

## ÁLTALÁNOS MŰSZAKI KÖVETELMÉNYEK

### 1. SZKENNELÉSI KÖVETELMÉNYEK

#### 1.1. Általános szkennelési követelmények

1.1.1. Az alkalmazni kívánt felbontás: 400 dpi.

1.1.2. Az alkalmazni kívánt képfórmátum: tömörítetlen TIFF.

1.1.3. *Szöveges kiadványok esetében:*

1.1.3.1. A színmélység: 8 bit.

1.1.3.2. A kötetek 80%-a szöveges kiadvány, melyek nem tartalmaznak színes anyagokat, illusztrációkat (mellékleteket, táblázatokat, ábrákat, stb.).

1.1.4. *Színes kiadvány(részek) esetében:*

1.1.4.1. A színmélység 24 bit.

1.1.4.2. Ajánlatkérő a kötetek Kísérőlapjain külön jelzi a színesben kért oldalak helyét.

1.1.4.3. Színes szkennelést csak kalibrált szkennereken lehet elvégezni. Színes szkennelés esetén Ajánlatkérő kéri a színtáblázatot is a kötet digitális változata mellett elhelyezni.

#### 1.2. Az oldalkép minősége

1.2.1. Az egy-egy dokumentumoldalról készített digitális felvételnek egy képnek kell lennie, nem lehet sem szerkesztett, sem összeillesztett kép.

1.2.2. Egy kép csak egy dokumentumoldalt tartalmazhat.

1.2.3. A képnek a teljes oldalt kell tartalmaznia.

1.2.4. A margók mérete ne legyen kisebb 50 pixelnél és ne legyen nagyobb 200 pixelnél.

1.2.5. Az oldalferdeség nem haladhatja meg a 0,5 fokot.

1.2.6. A nyomtatott oldal mérethűen az eredeti kötetnek megfelelően (1:1 arányban) legyen elmentve.

1.2.7. Az oldalméreték beállítása kötetenként legyen azonos (az oldalméreték közötti eltérés mértéke mind szélességben, mind hosszúságban ne haladja meg a +/- 5 pixelt). Az egyes kötetekben előforduló eltérő méretű dokumentumrészekről a Kísérőlap ad iránymutatást.

1.2.8. A dokumentumok oldaltájolását megváltoztatni nem szabad.

1.2.9. Minden digitalizált dokumentum címlapjáról készüljön színes bélyegkép.

1.2.10. Ajánlatkérő a következő képjavító eljárás elvégzését kéri: körbevágás.

## 2. A SZKENNELÉS MÓDJAI

### 2.1. Digitalizálási módok

2.1.1. A dokumentum fizikai *egységének megőrzésével* történő digitalizálás (a dokumentumot állománykímélő módon kötetben kell digitalizálni, könyvkímélő technológia alkalmazásával, nyithatóság minél kisebb dőlésszögben).

Teljes mennyiség: 1.700.000 oldal

2.1.2. A dokumentum fizikai *egységének részleges megőrzésével* történő digitalizálás (a dokumentum kiemelhető kötéséből, de ívei nem vágathatók fel).

Teljes mennyiség: 100.00 oldal

2.1.3. A dokumentum fizikai *egysége megőrzése nélkül* történő digitalizálás (a dokumentum lapra szedhető, laponként digitalizálható).

Teljes mennyiség: 200.000 oldal

### 2.2. A digitalizálás helyszínei

2.2.1. A szkennelési feladatokat az alábbi helyszíneken lehet elvégezni:

#### 2.2.1.1. *Helyszín 1*

2.2.1.1.1. Az Országház épületében, az Országgyűlési Könyvtár területén,

2.2.1.1.2. Hét közben és hétvégén egyaránt lehet munkát végezni, nemzeti és állami ünnepeken nincs lehetőség munkavégzésre.

