biztosítja a papíralapú dokumentum - joghatás kiváltása szempontjából lényeges - tartalmi és formai elemeinek megismerhetőségét”.*

A jogszabály a másolatkészítés megfelelőségi kritériumai között felsorolja még a hitelesítési záradék, a szükséges metaadatok, valamint az elektronikus aláírás meglétét is.

A 6.1 pontban megfogalmazott alapelvek, valamint az elektronikus iratok sajátosságai figyelembevételével ezen iratok hosszú távú megőrzését, illetőleg hitelességét többszintű garanciarendszerrel kell biztosítani.

- Mindenekelőtt a jogszabályalkotónak meg kell szüntetnie azt a szabályozási hiányt illetve ellentmondásokat, melyekre a 2.1 pontban felhívtuk a figyelmet. A Ltv. 35/A § (2) bekezdésében adott felhatalmazás alapján szabályozni kell az elektronikus iratok levéltárba adásának követelményeit.
- Jogszabályi szinten szükséges megerősíteni a levéltárak felelősségét és jogköreit az illetékességükbe tartozó iratképző szervezetek az elektronikus iratok kezelése –

keletkeztetése, használata, megőrzése - során folytatott gyakorlata felett. Mivel az elektronikus iratok biztonságos megőrzése a hagyományos gyakorlattól eltérően annak teljes életciklusa feletti ellenőrzést kell jelentsen (lifecycle records management), ennek jogszabályi feltételeit a vonatkozó jogszabályokban, különösen a Ltv-ben, valamint a szakmai követelményeket megfogalmazó 10/2002 NKÖM rendeletben, vagy ezek szintjén kell megteremteni (lásd még a 2.2 pontot!).

- Az egyes elektronikus dokumentumok hitelességét nem magában a dokumentumban, hanem az azt kialakító és kezelő informatikai környezetben kell biztosítani. Mindezekre kialakult informatikai eszközök állnak rendelkezésre (megfelelő jogosultsági rendszerek kialakítása, a változások naplózása, előző változatok tárolása stb.), melyeket az elektronikus dokumentumkezelő rendszerek tervezése és alkalmazása során fokozott gondossággal kell alkalmazni.
- Az elektronikus dokumentumkezelő rendszerekben olyan szabványos, nemzetközileg is elfogadott formátumokat és megoldásokat kell alkalmazni, melyek a papír alapú iratanyagról készült digitális másolat olvashatóságát, valamint képpontok szerinti megjeleníthetőségét hosszú távon is garantálni képes.
- A hagyományos iratkezelési és levéltári gyakorlathoz hasonlóan ki kell alakítani a maradandó értékű iratanyagok átvételével kapcsolatos eljárásrendet, biztosítani kell annak informatikai garanciáit, hogy a levéltár az átvétel pillanatában lezárt dokumentumok változatlanságáért felelősséget vállalhasson. Megfontolandó eljárás az iratátvétel során az átvételhez rendelendő ellenőrző összeg (checksum) alkalmazása az adatok integritásának biztosítására.
- Az elektronikus iratok őrzéséért felelős közgyűjteményeknek saját eljárásrendjüket és napi gyakorlatukat ki kell egészítsék az elektronikus iratok speciális őrzési és kezelési követelményeit tartalmazó előírásokkal.
- A közgyűjtemények fenntartóinak hosszú távú elkötelezettséget kell vállalniuk az elektronikus iratok megőrzésével járó megnövekedett infrastrukturális és személyi kiadások fedezésére.
- Az elektronikus aláírás, mint hitelesítési eszköz hosszú távú fenntartása véleményünk szerint a jelenlegi technikai körülmények között nem tartható elvárás a közgyűjteményekkel szemben. Az elektronikus iratanyagok történeti hitelességét a hagyományos levéltári alapelvek mentén, a szigorúan ellenőrzött és szabályozott, zárt eljárásrendek kialakítása révén célszerű garantálni.

## 7. Technikai feltételek

### 7.1 A másolattal szemben támasztott követelmények

A digitális másolatok készítésének kulcskérdése a megfelelő minőségű képi állományok előállítása. A megfelelő minőséget ebben az értelemben több tényező alapján, minden digitalizálási program során egyedileg kell meghatározni. Az elvárt képminőséget az alábbi tényezők befolyásolják:

- a másolatkészítés célja:
  - o hosszú távú megőrzés,
  - o publikálás,
  - o széleskörű hozzáférés biztosítása,
  - o az eredeti iratanyag védelme az elhasználódástól, szándékos vagy véletlen rongálástól, lopástól stb.
  - o a napi ügymenet támogatása;
- az eredeti iratanyag jellege, méretei, fizikai állapota;
- a rendelkezésre álló informatikai kapacitás (a digitalizálás eszközei, tárolókapacitás stb.).

A digitális másolatok megfelelő minőségét elsősorban a következő szempontok határozzák meg:

- **Szín.** A másolatok készítése során a színes digitalizálás ajánlott, mert ez még az egyébként fekete-fehér dokumentumok esetében is általában az eredeti képminőség korrektebb reprodukcióját valósítja meg. Nem hosszú távú megőrzésre szánt digitalizálás esetén elfogadható a kéttónusú, fekete-fehér szkennelés. Megjegyzendő ugyanakkor, hogy a színes digitális másolat megőrzi az eredeti dokumentum olyan sajátosságait is (például az eredeti adathordozó színárnyalatát), melyek kéttónusú digitalizálás esetén elvesznek, s amelyek a későbbi felhasználás során jelentős mértékben emelik a másolat használati értékét, növelik az eredetivel való megegyezés, a hitelesség érzetét.
- **Színmélység.** A színmélység vagy bitfelbontás azt mutatja meg, hogy egy képpont színét hány biten tárolja a rendszer, vagyis maximálisan hány szín használható a képen. A nagyobb színmélység több színt, az eredeti kép pontosabb színvisszaadását teszi lehetővé, de egyben a képfájl méretét is növeli. 8 bites színmélység 256, a 16 bites (High Color) színmélység 65,536, a 24 bites (True Color) színmélység 16,777,216 szín visszaadását jelenti. Ebben az esetben is a lehető legpontosabb (True Color) színvisszaadásra kell törekedni, maradandó értékű iratok szkennelése során ez alapkövetelmény.
- **Képfelbontás.** A képfelbontás a képen belüli képpont-távolságot határozza meg. Mértékegysége a képpont/hüvelyk (pixels per inch, ppi). A képfelbontás értelemszerűen

szoros összefüggésben van az elkészült képfájl reprodukciós hitelességével és természetesen a képfájl méreteivel is. A papír alapú iratok digitalizálása során legalább 300 ppi felbontású másolatot kell készíteni a hosszú távon megőrzendő anyagokról. Egyéb célú digitalizálás esetén az adott felhasználási terület (lásd feljebb!) egyedi igényei szerint alakítható a szkennelés felbontása és ezáltal a használati képfájlok mérete.

- **Szintér.** A fenti tulajdonságok mellett az egyes képek hiteles színvisszaadása szempontjából jelentősége van még az ún. szintérnek, vagyis az egyes színek definiálására használatos koordináta-rendszerek. A számítástechnikai képfeldolgozás terén általánosan elterjedt szinterek az RGB és ennek variációi (sRGB, Adobe RGB, ProPhot GB), melyek közül az eredeti anyagok digitális tárolására az AdobeRGB<sup>49</sup> szintér használata javasolt.
- **Tárolási formátum.** Az eltárolt képek minőségét, megbízhatóságát és hosszú távú használhatóságát a fentiekén túl alapvetően befolyásolja még az adott tárolási fájlformátum, melyek használatával kapcsolatban lásd az 5.2 fejezetet!

A papír alapú, illetve pozitív vagy negatív filmről készülő archiválási célú másolatokkal szemben támasztott legfontosabb minőségbeli követelmények a következők.

Dokumentumtípus	Felbontás	Szín	Színmélység	Fájlformátum	Tömörítés
Magas kontrasztú, szöveges vagy grafikai dokumentumok	300 ppi	színes	24 bit	PDF/A-1	tömörítetlen, vagy v.mentesen t. <sup>50</sup>
Színes, vagy alacsony kontrasztú, árnyalatgazdag dokumentumok	Min. 300 ppi	színes	24 bit	PDF/A-1	tömörítetlen, vagy v.mentesen t.
Fekete-fehér fényképek	600 ppi	színes	24 bit	TIFF	tömörítetlen, vagy v.mentesen t.
Színes fényképek	600 ppi	színes	24 bit	TIFF	tömörítetlen, vagy v.mentesen t.
Fekete-fehér negatív vagy pozitív film	2400 ppi	színes	24 bit	TIFF	tömörítetlen
Színes negatív vagy pozitív film	2400 ppi	színes	24 bit	TIFF	tömörítetlen

