



# Information Integrator for Content kezelése

*8.2 változat*





# Information Integrator for Content kezelése

*8.2 változat*

**Megjegyzés**

A kiadvány és a benne tárgyalt termék használatba vétele előtt olvassa el a “Megjegyzések” oldalszám: 131 részt.

**Második kiadás (2003. március)**

Ez a kiadás az IBM Enterprise Information Portal for Multiplatforms (termékszám: 5724-B43) 8.2 változatára és minden további kiadására és módosítására vonatkozik mindaddig, amíg ezt a későbbi kiadások másként nem jelzik.

A termék bizonyos részei: Copyright © 1990-2000 ActionPoint, Inc. és/vagy licencelői, 1299 Parkmoor Drive, San Jose, CA 95126 U.S.A. Minden jog fenntartva.

Outside In<sup>®</sup> Viewer Technology, © 1992-2000 Inso Corporation. Minden jog fenntartva.

© Szerzői jog IBM Corporation 1999, 2003. Minden jog fenntartva

# Tartalom

## Néhány szó a könyvről . . . . . v

Kiknek szól ez a kézikönyv? . . . . .	v
Az adminisztrátorok szaktudása . . . . .	v
Az üzleti elemzők vagy folyamatmodellezők szaktudása . . . . .	vi
További információforrások . . . . .	vi
Terméksomagban található információk . . . . .	vi
Támogatás az Interneten . . . . .	vii
Megjegyzések küldése . . . . .	vii
Az EIP 8.2 változatának újdonságai . . . . .	viii

## 1. fejezet Az Enterprise Information

### Portal bemutatása . . . . . 1

Ügyfélinformációk keresése . . . . .	1
Az igény . . . . .	1
A megoldás . . . . .	2
Áttekintés . . . . .	2
Enterprise Information Portal összetevők bemutatása . . . . .	2

## 2. fejezet Adminisztrációs ügyfél

### bemutatása . . . . . 7

Az EIP Első lépések használata a rendszeradminisztrációs ügyfél megértéséhez . . . . .	7
Az EIP adminisztrálása . . . . .	7
Felhasználók és csoportok kezelése . . . . .	7
Az adminisztrációs ügyfél eszközeinek használata . . . . .	7
Jogosultságok bemutatása . . . . .	9
Jogosultságok . . . . .	9
Termék nézetek és adatbázisok átkapcsolása . . . . .	10
Az adminisztrációs ügyfél javításai és bővítései . . . . .	10
Adminisztrációs ügyfél csatlakoztatása helyi adminisztrációs adatbázishoz . . . . .	11
Adminisztrációs ügyfél csatlakoztatása távoli adminisztrációs adatbázishoz . . . . .	11
1. lépés - Távoli adatbázis katalogizálása a DB2 konfigurációs segédlettel . . . . .	11
2. lépés - Kiszolgáló konfigurációs segédprogram használata . . . . .	12
3. lépés - Távoli adatbázis kapcsolatának ellenőrzése . . . . .	14
Dokumentumtípusok meghatározása . . . . .	14
A kiszolgáló MIME típus fájljának (cmbcc2mime.ini) módosítása . . . . .	14

## 3. fejezet Az EIP adminisztrációs ügyfél

### szolgáltatásainak használata . . . . . 17

Egyesített keresés létrehozása . . . . .	17
Kiszolgálók meghatározása . . . . .	17
Kiszolgálók meghatározásának irányelvei . . . . .	18
OnDemand kapcsolatok kezelése: TCP/IP hangolás és socketek . . . . .	23
Extended Search kapcsolat kezelése . . . . .	23
Egyesített példányok létrehozása . . . . .	23
Egyesített példányok – Áttekintés . . . . .	23
Az Egyesített példány létrehozása varázsló használata . . . . .	24
Egyesített szövegindexek létrehozása . . . . .	25

Keresési sablonok létrehozása . . . . .	25
Keresési sablon meghatározása . . . . .	25
Keresési feltételek meghatározása . . . . .	26
Keresési beállítások meghatározása . . . . .	26
Jogosultságok hozzárendelése . . . . .	26