2.2.1.1.3. Tervezhető napi műszakok száma: 3.

#### 2.2.1.2. *Helyszín 2*

2.2.1.2.1. Az Országgyűlési Könyvtár külső raktárában (Budapest, XVIII. Nagyenyed u. 18. sz.).

2.2.1.2.2. Tervezhető műszakok száma hét közben: 2.

2.2.1.2.3. Hétvégén munkát végezni nem lehet, nemzeti és állami ünnepeken nincs lehetőség munkavégzésre.

#### 2.2.2. *Munkarend*

2.2.2.1. Egyműszakos munkarend esetén tervezett munkaidő: 8.00-16.00 óra.

2.2.2.2. Kétműszakos munkarend esetében tervezett munkaidő: 06.00-14.00, 14.00-22.00 óra.

2.2.2.3. Három műszakos munkarend esetén tervezett munkaidő: 06.00-14.00, 14.00-22.00 és 22.00-06.00 óra.

### 2.3. A digitalizálás ütemezése

2.3.1. Ajánlatkérő a digitalizálás ütemezését Ajánlattevővel egyeztetve, un. Teljesítési ütemtervben rögzíti.

2.3.2. A tömeges méretű digitalizálás ütemezésében célszerű az egyes dokumentumcsoportokat külön digitalizálni (könyvek, döntvénytárak, országgyűlési dokumentumok, folyóiratok, közlönyök). Az egyes

dokumentumcsoportokon belül javasolt újabb csoportokat kialakítani, amelyekben fő válogatási szempont (lehet) az azonos módon szkennelhető dokumentumok összeválogatása.

## **2.4. Egyéb követelmények**

- 2.4.1. A könyvtári dokumentumokról egyedi, vagy csoportos állapotleírás készült, amelyen Ajánlatkérő rögzítette a dokumentumok terjedelmi, szerkezeti és átadás kori állapotának legfontosabb jellemzőit (továbbiakban: Kísérőlap). A Kísérőlap egy másolati példánya kíséri a dokumentumokat.
- 2.4.2. A Kísérőlap minden egyes dokumentum, vagy dokumentum csoport vonatkozásában meghatározza a digitalizálás módját, amelytől Ajánlattevő csak Ajánlatkérővel való konzultációt követően térhet el. A konzultációról mindkét fél által aláírt jegyzőkönyvet kell felvenni.
- 2.4.3. Ugyancsak a Kísérőlap határozza meg azon dokumentumok körét, amelyek fizikai egysége digitalizálás után helyreállítandó, vagyis újra be kell kötni őket. A Kísérőlap dokumentumonként meghatározza az újrakötés módját is.
- 2.4.4. Ajánlatkérő a munkafolyamat során 300.000 oldal tartalék dokumentumállományt biztosít annak érdekében, hogy az esetleg kieső dokumentumok cseréjével a 2.000.000 oldal digitalizálása fenntartható legyen. A csere szükségességét Ajánlattevő írásban jelzi Ajánlatkérő számára. A csere kizárólag az Ajánlatkérő támogatása mellett jöhet létre. Az országgyűlési dokumentumok esetében cserére nincs lehetőség.
- 2.4.5. Ajánlatkérő elvárja, hogy Ajánlattevő a digitalizálási eljárás során a legnagyobb gondossággal járjon el, olyan óvatossággal, amilyen elvárható a törekeny papírok kezelése folyamán. Ha az Ajánlattevő olyan súlyos sérülést észlel, amely érdemben befolyásolja a műszaki paraméterek teljesülését, vagy a munkafolyamat a dokumentum további súlyos károsodását okozza, akkor köteles egyeztetést kezdeményezni Ajánlatkérővel.
- 2.4.6. A dokumentum fizikai egységének részleges megőrzésével történő digitalizálást Ajánlatkérő egy speciális dokumentumtípusra, az oldalról fűzőkapoccsal összefogott 1-10 ív terjedelmű brosúrákra dolgozta ki (összesen kb. 200 db). A tűzész mód miatt ezen dokumentumok belső margója nagyon szűk, a fűzőkapocs miatt a kötetek lapozhatósága is nehézkes. Ajánlatkérő ezeket a köteteket előkészítve adja át Ajánlattevőnek (ívekre bontva, lapra szedve, vagy géppel lapozható módon).