<sup>49</sup> <http://www.adobe.com/digitalimag/pdfs/AdobeRGB1998.pdf>

<sup>50</sup> veszteségmentesen tömörített

## 7.2 A másolatkészítő rendszerrel szemben támasztott követelmények

- A másolatkészítő rendszernek támogatnia kell a monokróm és a színes szkennelést egyaránt.
- A másolatkészítő rendszernek legalább az alábbi szabványos képformátumba kell tudnia képeket menteni:
  - o TIFF
  - o JPEG
  - o PDF/A
- A másolatkészítő rendszernek képesnek kell lennie különböző felbontásban elkészíteni és elmenteni a képeket. Az egyes beállításokat profilokba kell tudni menteni, így az azonos típusú dokumentumok azonos paraméterekkel való szkennelése garantálható.
- A másolatkészítő rendszernek képesnek kell lennie arra, hogy minimálisan a hagyományos dokumentumméreteket (A4, A3) kezelje.
- A másolatkészítő rendszernek rendelkeznie kell optikai karakterfelismerő (OCR, Optical Character Recognition) funkcióval, de szerencsés, ha az OCR mellett az ICR (Intelligent Character Recognition) lehetőségét is támogatja. (Ez a funkció ugyanakkor nem elsősorban a másolatkészítést, hanem annak utólagos használhatóságát szolgálja.)
- Optikai karakterfelismerés alkalmazása esetén az iratkezelő rendszernek egyetlen iratként kell kezelnie a beszkenelt képet és az abból előállított szöveget.
- Tömeges szkennelés esetén a rendszernek képesnek kell lennie a különálló dokumentumok felismerésére, sorba rendezésére és rögzítésére.
- A másolatkészítő rendszernek lehetőséget kell biztosítania a képek szkennelés utáni ellenőrzésére és szükség szerinti korrekciójára.
- A másolatkészítő rendszernek rendelkezni kell a képek alapvető minőségjavítására (kontrasztarány és fényerő változtatása, gammakorrekció, retusálás stb.) alkalmas szolgáltatással, illetőleg támogatnia kell erre alkalmas külső megoldások alkalmazását.
- A rendszernek lehetőséget kell biztosítani arra, hogy az adatrögzítő megjegyzéseket fűzhessen az elkészült képekhez.
- A rendszernek a szkennelési munkamenetek alábbi részleteit kell automatikusan naplóznia:
  - o felhasználó neve;

- munkaállomás azonosítója;
  - be- és kilépés időpontja;
  - a munkamenet azonosítója;
  - kötegazonosító(k);
  - dokumentumok száma (amennyiben értelmezhető);
  - digitalizált képek száma;
- A rendszernek a következő metaadatokat automatikusan rögzítenie kell:
- a felhasználó azonosítója,
  - a másolást végző szervezeti egység azonosítója,
  - a szkennelés időpontja,
  - az eredeti dokumentum fizikai méretei,
  - a szkennelő eszköz illetve rendszer, valamint a másolatkészítési szabályzat megnevezése.
- A másolatkészítő rendszernek képesnek kell lennie a beszkennelt állományok bélyegképeit megjeleníteni a böngészés és keresés megkönnyítése érdekében.
- A másolatkészítő rendszernek együtt kell működnie az intézmény dokumentumkezelő rendszerével, amennyiben ilyen létezik.
- A rendszernek garantálnia kell, hogy az elektronikus másolat és annak papír alapú eredetije közti kapcsolat ténye a metaadatokban dokumentálódjon.
- A másolatkészítő rendszernek tudnia kell exportálni az iratokat metaadataikkal együtt kibocsátott információs csomag (Submission Information Package, SIP) formájában az OAIS szabvány szerint<sup>51</sup>.
- A másolatkészítő rendszernek az intézmény adatvédelmi politikájával összhangban garantálnia kell a rendszerbe bekerült személyes vagy minősített, illetve egyéb okból érzékeny adatok kellő védelmét.
- A digitális másolatok készítésére használatos eszközök nem okozhatnak helyrehozhatatlan károsodást az eredeti iratokban, különös tekintettel az archiválási céllal digitalizált, maradandó értékű iratanyagokra.
- Automatikus laptovábbító eljárások csak abban az esetben engedélyezhetők, ha azt az eredeti iratok fizikai állapota megengedi. A papír alapú dokumentumok nem kockáztathatók a tömeges másolatkészítés érdekében.

---

<sup>51</sup> ISO 14721:2003 [http://www.iso.org/iso/catalogue\\_detail.htm?csnumber=24683](http://www.iso.org/iso/catalogue_detail.htm?csnumber=24683)

- A másolatkészítő rendszernek lehetővé kell tennie a jogosultságok tartalmának differenciált meghatározását az egyes munkafolyamatokra lebontva.
- A másolatkészítő rendszer meg kell feleljen a 24/2006 BM-IHM-NKÖM együttes rendeletben megfogalmazott biztonsági, valamint a 2011. évi CXII. tv., illetve az 1998. évi VI. tv. adatvédelmi előírásainak.

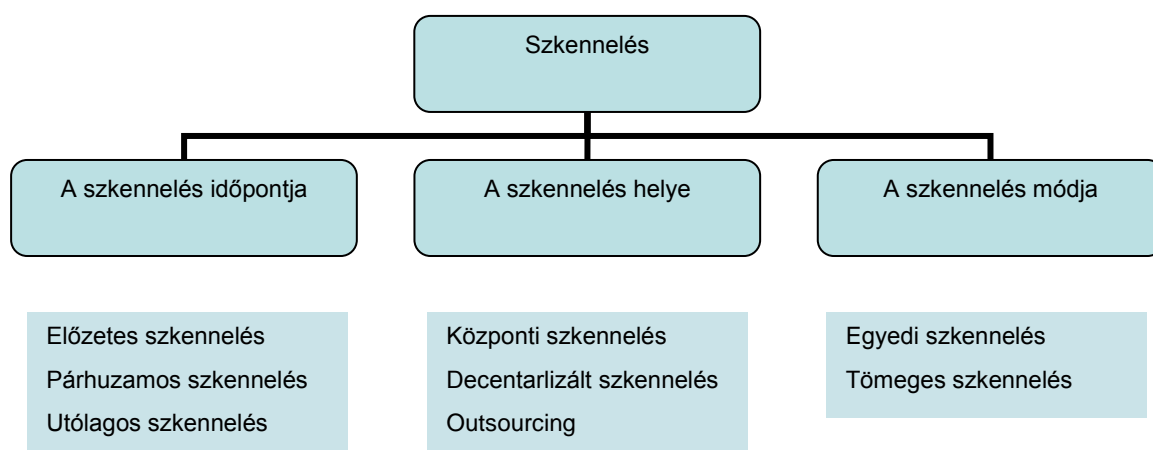
## 8. A digitális másolatkészítés folyamata

### 8.1 Általános szempontok

Az iratképző szerveknél a papír alapú másolatok készítésének alapvetően két fő célja van. (A levéltári elektronikus másolatkészítés kérdéskörét nem tekintjük jelen témánkba tartozónak.)

- Már papír alapon vagy mikrofilmen létező és őrzött iratanyag digitális másolatának elkészítése annak érdekében, hogy az elektronikus forma használható legyen egy dokumentumkezelő rendszerben (retrospektív digitalizálás).
- A papír alapon érkező vagy ilyen formában keletkező iratok elektronikus másolatainak szisztematikus elkészítése elektronikus iratkezelés céljából (folyamatos digitalizálás).

A fenti alapvető célok nagymértékben kihatnak a digitalizálás szervezeti kereteinek kialakítására és munkafolyamat-követelményeire.



#### A szkennelés lehetséges megoldásai

Fontos előfeltétele a digitális másolatok készítésének, hogy az eredeti, papír alapú iratok a hagyományos iratkezelési szabályok szerint rendezettek és rendszerezettek legyenek.



### 8.1.1 A szkennelés időpontja

A digitalizálási folyamat helye a teljes iratkezelési rendszeren belül különböző lehet

- **Előzetes szkennelés**, amikor a digitalizálás már a papír alapú dokumentumok beérkezésekor megtörténik a központi postabontóban vagy decentralizáltan az adott iratkezelési rendszertől függően. Az előzetes szkennelés céljai a következők lehetnek:
  - o az információk minél gyorsabb és minél szélesebb körű megosztásának lehetősége,
  - o az elektronikus továbbítás révén elérhető hatékonyságnövekedés,
  - o a beérkező iratok minél gyorsabb eljuttatása az ügyintézőhöz,
  - o az elektronikus ügyirat irat minél korábbi szakaszban való előállítása,
  - o a teljes ügyintézési szakasz gyorsítása.

Az előzetes szkennelés nagyon hatékony eszköze az elektronikus iratkezelésnek abban az esetben, ha a rendelkezésre álló eszközbeli és emberi kapacitások lehetővé teszik a szükséges munkafolyamatok gyors elvégzését. Különösen igaz ez az elektronikus formába konvertált irathoz szükséges metaadatok rögzítésére, ami rendkívül időigényes feladat, így könnyen válhat ez a munkafolyamat az egész elektronikus iratkezelés szűk keresztmetszetévé.

- **Párhuzamos szkennelés**, amikor nem előzetesen történik a digitalizálás, hanem ügyintéző, vagy az ezért felelős szervezeti egység adminisztrátora az ügyintézés megkezdésével egyidőben végzi el az elektronikus másolást. E megoldásnak előnye, hogy csak azoknak az anyagoknak a másolása történik meg, melyekre ténylegesen szükség van, ráadásul az elektronikus másolatok legfontosabb leíró metaadatai is ezen a szinten adhatók meg legprecízebben. A módszer hátránya ugyanakkor, hogy nem biztosít gyors és széleskörű hozzáférést a dokumentumokhoz, ezenkívül a másoláshoz szükséges technikai tudás „decentralizációjára” is szükség van, az egyedileg telepített másolatkészítő rendszerek pedig általában kevésbé költséghatékonyak, mint a központosított megoldások. Mindezekből következően ez a megoldás inkább kisebb szervezetek, intézmények számára ajánlható, ahol nem jellemző a bonyolult döntéshozatali folyamat és nincs szükség széleskörű hozzáférésre az egyes elektronikus dokumentumokhoz.
- **Utólagos szkennelés**, amikor az ügyirat, vagy egyes részeinek szkennelése csak a hagyományos iratkezelési procedúra lezárultával történik meg akár központilag, akár decentralizáltan. Ez a megoldás jelenti a legkisebb beavatkozást a napi iratkezelési

munkafolyamatokba, egyúttal azonban ebből származik a legkisebb haszon is annak támogatása terén. Az utólagos szkennelés akkor jelent megtérülő beruházást, ha nagy mennyiségű a lezárt ügyek iránti igény, érdeklődés.