## 4. fejezet Felhasználói hozzáférés

### kezelése . . . . . 29

Felhasználói azonosítók és jelszavak létrehozása . . . . .	29
DB2 adminisztratori jogosultság . . . . .	30
Csatlakozás a DB2 adatbázishoz INI fájlok használatával . . . . .	30
Könyvtárkiszolgáló és erőforráskezelő rendszeradminisztratori jelszó módosítása . . . . .	31
Adatbázis hozzáférési jelszavak módosítása . . . . .	31
Felhasználók importálása LDAP címtárból . . . . .	32
Jogosultságok bemutatása . . . . .	32
Jogosultságkészletek létrehozása . . . . .	33
Jogosultsági csoportok létrehozása . . . . .	33
Jogosultságkészlet felhasználókhoz rendelése . . . . .	34
Adományozható jogosultságkészlet felhasználói azonosítóhoz rendelése . . . . .	34
Felhasználók erőforráskezelőkhöz rendelése . . . . .	34
Felhasználók gyűjteményekhez rendelése . . . . .	34
Felhasználói csoportok létrehozása . . . . .	34
Hozzáférés felügyeleti listák létrehozása . . . . .	35
Jogosultságkészlet hozzáférés felügyeleti listához rendelése . . . . .	35
Tartományok létrehozása . . . . .	35
Tartományok felügyelete . . . . .	36
Tartományok elérése . . . . .	36
Felhasználó tartományhoz rendelése . . . . .	37
Felhasználói csoport tartományhoz rendelése . . . . .	37
Jogosultságkészlet tartományhoz rendelése . . . . .	37
Objektumkiszolgáló tartományhoz rendelése . . . . .	37
Gyűjtemény tartományhoz rendelése . . . . .	37
Felhasználó áthelyezése egyik tartományból a másikba . . . . .	38
Felhasználói csoport áthelyezése egyik tartományból a másikba . . . . .	38
Erőforráskezelő áthelyezése egyik tartományból a másikba . . . . .	39
Gyűjtemény áthelyezése egyik tartományból a másikba . . . . .	39
Jogosultságkészlet áthelyezése egyik tartományból a másikba . . . . .	39
Hozzáférés felügyeleti lista áthelyezése egyik tartományból a másikba . . . . .	39

## 5. fejezet Információbányászat kezelése 41

Mi az Információbányászat? . . . . .	41
Az Enterprise Information Portal Információbányászat szolgáltatásai . . . . .	41
Az Információbányászat szolgáltatások összetevői . . . . .	42
Az Információbányászat használata üzleti környezetben . . . . .	44
Egy példa az Információbányászat felhasználására . . . . .	46
Támogatott nyelvek és formátumok . . . . .	49

Alapelvek . . . . .	49
A rendszer felépítése . . . . .	49
Az Információbányászat alapelvei . . . . .	51
Információbányászati eszközök . . . . .	52
Programozási illesztők . . . . .	59
Első lépések . . . . .	61
Osztályozás összeállítása . . . . .	61
Az Információstrukturálási eszköz telepítése . . . . .	61
Kezdeti lépések . . . . .	62
Hozzáférési jogok . . . . .	62
Osztályozás meghatározása . . . . .	62
Betanítási dokumentumok kiválasztása . . . . .	64
Betanítási dokumentumok feltöltése . . . . .	64
Kategorizálási modell kiértékelése . . . . .	67
Katalógus betanítása . . . . .	70
Teljesítményhangolás . . . . .	71
Az IBM Web Crawler használata . . . . .	71
Az IBM Web Crawler képességei . . . . .	72
Az IBM Web Crawler webes változatának beállítása és futtatása . . . . .	72
Az IBM Web Crawler konfigurációs fájl . . . . .	75
Az IBM Web Crawler naplózása . . . . .	83
Hibaelhárítás . . . . .	84
Összegzők kiválasztása . . . . .	85
IBM Web Crawler Notes változata . . . . .	86
Az IBM Web Crawler kizárása a kiszolgálókról . . . . .	90