## **3. A KARAKTERFELISMERÉS PARAMÉTEREI**

### **3.1. Általános követelmények**

3.1.1. Teljes mennyiség: 2.000.000 oldal

3.1.2. A főszöveg karakterfelismerési minősége legalább 95%, azaz 100 karakter maximum 5 hibát tartalmazhat.

3.1.3. A hibaarány ennél nagyobb lehet speciális esetekben: hibás, rongált oldalak, a törzsszövegnél jelentősen kisebb betűkkel nyomtatott részek valamint többirányú

szöveget tartalmazó táblázatok esetében, de itt is el kell érnie a legalább 90%-os minőséget (100 karakter maximum 10 hibát tartalmazhat).

- 3.1.4. Amennyiben a tesztelés során megállapítást nyer, hogy egyes oldalak esetében a hibaarány meghaladja az Ajánlatkérő által előírt mértéket, úgy Ajánlattevő – külön díj felszámítása nélkül – köteles a szolgáltatás újbóli elvégzésére. Amennyiben Ajánlatkérő az előírt hibaarány teljesítéséhez szükséges minőségű digitalizálendő dokumentumot nem tud biztosítani, és ennek következtében a hibák a szkennelés megismétlésével sem küszöbölhetők ki, Ajánlatkérő elfogadja az adott teljesítést.
- 3.1.5. A digitalizált oldalak automatikus karakterfelismerésre kerülnek (OCR) és ezt követően kétrétegű PDF formátumban kerülnek elmentésre.
- 3.1.6. A karakterfelismerés a mellékletek és a képek címére is terjedjen ki.
- 3.1.7. Ajánlattevőnek meg kell valósítani az infokommunikációs akadálymentesítést.
- 3.1.8. Alkalmazott karakterkészlet: UNICODE karakterkészlet, UTF-8 kódolásban, normalizált formában, az ékezetes karakterek: „precombined” formában.
- 3.1.9. A szkennelési hibák kézi javítása nem követelmény. Ezen kivétel azonban nem érintheti a speciális OCR követelmények betartását.

### 3.2. Speciális OCR követelmények

- 3.2.1. Az Ajánlatkérő az általános OCR minőségi követelményekhez képest szigorúbb követelményrendszert állapít meg (azaz a karakterfelismerés 995 ezrelékes pontosságú kell legyen, vagyis 1000 karakterből 5 lehet hibás.), az alábbi dokumentum csoportok, illetve dokumentumrészletek esetében:
  - 3.2.1.1. **Könyvek:** A kötet elején vagy a végén található tartalomjegyzék, mutatók.
  - 3.2.1.2. **Jogforrások (döntvénytárak, közlönyök):** A jogszabályok címe, a jogszabályok száma, a kibocsátás éve, a kibocsátó neve (ez általában egy rövidítés), jogszabály típusa (ez lehet rövidítés). Az azonosítására szolgáló elemek az egyes jogszabályok előtt található. Az előforduló jogszabályok lehetséges típusairól, a kibocsátókról Ajánlatkérő külön listát bocsát rendelkezésre.
  - 3.2.1.3. **Lexikonok:** A lexikon szócikkek címe (a szócikkek elején található legelső szó, általában félkövér kiemeléssel).

### 3.3. Az OCR pontosságának ellenőrzése

- 3.3.1. Ajánlatkérő a 3.1.2 pontban megfogalmazott követelményeket dokumentumcsoportonként véletlenszerű mintavétellel kívánja ellenőrizni.

## 4. MENTÉS, ARCHIVÁLÁS

### 4.1. Fájlnev

- 4.1.1. Ajánlatkérő a képeket egyedi azonosító jelzettel kéri ellátni, ez a fájlnev.
- 4.1.2. A képfájlok elnevezése írásjeleket és ékezetes betűket nem tartalmazhat. Kivéve alsó vonal jelet: \_ UNICOD UTF-8 name: LOW LINE, cod: U+005F

- 4.1.3. Ajánlatkérő a fájlok elnevezését a Műszaki dokumentáció 6. számú mellékletében meghatározott elvárások szerint kéri elnevezni.
- 4.1.4. A szerződés teljesítése során a 4.1.1 és 4.1.2 pontokban felsorolt kérdésekben a teljesítés megkezdése előtt a könyvtár által kijelölt munkatárssal konzultálni kell.