### 8.1.2 A szkennelés helye

A szkennelés helyét tekintve alapvetően három megoldás létezik, bár ezek kombinációja is megvalósítható.

- **Központi szkennelés**, mely alapvetően nagy mennyiségben végzett digitalizálásra alkalmas. Tipikusan nagy teljesítményű, így költséghatékony eszközökkel dolgozik és könnyebben biztosítható a szkenneléshez szükséges szakértők jelenléte.
- **Decentralizált szkennelés**, melyet elsősorban kisebb vagy területileg széttagolt szervezeteknél célszerű alkalmazni, illetőleg amikor a digitalizáláshoz szükséges döntéshez a konkrét ügy ismerete szükséges. Ebben az esetben általában kisebb teljesítményű, lassabb és rossz költséghatékonyágú eszközökkel történik a digitális másolatok készítése.
- **Outsourcing**, amikor a digitalizálást és az ezzel kapcsolatos munkafolyamatokat részben vagy teljes egészében külső vállalkozás végzi. Ezt a megoldást azokban az esetekben célszerű alkalmazni (azokban az esetekben költséghatékony), amikor
  - o nagy mennyiségű, korábban már lezárt ügyiratot kell szkennelni,
  - o amikor az iratképző szervnél nem áll rendelkezésre a szükséges személyi erőforrás,
  - o amikor az iratképző szervnél nem áll rendelkezésre megfelelő eszközállomány vagy nincs alkalmas hely a digitalizálás végzésére,
  - o amikor a szkennelés és/vagy a tárolás nem helyszíntől függő.

### 8.1.3 A szkennelés módja

- **Egyedi szkennelés** esetén a digitális másolást egyedi dokumentumonként, síkágyas szkennerekkel végzik általában helyi szinten. Hatékonysága meglehetősen alacsony, ugyanakkor sokkal kisebb előkészítő munkát igényel és a folyamatos emberi részvétel miatt alacsonyabb lehet a hibások végzett digitalizálás aránya is.
- **Tömeges szkennelés általában** centralizáltan, automatikus adagolású dokumentumskennerekkel történik. Ebben az esetben értelemszerűen növekszik az egységnyi idő alatt elvégezhető digitalizálások mennyisége, ugyanakkor rendkívül

komoly előkészítő munkát igényel a papír alapú eredeti anyagok előkészítése, ezenkívül komoly erőforrásra van szükség az elkészült digitális másolatok minőségellenőrzéséhez, esetleges korrekciójához, újraszkeneléséhez stb.

#### **8.1.4 A digitális másolatkészítés eszközei**

A digitális másolat készítéséhez több – a fenti megoldások függvényében alkalmazandó – eszköz illetve feltétel szükséges. Ezek a következők:

- digitális szkennelő (a 8.1.2 ill. 8.1.3 fejezetben leírtak függvényében alkalmazható síkágyas szkennelő, dokumentumszkennelő, könyvszkennelő, rollszkennelő, digitális fényképezőgép vagy mikrofilm-szkennelő),
- szkennelő és képfeldolgozó szoftver,
- elektronikus dokumentumkezelő rendszer,
- elektronikus tárolóeszköz, szerver számítógép.

#### **8.1.5 Dokumentációs követelmények**

Az elektronikus iratok hitelességének egyik kulcsa a szabályozott és alkalmazott eljárásrend, melyben azokat előállítják, kezelik és tárolják. A megbízható digitális másolatokat eredményező digitalizálási programok minimális dokumentációs követelményrendszere a következő:

- Az illetékes levéltár által jóváhagyott Iratkezelési Szabályzat és Irattári Terv,
- Informatikai Szabályzat,
- Adatvédelmi és Adatbiztonsági Szabályzat,
- Másolatkészítési Szabályzat, mely tartalmazza a digitális másolatok készítésére vonatkozó technikai követelményrendszert is,
- a másolatkészítő rendszer részletes technikai specifikációja,
- az elektronikus iratok leírására szolgáló metaadat-séma és kitöltési útmutató.

#### **8.1.6 Személyi feltételek**

A digitális másolatok készítésének személyi feltételei alapvetően a következőktől függenek:

- a digitalizálandó iratanyag mennyisége,
- a munkafolyamat centralizált vagy decentralizált jellege,

- a választott szervezeti megoldás (saját erőből, ill. részben vagy egészben külső vállalkozóval végzett digitalizálás).

Bármilyen szervezeti megoldást választ az intézmény, a digitalizálási projekt következő szintjeit feltétlenül saját munkatársaival kell végeznie:

- projektmenedzsment,
- az eredeti anyagok előkészítése, illetve ennek felügyelete,
- az elkészült képek minőségi ellenőrzése.

A digitális másolatkészítési folyamat során az alábbi minimális személyi feltételeket meg kell teremteni.

- Legalább egy felsőfokú végzettségű informatikai szakember rendszeres jelenléte és a digitalizálási folyamat állandó informatikai felügyelete.
- Iratkezelő/irattáros végzettségű (54 322 02 0100 52 01 OKJ) munkatárs részvétele az elektronikus másolatkészítés folyamatában, ennek hiányában a digitális másolatkészítés és metaadat-kezelés speciális feladataival kapcsolatos oktatáson részt vett, legalább középfokú végzettségű munkatárs alkalmazása.
- A másolatkészítés bármely szakaszában részt vevő munkatársak egyedi, a másolatkészítés technikai részleteit, valamint a metaadatok kezelését is magában foglaló továbbképzése.
- Az intézmény legalább egy saját dolgozójának megbízása az elkészült digitális másolatok megfelelőségének ellenőrzésével.

## **8.2 A másolatkészítés folyamata**

### **8.2.1 A digitalizálandó anyagok kiválasztása**

A digitális másolatkészítés rendszere soha nem működhet önmagában, hanem mindig csak szigorú összhangban az intézmény iratkezelési gyakorlatával. Az elektronikus másolatok ennek az iratkezelési rendszernek a tárgyai lesznek, így alapvetően ennek a szabályai határozzák meg a digitalizálandó iratok kiválasztásának szempontjait is.

Az iratképző szervekhez érkező vagy ott keletkező, papír alapú dokumentumok nem digitalizálhatók válogatás nélkül, hiszen ez aránytalanul megterhelné a dokumentumkezelő rendszert és az azt működtető személyzetet is. A digitális másolásra való kiválasztás legfontosabb szempontjai a következők:

- az irat megőrzési ideje az Irattári Terv alapján, hiszen rövid időn belül selejtezhető és a napi ügyvitelben átmenetileg sem nélkülözhetetlen dokumentumok digitalizálása felesleges;
- a digitalizálandó iratanyag iránt megmutató hivatali vagy társadalmi érdeklődés;
- saját iratanyag digitalizálása esetén csak elkészült, befejezett, utólag már nem módosuló dokumentumok digitalizálhatók;
- kiadványok vagy nyilvánosan megjelent dokumentumok (pl. könyvek, újságok, jogszabályok), illetve nyilvánvaló duplumok digitalizálása indokolatlan, amennyiben ezek eredetije rendelkezésre áll;
- az eredeti iratanyag méretei és adathordozója szintén érdemben befolyásolják a digitalizálási lehetőségeket, hiszen speciális méretű vagy típusú adathordozók (térképek, tervrajzok, nagyméretű lapok, szorosan bekötött dokumentumok) speciális digitalizálási módszereket, eszközöket és szaktudást igényelnek;
- különösen fontos kiválasztási szempont az eredeti iratok fizikai állapota, hiszen például töredezett, szakadt, vagy egyéb okból sérülékeny iratok digitalizálása egyáltalán nem, vagy csak speciális eszközök igénybevételével, szakszerű előkészületek után történhet meg.

### **8.2.2 Az iratok előkészítése a digitális másolásra**

A papír alapú eredeti dokumentumokat szakszerű előkészítése a digitalizálási program egyik kulcsa, hiszen ennek révén kerülhető el a hibás, használhatatlan elektronikus állományok létrejötte, illetőleg az eredeti anyagok károsodása.

- A szkennelés megkezdése előtt minden idegen irodaszert (kapcsok, elválasztólapok, műanyag borítók, madzagok stb.) el kell távolítani a dokumentumokról. Nem vonatkozik ez a közjegyzői vagy egyéb hitelesítéssel ellátott iratok összefűzésére.
- Ki kell emelni az iratokból azokat az elemeket, melyeket nem szükséges szkennelni (lásd a 8.2.1 pontot!), illetőleg amelyeket saját kezű felbontásra szántak.
- Speciális méretű (pl. A3-as méretnél nagyobb vagy nagyon kicsi), egyedi formátumú (pl. bekötött) vagy egyedi adathordozójú (pl. mikrofilmen tárolt) iratok esetén a rendelkezésre álló szkennelő berendezés lehetőségei függvényében egyedi döntést kell hozni a másolás manuális elvégezhetőségéről. Ilyen esetekben egyaránt alkalmazható rollskenner, könyvskenner, mikrofilmskenner vagy digitális fényképezőgép, amennyiben az általa készített digitális másolat az elkészült képekkel szemben támasztott, 7.1 pontban részletezett követelményeknek megfelel.

- El kell távolítani a műanyag vagy karton elválasztólapokat is. Amennyiben ezek érdemi információkat tartalmaznak, ezek fénymásolatát kell a digitalizálandó anyagok között elhelyezni. Ugyanez vonatkozik az iratokra ragasztott jegyzetlapokra (post-it) is.
- Az iratok fizikai állapota alapján – szükség esetén szakemberek bevonásával - egyedi döntést kell hozni a szkennelés kockázatairól. A nem szkennelhető anyagokat ki kell emelni és szükség esetén állományvédelmi intézkedéseket kell kezdeményezni.
- Szintén egyedi döntést kell hozni a szkennelés megkezdése előtt az adott irat történeti, művészeti stb. értéke alapján (pl. ha az iratok fényképeket, grafikákat, különösen fontos kéziratokat stb. tartalmaznak).
- Fényképeket, grafikákat stb. célszerű levéltári műanyag borítóban elhelyezni digitalizálás előtt.
- Automatikus laptovábbító eszközök használata esetén különös gondossággal kell mérlegelni a begyűrdés, szakadás kockázatait.
- Borítékban lévő iratokat ki kell emelni a borítékból. Amennyiben a boríték érdemi információt tartalmaz, azt üresen kell a digitalizálandó anyag elé helyezni, vagy annak fénymásolatával helyettesíteni.
- Amennyiben az iratok speciális adathordozókat (CD, DVD stb.) is tartalmaznak, annak további kezelése kapcsán egyedi döntést kell hozni az iratkezelésért felelős szakember(-ek) bevonásával.
- Automatikus adagolóval történő digitalizálás előtt az iratokat le kell számozni, hogy az ellenőrzés során megállapítható legyen az esetleges hiány (összeragadt lapok stb.).
- Tömeges szkennelés esetén - az adott szkennelő szoftver működésétől függően – üres vagy vonalkóddal ellátott elválasztólapokat kell az iratok közé helyezni.

### **8.2.3 Másolás és indexelés**

A fentiek szerint megfelelően előkészített iratokat a digitalizálási megoldástól függően egyedileg vagy automatizáltan beolvassa a szkennel.

A folyamat során, a digitalizáló rendszer képességei függvényében elhelyezhető kiegészítő információ is a képen (pl. vízjel, lapszám, dátum stb.), amennyiben az az eredeti anyag olvashatóságát nem zavarja.

A másolatkészítés során létrejött digitális képen utólag végrehajthatók képjavítási eljárások (pl. kontraszt-növelés, világosítás/sötétítés stb.) a következő feltételekkel:

- a képjavítás kizárólag az eredeti anyag hitelesebb képi visszaadása, illetőleg jobb olvashatósága érdekében történhet,
- a módosítások nem járhatnak az eredeti dokumentumokban szereplő érdemi adatok elvesztésével, illetőleg azon nem szerepelt információk megjelenítésével.
- a képjavítás csak dokumentáltan, utólag ellenőrizhető módon történhet.

A digitális másolatkészítő rendszernek a szkennelt iratok bizonyos metaadatait automatikusan rögzítenie kell (a felhasználó azonosítója, a másolást végző szervezeti egység azonosítója, a szkennelés időpontja, az eredeti dokumentum fizikai méretei, a szkennelő eszköz illetve rendszer, valamint a másolatkészítési szabályzat megnevezése).

Minél több további metaadatot rögzít a rendszer, annál kevesebb további manuális beavatkozásra van szükség és annál hatékonyabb lesz az elkészült digitális állományok használata.

A metaadatok automatikus rögzítésének kulcsa az iratok strukturáltsága és homogén jellege: minél egységesebb az iratanyag (pl. azonos formátumú kérdőívek, formanyomtatványok), annál hatékonyabb az automatikus indexelés.

A digitalizálási program első fázisában meg kell határozni azt a legmélyebb szintet, melyen az adott iratot keresni fogják, és ennek megfelelően kell kialakítani az ehhez szükséges metaadat-szerkezetet. Ebben a szakaszban döntendő el az is, hogy a rendszer mely metaadatokat tároljon a fájlban belül és melyeket adatbázisban.

Az automatikus indexelés soha nem helyettesítheti teljes mértékben az iratkezeléssel foglalkozó munkatársak manuális adatrögzítési munkáját.

Az elektronikus másolatok metaadatokkal való ellátása több szinten történhet a választott szervezeti megoldástól függően. Hangsúlyozandó, hogy az elkészült elektronikus másolat egy ügyiratnak csak része, így a legfontosabb metaadatokat az iratkezelés helyi gyakorlatától függően az iratkezelés adott mozzanatát végző személyek rögzítik.

Digitális másolatkészítés során a másolatkészítéssel kapcsolatos legfontosabb technikai metaadatokat a központi szkennelő szolgáltatás automatikusan, vagy annak működtetői manuálisan rögzítik.

Az iratanyag leíró és adminisztratív metaadatait az Iratkezelési Szabályzatban meghatározott, az iratkezelés megfelelő szintjéért felelős személyek rögzítik.

A digitalizálás megtörténtével párhuzamosan, vagyis az elektronikus másolat életciklusának elején garantálni kell a másolat és eredetije nyilvántartási adatainak összetartozását. Ez leginkább úgy biztosítható, ha a másolat egy elektronikus ügyirat része lesz, mely a papír alapú irat nyilvántartási és egyéb metaadatait is tartalmazza.

### 8.2.4 A másolatkészítés során rögzítendő metaadatok

Hangsúlyozandó, hogy az elektronikus iratok általános metaadat-szerkezetének definiálása nem jelen tanulmány témája, alább szigorúan csak a másolatkészítéssel kapcsolatban rögzítendő metaadatok köréről lesz szó. Ezeket a metaadatokat a fájlon belül, vagy a teljes ügyirat metaadat-strukturájának technikai metaadatai között kell megjeleníteni.

Típus	Metaadat-elem	Kötelező/Opcionális
A kép tulajdonságaira vonatkozó metaadatok	Színmélység (bit)	K
	Tömörítés	K
	Felbontás (ppi)	K
	Szintér	K
	A kép szélessége (cm)	K
	A kép magassága (cm)	K
	A kép orientációja	O
A kép készítésével kapcsolatos metaadatok	Az eredeti dokumentum jellemzőinek leírása	O
	A szkenneléshez (digitális fényképezéshez) használt eszköznek és gyártójának a pontos megnevezése	K
	A képfeldolgozáshoz használt szoftver és verziójának megnevezése	K
	A képfeldolgozás során alkalmazott képjavítási eljárások megnevezése	K
	A másolatkészítési szabályzat és verziójának megnevezése	K
	A másolatot készítő felhasználó	K



	A másolatot készítő szervezeti egység	K
	A másolat készítésének időpontja	K
	Érvényességi idő, vagy annak jelzése, hogy az érvényességi idő nem meghatározott	K

Azon digitális másolatok esetében, melyeket nem dokumentumkezelő rendszerben, hanem fájlrendszerben tárolnak, a fenti metaadatokon túl egyedi nyilvántartásban rögzíteni, vagy abba importálni kell az elektronikus ügyirat kötelező metaadat-elemeit, illetőleg egyértelműen meg kell teremteni a kapcsolatot a létező nyilvántartás és az abban tárolt metaadatok, valamint az ahhoz tartozó digitális képfájl között.

#### 8.2.5. A digitális képfájlok elnevezésével kapcsolatos megoldások

A számítógépes rendszer önmagában nem képes a későbbi azonosítást vagy rendszerezést segítő fájlneveket adni az elkészült digitális állományoknak. A fájlok elnevezésének többféle gyakorlat is kialakult:

- olyan számrendszer alkalmazása a névadásban, mely az elektronikus nyilvántartási rendszerben használatos számozást – vagy azt is – tükrözi,
- a kép készítésével (szkennelő neve, szkennelés dátuma stb.) kapcsolatos adatokat megjelenítő fájlnevek használata,
- a fenti megoldások kombinációja.

A fenti megoldások közül a nyilvántartási rendszerben alkalmazott számozásra utaló fájlnévadás alkalmasabbnak tűnik arra, hogy további biztosítékot jelentsen az eredeti dokumentum és digitális másolata összetartozásának fennmaradására. A fájlnevezési konvenciónak ezek szerint a dokumentum provenienciáját és a regisztratúrában elfoglalt helyét kell kifejeznie. A fájlok elnevezési szabályait a Másolatkészítési Szabályzatban részletesen rögzíteni és azt a gyakorlatban alkalmazni kell, a dokumentumkezelő rendszer pedig lehetőleg ki kell kényszerítenie a fájlnevezési konvenciók betartását.

## 9. A digitális másolat elkészülte utáni munkafolyamatok

### 9.1 Az elkészült képfájlok OCR-es feldolgozása, az ICR, IWR alkalmazása

Az OCR technológia révén az elmentett és raszteres képként tárolt állományok szöveges dokumentumokká alakíthatók. Figyelembe kell venni ugyanakkor, hogy ez a technológia még nyomtatott szövegek esetében sem szolgáltat 100 %-os felismerési arányt az eltérő dokumentumtípusok, az azokon használt különböző fonttípusok és betűméretek, a szkennerek és az OCR-szoftver egyedi beállításai miatt.