## **4.2. A digitalizált állományok átadása**

- 4.2.1. Szkenelt állomány: Vállalkozónak a szkennelt állományokat tömörítetlen TIFF formátumban, két példányban, 3,5”, minimum 1 TB (terabyte) SATA, USB 2.0 csatlakozású külső meghajtó egységen kell átadnia. A külső meghajtó egységen fel kell tüntetni a lemez tartalmát.
- 4.2.2. OCR-ezett állomány: Vállalkozónak az optikai karakterfelismeréssel (OCR) elkészített állományokat PDF formátumban, két példányban, 3,5”, minimum 1 TB (terabyte) SATA, USB 2.0 csatlakozású külső meghajtó egységen kell átadnia. A külső meghajtó egységen fel kell tüntetni a lemez tartalmát. A Vállalkozó által rendelkezésre bocsátott adathordozó a Megrendelő tulajdonává válik.
- 4.2.3. A külső meghajtó egység átadása során jegyzőkönyvet kell felvenni, amely tartalmazza:
- az átadás idejét,
  - az átadás helyét,
  - a résztvevők személyét, aláírását,
  - az átadott külső meghajtó egység jellemzőit, mennyiségét, tartalmát,
  - és a külső meghajtó egység működőképességét jellemző leírást.
- 4.2.4. Külső meghajtónként Ajánlatkérő kéri megjelölni és megadni:
- külső jelöléssel: meghajtó azonosítója/átadás dátuma. pl. OGYK1/20100721
  - elektronikusan a gyökér könyvtárban: meghajtó azonosítója/ átadás dátuma/dokumentum típusonként/dokumentumok darabszám
  - CD-n és kinyomtatva is az átvételi listát, amely tartalmazza: meghajtó azonosítója/ átadás dátuma/dokumentum típusonként/dokumentumok lemezre rögzített fájlok listája, amit átvételkor, de később is össze illetve be lehet azonosítani a lemez-tartalommal.

## **5. ADATBÁZISBA SZERVEZÉS**

### **5.1. Általános követelmények**

- 5.1.1. Ajánlatkérő a digitalizált állományokat adatbázisba szervezve kéri, azaz a digitális adatbázisban kereshetően álljanak rendelkezésre a metarekordok.
- 5.1.2. Az adatbázisba szervezés részkérdéseit a *Műszaki Dokumentáció 7 számú melléklete (Útmutató az adatbázisba szervezés kérdéseihez)* tartalmazza.
- 5.1.3. Ajánlatkérő – a Vállalkozói szerződés aláírását követő 30 naptári napon belül – reprezentatív mennyiségű (kb. 500 db) metarekord átadását *kéri előzetes vizsgálat céljára*. A vizsgálat célja a megtalált hiányosságok kiküszöbölése. A

metarekordokat az első szállítmány feldolgozását követően kell átadni a Megrendelőnek.

## 5.2. Új metarekordok létrehozása

5.2.1. Teljes mennyiség: 70.000 metarekord

5.2.2. A könyvtár metarekordot biztosít a digitalizálandó dokumentumok alábbi csoportjaihoz:

- könyvek (összesen max. 4.000 db),
- országgyűlési dokumentumok,
- időszaki kiadványok részdokumentumainak (folyóiratcikkek) egy részéről (összesen max. 30.000 db, +–10%).

5.2.3. Ajánlatkérő adatbázisában 70.000 folyóiratcikk metarekordja hiányzik – ezekről az Útmutató szerint Ajánlattevőnek kell elkészítenie az új metarekordokat.

## 5.3. Adatbázis kapcsolatok létrehozása

5.3.1. Teljes mennyiség: 100.000 metarekord

5.3.2. Elvégzendő feladat a dokumentumok beazonosítása, és a DigiTool-ban összekapcsolásuk a metaadatokkal, majd a digitális dokumentum feltöltése az adatbázisba.