Figyelembe kell venni, hogy ez a technológia csak bizonyos feltételek teljesülése esetén vezet a kívánt eredményre. Mindenekelőtt el kell döntenie a következőket:

- a dokumentum teljes szövegét fel kívánjuk dolgozni a későbbi teljes szövegű keresés lehetőségének biztosítására, vagy
- csak papír alapú ill. mikrofilmen tárolt iratok egy előre definiált, jól meghatározott részterületét.

Ez utóbbi megoldás akkor lehet különösen hatékony, ha egységes formájú dokumentumokkal dolgozunk, melyek meghatározott helyen meghatározott struktúrájú információkat tartalmaznak. Ez történhet formanyomtatványok, kérdőívek, egységes formátumú számlák stb. feldolgozása során. Ebben az esetben pontosan definiálni kell azt a területet az iraton belül, melyet fel kívánunk dolgozni, és meg kell határozni a még elfogadható hibaarány tűréshatárát. Mindezek révén kellő mértékben javítani lehet az automatikus indexelés és feldolgozás hatékonyságát.

Amennyiben nem elvárás az eredeti információtartalom hibátlan reprodukálása, a teljes szövegű OCR-es feldolgozás is használható eljárás. Ebben az esetben a feldolgozott szöveg lehetőséget nyújt a teljes szövegű keresésre, ami semmiképpen sem helyettesítheti az elektronikus másolat metaadatokkal való szakszerű ellátását, leírását.

Az utóbbi esetben megoldást jelenthet az is – amennyiben azt irat fontossága megköveteli és a rendelkezésre álló munkaerő megengedi -, hogy a képként tárolt iratot a feldolgozott szöveggel összevetve korrigáljuk a hibásan feldolgozott szövegrészeket. Ez utóbbi meglehetősen nagy időráfordítást igénylő eljárás, így alkalmazása csak különösen indokolt esetekben javasolható.

Az újabb OCR-alkalmazások már taníthatóak, vagyis egyes pontthalmazok értelmezésében beavatkozási lehetőséget kínálnak a felhasználóknak. Ez nagymértékben növeli a felismerés hatékonyságát, de elsősorban hasonló eszközökkel és formában készült anyagok esetén működőképes eljárás.

Általánosságban megállapítható, hogy az OCR-technológia olyan iratok esetében vezet kielégítő eredményre, melyek a lehető legkisebb arányban tartalmaznak kézírást és melyek egységes formátumban, hasonló adatstruktúrában készültek.

Az OCR-technológia továbbfejlesztett változata, mely a szöveg kontextusának értelmezésére, kézírásos nyomtatott betűk felismerésére stb. szolgál az ICR (Intelligent Character Recognition) nevet viseli. Jelenleg ez az eljárás nem működik olyan hatékonysággal, hogy az elektronikus képfeldolgozás során – szigorúan behatárolt adatbeviteli sémák (pl. csak kézzel írt számokat tartalmazó adatlapok) alkalmazásától eltekintve – széles körben alkalmazni lehetne. Még inkább érvényes ez a folyó írások értelmezését (is) célul tűző IWR (Intelligent Word Recognition) eljárás esetében, mely egyelőre csak a jövő reménybeli technológiájának tekinthető.

## 9.2 Minőségi ellenőrzés

A digitális másolatok, valamint az azt létrehozó rendszer(-ek) minőségének folyamatos ellenőrzése a következőket foglalja magában

- a digitalizáló eszköz rendszeres ellenőrzését és tisztítását,
- a szkennerek és a monitorok rendszeres kalibrálását,
- az elkészült képállományok minőségi ellenőrzését.

Az ennél részletesebb követelményrendszer kialakítása az adott másolatkészítő rendszer (szkennerek, szoftverek, dokumentumkezelő rendszer) függvénye.

Az elkészült képek ellenőrzése során a képfeldolgozás, rendszerezés vagy indexelés eredményeit is folyamatosan ellenőrizni kell.

Tömeges szkennelés esetén véletlenszerű, 5-10%-os mintavétel alapján célszerű végezni az ellenőrzést.

A minőségi ellenőrzés magában kell foglalja:

- az összes kijelölt oldal szkennelésének megtörténtét (a papír alapú oldalszámok és a digitális képfájlok számának összevetésével)
- az eredeti irat legapróbb információi rögzítésének ellenőrzését,
- a digitális kép helyes orientációjának vizsgálatát ,
- a szkennerek által okozott esetleges képhibák létének ellenőrzését,
- a teljes képterület rögzítettségének vizsgálatát,
- színes szkennelés esetén a megfelelő színvisszaadás ellenőrzését.

### **9.3 A digitális képfájlok és szöveges állományok további kezelése**

Az elkészült és ellenőrzött digitális képi és szöveges állományokat ezután az elektronikus dokumentumkezelő rendszer tárolja és kezeli.

Az intézmény Iratkezelési Szabályzatának részletes szabályokat kell tartalmaznia az elektronikus iratok további kezelését és megőrzését illetően. Általánosságban megállapítható, hogy a dokumentumkezelő rendszerbe került elektronikus másolatok őrzési ideje meg kell egyezzen az eredeti papír alapú dokumentuméval.

Az elektronikus iratok selejtezése azoknak az elektronikus dokumentumkezelő rendszerből – az Irattári Tervben előírt határidő lejárta és levéltári jóváhagyás után – történő törlésével valósítható meg. Az elektronikus formában vezetett iktatásból a kiselejtezett iratok adatai nem törölhetők.

### **9.4 Az eredeti iratok további kezelése**

Az eredeti iratok további keletkezését az intézményi Iratkezelési Szabályzat és Irattári Terv határozza meg. Három alapvető megoldás lehetséges ezen a téren az iratanyag értéke, valamint az intézmény iratkezelési gyakorlata függvényében:

- az eredeti anyag a másolás után kikerül az ügymenetből és azt irattározzák, vagyis a másolat teljes körűen helyettesíti az eredeti anyagot,
- az eredeti iratanyag végigkíséri az ügyintézés folyamatát, ebben az esetben az elektronikus másolat nem helyettesíti, csak kiegészíti az eredetit,
- a papír alapú dokumentumot másolás után megsemmisítik, amennyiben ezt az iratkezelési szabályok és az irat értéke megengedi (ez utóbbi esetben is célszerű két-három hónapos „türelmi időt” hagyni az esetleges szkennelési hibák kijavítására).

## Függelék

### 1. sz. függelék A megbízható elektronikus másolattal szemben támasztott követelmények

#### 1. Bevezetés

A digitális másolatok készítésének alábbi kritériumai azt a célt szolgálják, hogy garantálják az elkészült elektronikus másolatok és az eredeti iratok képi valamint tartalmi megfelelését, az elektronikus másolatok hosszú távú használhatóságát, értelmezhetőségét és esetleges változásainak nyomon követhetőségét.

Az alábbi követelmények csak néhány, a nemzetközi és hazai szabványok, ajánlások és gyakorlat szerint megbízható, stabil, rendszerfüggetlen és hosszú távon fenntartható fájlformátum használatát engedik meg annak érdekében, hogy az elektronikus iratkezelés egységes alapon történjék és a későbbiekben a levéltárak is egységesen tudják kezelni a hosszú távon megőrzendő iratokat.

Hasonlóan szigorú követelményeket kell érvényesíteni a másolatok egyéb technikai követelményei (felbontás, szín stb.), valamint a későbbi feldolgozás és visszakereshetőség alapját képező metaadatok rögzítése tekintetében is.

#### 2. Dokumentációs követelmények

Digitális másolatkészítési folyamat alkalmazásának minimális dokumentációs feltételei a következők:

- Az illetékes levéltár által jóváhagyott Iratkezelési Szabályzat és Irattári Terv.
- Informatikai Szabályzat, mely meghatározza a digitális másolatkészítési folyamat, valamint az elkészült digitális állományok helyét az intézmény informatikai rendszerében.
- Adatvédelmi és Adatbiztonsági Szabályzat, mely az intézmény adatvédelmi politikájának keretében meghatározza a másolatkészítés során kezelt vagy keletkező, védendő személyes adatok sorsát.
- A 13/2005 IHM rendelet szerinti Másolatkészítési Szabályzat. Ez részletezi a másolatkészítési folyamat szervezeti kereteit, személyi feltételeit, illetve magában foglalja a digitális másolatok készítésére vonatkozó technikai követelményrendszert, mely a fenti szabályok alapján részletesen meghatározza az egyes dokumentumtípusok másolása során alkalmazandó követelményeket.

- A másolatkészítő rendszer részletes technikai specifikációja, beleértve a szkennerek, a képfeldolgozó szoftver, a dokumentumkezelő rendszer (ha van ilyen), valamint az elkészült digitális másolatok tárolására és használatára szolgáló eszközök leírását.
- Az elektronikus iratok leírására szolgáló metaadat-séma, illetőleg az egyes metaadat-elemek kitöltésének részletes előírásait tartalmazó utasítás.

### **3. Személyi feltételek**

A digitális másolatkészítési folyamat során az alábbi minimális személyi feltételeket meg kell teremteni.

- Legalább egy felsőfokú végzettségű informatikai szakember rendszeres jelenléte és a digitalizálási folyamat állandó informatikai felügyelete.
- Iratkezelő/írártáros végzettségű (54 322 02 0100 52 01 OKJ) munkatárs részvétele az elektronikus másolatkészítés folyamatában, ennek hiányában a digitális másolatkészítés és metaadat-kezelés speciális feladataival kapcsolatos oktatáson részt vett, legalább középfokú végzettségű munkatárs alkalmazása.
- A másolatkészítés bármely szakaszában részt vevő munkatársak egyedi, a másolatkészítés technikai részleteit, valamint a metaadatok kezelését is magában foglaló továbbképzése.
- Az intézmény legalább egy saját dolgozójának megbízása az elkészült digitális másolatok megfelelőségének – alább részletezendő – ellenőrzésével.
- Az elkészült digitális másolatok szakszerű kezelése és biztonságos tárolása az intézmény informatikai szakembergárdájának felelősségi körébe kell tartozzon.
- Az eredeti iratok előkészítését, az elkészült másolatok utólagos minőségellenőrzését, valamint a teljes másolatkészítési folyamat irányítását az intézmény saját munkatársai kell végezzék.

### **4. A másolatkészítő rendszer működésére vonatkozó általános követelmények**

- A másolatkészítő rendszernek támogatnia kell a monokróm és a színes szkennelést egyaránt.
- A másolatkészítő rendszernek legalább a következő szabványos képformátumba kell tudnia képeket menteni és konvertálni: TIFF, JPEG, PDF/A-1a.

- A másolatkészítő rendszernek képesnek kell lennie különböző felbontásban elkészíteni és elmenteni a képeket. Az egyes beállításokat profilokba kell tudni menteni, így az azonos típusú dokumentumok azonos paraméterekkel való szkennelése garantálható.
- A másolatkészítő rendszernek képesnek kell lennie arra, hogy minimálisan a hagyományos dokumentumméreteket (A4, A3) kezelje.
- A másolatkészítő rendszernek rendelkeznie kell optikai karakterfelismerő (OCR, Optical Character Recognition) funkcióval, de szerencsés, ha az OCR mellett az ICR (Intelligent Character Recognition) lehetőségét is támogatja.
- Tömeges szkennelés esetén a rendszernek képesnek kell lennie a különálló dokumentumok felismerésére, sorba rendezésére és rögzítésére.
- A másolatkészítő rendszernek lehetőséget kell biztosítani a képek szkennelés utáni ellenőrzésére és szükség szerinti korrekciójára.
- A másolatkészítő rendszernek rendelkezni kell a képek alapvető minőségjavítására (kontrasztarány és fényerő változtatása, gammakorrekció, retusálás stb.) alkalmas szolgáltatással, illetőleg támogatnia kell erre alkalmas külső megoldások alkalmazását.
- A rendszernek lehetőséget kell biztosítani arra, hogy az adatrögzítő megjegyzéseket fűzhessen az elkészült képekhez.
- A rendszernek a szkennelési munkamenetek alábbi részleteit kell automatikusan naplóznia:
  - o felhasználó neve;
  - o munkaállomás azonosítója;
  - o be- és kilépés időpontja;
  - o a munkamenet azonosítója;
  - o kötegazonosító(k);
  - o dokumentumok száma (amennyiben értelmezhető);
  - o digitalizált képek száma;
- A rendszernek a következő metaadatokat automatikusan rögzítenie kell:
  - o a felhasználó azonosítója,
  - o a másolást végző szervezeti egység azonosítója,
  - o a szkennelés időpontja,
  - o az eredeti dokumentum fizikai méretei,

- a szkennelő eszköz illetve rendszer, valamint a másolatkészítési szabályzat megnevezése.
- A másolatkészítő rendszernek képesnek kell lennie a beszkennelt állományok bélyegképeit megjeleníteni a böngészés és keresés megkönnyítése érdekében.
- A másolatkészítő rendszernek együtt kell működnie az intézmény dokumentumkezelő rendszerével, amennyiben ilyen létezik.
- A rendszernek garantálnia kell, hogy az elektronikus másolat és annak papír alapú eredetije közti kapcsolat ténye a metaadatokban dokumentálódjon.
- A másolatkészítő rendszernek tudnia kell exportálni az iratokat metaadataikkal együtt kibocsátott információs csomag (Submission Information Package, SIP) formájában az OAIS szabvány szerint.
- A másolatkészítő rendszernek az intézmény adatvédelmi politikájával összhangban garantálnia kell a rendszerbe bekerült személyes vagy minősített, illetve egyéb okból érzékeny adatok kellő védelmét.
- A digitális másolatok készítésére használatos eszközök nem okozhatnak helyrehozhatatlan károsodást az eredeti iratokban, különös tekintettel az archiválási céllal digitalizált, maradandó értékű iratanyagokra.
- Automatikus laptovábbító eljárások csak abban az esetben engedélyezhetők, ha azt az eredeti iratok fizikai állapota megengedi. A papír alapú dokumentumok nem kockáztathatók a tömeges másolatkészítés érdekében.
- A másolatkészítő rendszer meg kell feleljen a 24/2006 BM-IHM-NKÖM együttes rendeletben megfogalmazott biztonsági, valamint a 2011. évi CXII. tv., illetve az 1998. évi VI. tv. adatvédelmi előírásainak.

## **5. Az iratok digitalizálásra való előkészítésével kapcsolatos követelmények**

A papír alapú eredeti dokumentumokat szakszerű előkészítése a digitalizálási program egyik kulcsa, hiszen ennek révén kerülhető el a hibás, használhatatlan elektronikus állományok létrejötte, illetőleg az eredeti anyagok károsodása.

- A szkennelés megkezdése előtt minden idegen irodaszert (kapcsok, elválasztólapok, műanyag borítók, madzagok stb.) el kell távolítani a dokumentumokról. Nem vonatkozik ez a közjegyzői vagy egyéb hitelesítéssel ellátott iratok összefűzésére.
- Ki kell emelni az iratokból azokat az elemeket, melyeket nem szükséges szkennelni (pl. kiadványok vagy nyilvánosan megjelent dokumentumok), illetőleg amelyeket saját kezű felbontásra szántak.



- Speciális méretű (pl. A3-as méretnél nagyobb vagy nagyon kicsi), egyedi formátumú (pl. bekötött) vagy egyedi adathordozójú (pl. mikrofilmen tárolt) iratok esetén a rendelkezésre álló szkennelő berendezés lehetőségei függvényében egyedi döntést kell hozni a másolás manuális elvégezhetőségéről. Ilyen esetekben egyaránt alkalmazható rollskenner, könyvskenner, mikrofilmskenner vagy digitális fényképezőgép, amennyiben az általa készített digitális másolat az elkészült képekkel szemben támasztott követelményeknek megfelel.
- El kell távolítani a műanyag vagy karton elválasztólapokat is. Amennyiben ezek érdemi információkat tartalmaznak, ezek fénymásolatát kell a digitalizálandó anyagok között elhelyezni. Ugyanez vonatkozik az iratokra ragasztott jegyzetlapokra (post-it) is.
- Az iratok fizikai állapota alapján – szükség esetén szakemberek bevonásával - egyedi döntést kell hozni a szkennelés kockázatairól. A nem szkennelhető anyagokat ki kell emelni és szükség esetén állományvédelmi intézkedéseket kell kezdeményezni.
- Szintén egyedi döntést kell hozni a szkennelés megkezdése előtt az adott irat történeti, művészeti stb. értéke alapján (pl. ha az iratok fényképeket, grafikákat, különösen fontos kéziratokat stb. tartalmaznak).
- Fényképeket, grafikákat stb. célszerű levéltári műanyag borítóban elhelyezni digitalizálás előtt.
- Automatikus laptovábbító eszközök használata esetén különös gondossággal kell mérlegelni a begyűrődés, szakadás kockázatait.
- Borítékban lévő iratokat ki kell emelni a borítékból. Amennyiben a boríték érdemi információt tartalmaz, azt üresen kell a digitalizálandó anyag elé helyezni, vagy annak fénymásolatával helyettesíteni.
- Amennyiben az iratok speciális adathordozókat (CD, DVD stb.) is tartalmaznak, annak további kezelése kapcsán egyedi döntést kell hozni az iratkezelésért felelős szakember(-ek) bevonásával.
- Automatikus adagolóval történő digitalizálás előtt az iratokat le kell számozni, hogy az ellenőrzés során megállapítható legyen az esetleges hiány (összeragadt lapok stb.).
- Tömeges szkennelés esetén - az adott szkennelő szoftver működésétől függően – üres vagy vonalkóddal ellátott elválasztólapokat kell az iratok közé helyezni.

## **6. A digitális másolatok készítésével és indexelésével kapcsolatos követelmények**

- A folyamat során, a digitalizáló rendszer képességei függvényében elhelyezhető kiegészítő információ is a képen (pl. vízjel, lapszám, dátum stb.), amennyiben az az

eredeti anyag olvashatóságát, illetve a rajta szereplő információk megjeleníthetőségét nem zavarja.

- Digitális másolatkészítés során a másolatkészítéssel kapcsolatos legfontosabb technikai metaadatokat a központi szkennelő szolgáltatás automatikusan, vagy annak működtetői manuálisan rögzítik.
- Az iratanyag leíró és adminisztratív metaadatait az Másolatkészítési Szabályzatban meghatározott, az iratkezelés megfelelő szintjéért felelős személyek rögzítik.
- A digitalizálás megtörténtével párhuzamosan, vagyis az elektronikus másolat életciklusának elején garantálni kell a másolat és eredetije nyilvántartási adatainak összetartozását. Ez leginkább úgy biztosítható, ha a másolat egy elektronikus ügyirat része lesz, mely a papír alapú irat nyilvántartási és egyéb metaadatait is tartalmazza.
- Az elkészült fájlok elnevezése rendkívül fontos része a folyamatnak. Az egyes fájlok nevének a dokumentum provenienciáját és a regisztratúrában elfoglalt helyét kell kifejeznie, hogy ez biztosítékot jelentsen az eredeti dokumentum és digitális másolata összetartozásának fennmaradására. A fájlok elnevezési szabályait a Másolatkészítési Szabályzatban részletesen rögzíteni és azt a gyakorlatban alkalmazni kell, a dokumentumkezelő rendszer pedig lehetőleg ki kell kényszerítenie a fájlnevezési konvenciók betartását.

## **7. A másolatkészítés során alkalmazható fájlformátumok**

- Szöveges (nyomtatott vagy írott) dokumentumok digitális másolatának elkészítése során a másolatkészítő rendszernek PDF/A-1 formátumú fájlokat kell létrehoznia (ISO 19005-1:2005). A digitalizálás során elsősorban az irat képi megfelelőjét kell eltárolni, optikai karakterfelismerés esetén az így készült szöveges állományt ugyanabban a PDF-fájlban (PDF/A-1a) kell tárolni.
- Fekete-fehér vagy színes fényképek, grafikák, tervrajzok, térképek szkennelése során a másolatkészítő rendszernek TIFF-formátumú képeket kell előállítania (ISO 12639:2004) az alábbi megkötésekkel:
  - o a 6.0 vagy újabb verziójú TIFF-képek használata, valamint
  - o kizárólag tömörítetlen, vagy veszteségmentes tömörítésű TIFF-képek elfogadhatóak a hosszú távú megőrzés szempontjából.

A korábban egyéb formátumban (JPEG, JPEG2000, PNG, GIF stb.) készült képeket TIFF formátumba kell konvertálni.

- Olyan dokumentumok esetében, melyeknél kizárólag a tartalmi és nem a formai megőrzés és használat a cél, alkalmazható az egyszerű TXT-formátum is, de kizárólag az általánosan elfogadott karakterkódolású (US-ASCII, ISO-8859-1, UNICODE Utf-8, Utf-16) TXT-állományok készítése elfogadható levéltári szempontból.

## 8. Az elektronikus másolatok felbontása

Az elektronikus másolatokat felbontása egyfelől a mind pontosabb képi megfelelést, másfelől az eredményesebb karakterfelismerési folyamat megalapozását szolgálja.

A fentiek érdekében a digitális másolatkészítési eljárások során a következő felbontásokat kell alkalmazni:

Dokumentum típusa	Minimális felbontás <sup>52</sup>
Magas kontrasztú, nyomtatott szöveges dokumentumok	200 ppi
OCR-ezendő szöveges dokumentumok	300 ppi
Színes, vagy alacsony kontrasztú, árnyalatgazdag szöveges dokumentumok	300 ppi
Kézírásos, vagy kézírást is tartalmazó nyomtatott dokumentumok	300 ppi
Grafikák, tervrajzok, térképek	300 ppi
Fekete-fehér vagy színes fényképek	600 ppi
Fekete-fehér v. színes negatív vagy pozitív film	2400 ppi

## 9. Az elektronikus másolatok színe, színmélysége, szintere

A digitális másolatok képi megfelelőségének egyik garanciája, hogy a számítógépes monitoron megjelenő kép minél hitelesebben reprodukálja az eredeti dokumentum látványát. Ennek megfelelően a hosszú távú megőrzésre, illetőleg ez eredeti papír alapú dokumentumok teljeskörű helyettesítése céljából készülő másolatokat 24 bites True Color színmélységben, AdobeRGB színtérben kell rögzíteni.

## 10. A másolatkészítés során minimálisan rögzítendő metaadatok

<sup>52</sup> Az egyedi igények és lehetőségek függvényében az itt meghatározott felbontási paraméterektől „felfelé” el lehet térni.

A digitális másolatok készítése során a másolatkészítés folyamatára vonatkozó alábbi metaadatokat kell rögzíteni.

- A kép tulajdonságaira vonatkozó metaadatok:
  - o Színmélység (bit)
  - o Tömörítés típusa
  - o Felbontás (ppi)
  - o Színtér
  - o Az eredeti dokumentum szélessége (cm)
  - o Az eredeti dokumentum magassága (cm)
  - o A kép orientációja
  
- A kép készítésével kapcsolatos metaadatok
  - o Az eredeti dokumentum jellemzőinek leírása
  - o A szkenneléshez (digitális fényképezéshez) használt eszköznek és gyártójának a pontos megnevezése
  - o A képfeldolgozáshoz használt szoftver és verziójának megnevezése
  - o A képfeldolgozás során alkalmazott képjavítási eljárások megnevezése
  - o A másolatkészítési szabályzat és verziójának megnevezése
  - o A másolatot készítő felhasználó
  - o A másolatot készítő szervezeti egység
  - o A másolat készítésének időpontja
  - o Érvényességi idő, vagy annak jelzése, hogy az érvényességi idő nem meghatározott

A fenti metaadatok kizárólag a másolatok készítésével kapcsolatban rögzítendő információkat tartalmazzák, az elektronikus ügyirat teljes metaadat-szerkezetét a levéltári ajánlások alapján kialakított és jóváhagyott séma alapján kell alkalmazni.

## **11. A másolatok utólagos kezelésére és minőségi ellenőrzésére vonatkozó követelmények**

- Optikai karakterfelismerés alkalmazása esetén az iratkezelő rendszernek egyetlen iratként kell kezelnie a digitalizált képet és az abból előállított szöveget.

- A másolatkészítés során létrejött digitális képen utólag végrehajthatók képjavítási eljárások (pl. kontraszt-növelés, világosítás/sötétítés stb.) a következő feltételekkel:
  - o a képjavítás kizárólag az eredeti anyag hitelesebb képi visszaadása, illetőleg jobb olvashatósága érdekében történhet,
  - o a módosítások nem járhatnak az eredeti dokumentumokban szereplő érdemi adatok elvesztésével, illetőleg azon nem szerepelt információk megjelenítésével.
  - o a képjavítás csak dokumentáltan, utólag ellenőrizhető módon történhet.
  
- A digitális másolatok utólagos minőségi kontrollja elengedhetetlen része a másolatkészítési folyamatnak. A minőségi ellenőrzés során minimálisan az alábbi követelményeket teljesíteni kell.
  - o Legalább egy felelős személyt meg kell bízni az elkészült másolatok rendszeres minőségi ellenőrzésével.
  - o Az ellenőrzés során folyamatosan felügyelni kell a szkennelésre vonatkozó, a Másolatkészítési Szabályzatban megfogalmazott előírások teljesítését.
  - o Az elkészült képfájlokat legalább 5%-os mintavételi gyakorisággal ellenőrizni kell a következő szempontok szerint:
    - tömeges szkennelés esetén nem történik-e kihagyás vagy üres lapok szkennelése,
    - teljesültek-e a képek minőségi követelményeire vonatkozó előírások,
    - az eredeti irat minden érdemi információját rögzítették-e,
    - helyes-e a digitális kép orientációja,
    - okozott-e képhibát a szkennер állapota, tisztasága,
    - a teljes képterület rögzítették-e,
    - az elkészült fájlok megnyithatók és használhatók-e?
  
- Hibás vagy hiányosan elvégzett digitális másolás esetén az ezzel megbízott munkatárs köteles azonnal intézkedni a korrekcióról.
  
- Ugyanez a lehetőség minden, az ügyintézés érdemi szakaszában részt vevő munkatársnak rendelkezésére kell álljon.

## 2. sz. függelék Összefoglaló táblázat a digitalizálási formátumok követelményeiről

Dokumentumtípus	Felbontás	Szín	Színmélység	Fájlformátum	Tömörítés
Magas kontrasztú, nyomtatott szöveges dokumentumok	Min. 200 ppi	színes	24 bit	PDF/A-1	tömörítetlen, vagy v.mentesen t. <sup>53</sup>
OCR-ezendő szöveges dokumentumok	300 ppi	Színes	24 bit	PDF/A-1	tömörítetlen, vagy v.mentesen t.
Színes, vagy alacsony kontrasztú, árnyalatgazdag dokumentumok	300 ppi	színes	24 bit	PDF/A-1	tömörítetlen, vagy v.mentesen t.
Kézírásos, vagy kézírást is tartalmazó nyomtatott dokumentumok	Min. 300 ppi	színes	24 bit	PDF/A-1	tömörítetlen, vagy v.mentesen t.
Grafikák, tervrajzok, térképek	Min. 300 ppi	színes	24 bit	PDF/A-1	tömörítetlen, vagy v.mentesen t.
Fekete-fehér fényképek	600 ppi	színes	24 bit	TIFF	tömörítetlen, vagy v.mentesen t.
Színes fényképek	600 ppi	színes	24 bit	TIFF	tömörítetlen, vagy v.mentesen t.
Fekete-fehér negatív vagy pozitív film	2400 ppi	színes	24 bit	TIFF	tömörítetlen
Színes negatív vagy pozitív film	2400 ppi	színes	24 bit	TIFF	tömörítetlen

<sup>53</sup> veszteségmentesen tömörített

### 3. sz függelék Digitalizálási ellenőrzési lista

<b>1.</b>	<b>A digitalizálás tervezése</b>	
1.1	Megtörtént-e a digitalizálásból származó előnyök, valamint a ráfordítandó erőforrások rendelkezésre állásának értékelése?	
1.2	Rögzítették-e a digitális másolatok hitelességét, integritását és hozzáférhetőségét megfogalmazó követelményrendszert?	
1.3	Létezik-e érvényes Iratkezelési Szabályzat és Irattári Terv, illetve Másolatkészítési Szabályzat?	
1.4	A különösen védendő adatok (személyes adatok, minősített adatok) elektronikus kezelésének kérdései tisztázottak-e?	
1.5	A későbbi használathoz és megőrzéshez szükséges metaadat-követelmények rögzítettek-e?	
1.6	A digitalizálási folyamat dokumentációs követelményei teljesültek-e?	
1.7	A megfelelő képminőség technikai specifikációja rendelkezésre áll-e?	
1.8	Az elektronikus tárolás követelményeit és költségeit meghatározták és kalkulálták?	
1.9	Meghatározták-e a digitalizálás módszereit (saját erőből, részben külső segítséggel vagy teljesen kiszervezett formában)?	
1.10	Meghatározták-e a digitalizálást végző személyzet felkészítésének követelményeit?	
<b>2</b>	<b>A digitalizálást megelőző folyamatok</b>	
2.1	Rendelkezésre áll a helyszínen elektronikus dokumentumkezelő rendszer, vagy ennek hiányában a másolatok használatára alkalmas szoftveres és hardveres környezet?	
2.2	Az érvényes iratkezelési szabályok érvényesek az elektronikus iratkezelésre is?	
2.3	Az iratanyag kellő rendezettségű, rendszerezettségű a digitalizáláshoz?	
2.4	A különösen védendő adatok kezelésére vonatkozó szabályokat alkalmazták-e az elektronikus iratkezelésben is?	
2.5	A szükséges metaadat-séma rendelkezésre áll-e mindenütt, ahol a metaadatok rögzítése meg kell történnjen?	
2.6	A szkennelést végzők megkapták-e a szükséges felkészítést?	
2.7	A minőségi ellenőrzés követelményei és ellenőrzési rendszerét kialakították-e?	
2.8	A digitális képek technikai specifikációját integrálták-e a rendszerbe?	
2.9	Az OCR szoftvert (ha alkalmaznak ilyet) integrálták-e a rendszerbe?	
2.10	Megfelelő és megfelelően méretezett elektronikus tárolókapacitás áll-e rendelkezésre?	

<b>3</b>	<b>Digitális másolás</b>	
3.1	A hagyományos adathordozójú iratot megfelelően előkészítették-e a digitalizálásra?	
3.2	A törékeny vagy más okból veszélyeztetett iratok azonosítása megtörtént-e?	
3.3	A digitális irat és metaadatai bekerülnek-e az elektronikus dokumentumkezelő rendszerbe?	
<b>4</b>	<b>Digitalizálás utáni feladatok</b>	
4.1	A digitalizált képek és az eszközök minőségi ellenőrzése a folyamat minden szintjén megtörtént-e?	
4.2	A digitális képállományok használatosságát és visszakereshetőségét ellenőrizték, tesztelték-e?	
4.3	Az iratkezelési szabályokat alkalmazzák-e a hagyományos és elektronikus iratokra egyaránt?	



## 4. sz. függelék A PDF/A-1 formátum specifikációja a PRONOM alapján

<b>Name</b>	Acrobat PDF/A - Portable Document Format
<b>Version</b>	1a
<b>Other names</b>	PDF/A (1)
<b>Identifiers</b>	MIME: application/pdf, Apple Uniform Type Identifier: com.adobe.pdf, PUID: fnt/95
<b>Family</b>	
<b>Classification</b>	Page Description
<b>Disclosure</b>	Full
<b>Description</b>	The Portable Document Format/ Archive is a format designed for long term preservation by Adobe Systems. PDF/A is a simplified version of PDF 1.4, with all of the features from PDF 1.4 that would impede long term preservation removed. Removed features include Javascript, Audio/Visual content, LZW compression and encryption. A major principle of PDF/A is that it is self contained and not reliant on externalities thus all font and colour information is encoded into the file. PDF/A files are larger than other types of PDF files due to the need for embedded information. PDF/A supports two levels of compliance PDF/A1-a (Accessible) and PDF/A1-b (Basic). PDF/A1-a is fully ISO 19005-1:2005 (PDF/A-1) compliant whereas PDF/A1-b is less stringent and not compliant. PDF/A1-a requires tagged PDF and Unicode whereas PDF/A1-b does not.
<b>Orientation</b>	Binary
<b>Byte order</b>	Big-endian (Motorola)
<b>Related file formats</b>	Has priority over Acrobat PDF 1.4 - Portable Document Format (1.4)
<b>Technical Environment</b>	
<b>Released</b>	01 Oct 2005
<b>Supported until</b>	
<b>Format Risk</b>	
<b>Developed by</b>	Adobe Systems
<b>Supported by</b>	
<b>Source</b>	Digital Preservation Department / The National Archives

## 5. sz. függelék A TIFF formátum specifikációja a PRONOM alapján

<b>Name</b>	Tagged Image File Format
<b>Version</b>	
<b>Other names</b>	TIFF
<b>Identifiers</b>	PUID: fmt/353, MIME: image/tiff, Apple Uniform Type Identifier: public.tiff
<b>Family</b>	
<b>Classification</b>	Image (Raster)
<b>Disclosure</b>	Full
<b>Description</b>	The Tagged Image File Format (TIFF) is a raster image format originally developed by the Aldus Corporation, primarily for use in scanning and desktop publishing. When Adobe Systems Incorporated purchased Aldus in 1994, they acquired the rights to the TIFF format and have maintained it since then. TIFF files comprise three sections: an Image File Header (IFH), an Image File Directory (IFD), and the image data. TIFF files can contain multiple images (multi-page TIFF), and each image has a separate IFD. The IFH always appears at the beginning of the file, and is immediately followed by a pointer to the first IFD. The IFD contains metadata which describes the associated image, stored as a series of tags. The IFD also contains a pointer to the actual image data. TIFF supports colour depths from 1 bit to 24 bit (e.g. monochrome to true colour), and a wide range of compression types (RLE, LZW, CCITT Group 3 and Group 4, and JPEG), as well as uncompressed data.
<b>Orientation</b>	
<b>Byte order</b>	Little-endian (Intel) and Big-endian (Motorola)
<b>Related file formats</b>	None.
<b>Technical Environment</b>	
<b>Released</b>	01.aug.86
<b>Supported until</b>	
<b>Format Risk</b>	
<b>Developed by</b>	Aldus Corporation
<b>Supported by</b>	None.
<b>Source</b>	Digital Preservation Department / The National Archives

## 6. sz. függelék A JPEG formátum specifikációja a PRONOM alapján

<b>Name</b>	JPEG File Interchange Format
<b>Version</b>	02.jan
<b>Other names</b>	JFIF (1.02)
<b>Identifiers</b>	PUID: fmt/44, MIME: image/jpeg, Apple Uniform Type Identifier: public.jpeg
<b>Family</b>	
<b>Classification</b>	Image (Raster)
<b>Disclosure</b>	Full
<b>Description</b>	The JPEG File Interchange Format (JFIF) is a file format for storing JPEG-compressed raster images. It was developed by the Independent JPEG Group and C-Cube Microsystems, in the absence of any such format being defined in the JPEG standard, and rapidly became a de facto standard. A JFIF file comprises a JPEG data stream together with a JFIF marker. It begins with a Start of Image (SOI) marker, immediately followed by a JFIF Application (APP0) marker and one or more optional application extension markers. This is followed by the JPEG image data, which is terminated by an End of Image (EOI) marker. JFIF supports up to 24-bit colour and uses lossy compression (based on the Discrete Cosine Transform algorithm). Other types of compression are available through JPEG extensions, including progressive image buildup, arithmetic encoding, variable quantization, selective refinement, image tiling, and lossless compression, but these may not be supported by all JFIF readers and writers.
<b>Orientation</b>	Binary
<b>Byte order</b>	Big-endian (Motorola)
<b>Related file formats</b>	Has priority over Raw JPEG Stream , Is subsequent version of JPEG File Interchange Format (1.01)
<b>Technical Environment</b>	
<b>Released</b>	01 Sep 1992
<b>Supported until</b>	
<b>Format Risk</b>	
<b>Developed by</b>	C-Cube Microsystems, Independent JPEG Group
<b>Supported by</b>	None.
<b>Source</b>	Digital Preservation Department / The National Archives

**7. sz. függelék: a JPEG2000 formátum specifikációja a PRONOM alapján**

<b>Name</b>	JPEG2000
<b>Version</b>	
<b>Other names</b>	
<b>Identifiers</b>	public.jpeg-2000
<b>Family</b>	
<b>Classification</b>	Graphics file format
<b>Disclosure</b>	Full
<b>Description</b>	JPEG2000 is The National Archives' preferred format for digitisation. The qualities that the format presents by being a ratified ISO/IEC open source standard (ISO/IEC 15444) and providing high levels of compression for many image types suggest that it is suitable economically and structurally for preservation over long periods of time. Our primary requirement of any JPEG2000 file accessioned as part of this process is that it uses lossless compression, that is, the image should be mathematically lossless from that captured by the digital imaging device. Please see Appendix 1 for a more detailed description of our required JPEG2000 profile.
<b>Related file formats</b>	JPEG File Interchange Format
<b>Technical Environment</b>	
<b>Released</b>	2000
<b>Supported until</b>	
<b>Format Risk</b>	
<b>Developed by</b>	Joint Photographic Experts Group
<b>Supported by</b>	None.
<b>Source</b>	Digital Preservation Department / The National Archives