## 6. fejezet Bevezetés a munkafolyamatokba . . . . . 93

Munkafolyamatok magyarázata . . . . .	93
Munkafolyamatok alkalmazása . . . . .	93
Felhasználói azonosítók és csoportok szinkronizálása . . . . .	94
EIP kiszolgáló újratelepítése a munkafolyamatok engedélyezése esetén . . . . .	95
Felhasználói azonosítók és csoportok frissítése az MQSeries Workflow és az EIP adatbázis között . . . . .	95
Munkafolyamat tervezése . . . . .	96
Feldolgozott információk . . . . .	96
Információkezelés módja . . . . .	97
Végrehajtandó tevékenységek . . . . .	97
Hogyan haladnak végig az információk a folyamaton? . . . . .	97
A részek összekapcsolódása . . . . .	98
Enterprise Information Portal munkafolyamat összetevők használata . . . . .	98
Munkafolyamat tervező használata . . . . .	98
Munkafolyamat szolgáltatások használata . . . . .	99
Feladatlisták meghatározása . . . . .	99
Tevékenységlisták meghatározása . . . . .	100
Folyamatok létrehozása . . . . .	100
Munkafolyamat tervező engedélyezése . . . . .	100
MQSeries Workflow kiszolgáló indítása . . . . .	100

## 7. fejezet IBM Web Crawler mintafájlok 103

config-sample2.xml példa . . . . .	103
IBM Web Crawler napló elemzési fájl példa . . . . .	105

## 8. fejezet A szöveges keresés és a QBIC használata . . . . . 109

Dokumentumok keresése a szöveg keresőmotorral . . . . .	109
Szöveg keresési kiszolgáló engedélyezése . . . . .	109
Képek keresése a képek tartalma alapján (QBIC) . . . . .	109
Bevezetés a kép keresésbe . . . . .	109
Kép keresés beállítása . . . . .	110
Mintaadatok betöltése és indexelése . . . . .	112
Mielőtt betöltené a mintaadatokat . . . . .	112
Szöveg keresési index létrehozása . . . . .	113
Kép keresési adatbázis, katalógus és szolgáltatások létrehozása . . . . .	114
A betöltési program futtatása . . . . .	115
A példa szöveges adatainak indexelése . . . . .	116

## 9. fejezet Dokumentumformátumok 117

Információbányászat dokumentumformátumai . . . . .	117
Szövegszerkesztők: Általános . . . . .	117
Szövegszerkesztők: DOS . . . . .	117
Szövegszerkesztők: Nemzetközi . . . . .	118
Szövegszerkesztők: Windows . . . . .	118
Szövegszerkesztők: Macintosh . . . . .	119
Táblázatkezelő formátumok . . . . .	119
Adatbázis formátumok . . . . .	119
Szabványos grafikus formátumok . . . . .	120
Professzionális grafikus formátumok . . . . .	122
Bemutató formátumok . . . . .	122
Tömörített és kódolt formátumok . . . . .	122
Egyéb . . . . .	123

## 10. fejezet Jogok kezelése . . . . . 125

Szellemi tulajdon védelme . . . . .	125
Megjelölési technikák alkalmazása . . . . .	126
Látható megjelölés . . . . .	127
Láthatatlan megjelölés . . . . .	127

## 11. fejezet Kisegítő lehetőségek . . . . . 129

Billentyűzetről végzett bevitel és navigáció . . . . .	129
Kisegítő megjelenítőkkel kapcsolatos lehetőségek . . . . .	129
Kompatibilitás a kisegítő technológiákkal . . . . .	129
Hozzáférhető dokumentáció . . . . .	130

## Megjegyzések . . . . . 131

Védjegyek . . . . .	133
---------------------	-----

## Szószedet . . . . . 135

## Tárgymutató . . . . . 141

---

## Néhány szó a könyvről

Ez a kézikönyv az Enterprise Information Portal (EIP) rendszer kezeléséhez szükséges alapvető ismereteket írja le. Mivel az EIP több, az adminisztrációs ügyfélből kezelhető összetevővel is rendelkezik, továbbá mivel az EIP használatával más termékek funkciói is elérhetők, ez a kézikönyv nem egy hagyományos értelemben vett rendszeradminisztrációs útmutató. A könyv az alábbi témaköröket dolgozza fel:

- Az EIP használata az üzleti igények kielégítésére.
- Az adminisztrációs ügyfél indítása és használata.
- Felhasználói hozzáférés kezelése.
- Az EIP használata több tartalomkiszolgáló tartalmának keresésére, ideértve a relációs adatbázisok adatait, a szervezetlen vagy multimédiás tartalmat vagy a szöveges dokumentumokat.
- Munkafolyamatok tervezése, megvalósítása és kezelése.

---

## Kiknek szól ez a kézikönyv?

A kézikönyv az EIP adminisztrátorokat segíti a következő feladatok végrehajtása során:

### **Rendszeradminisztráció**

Beleértve az adatbázis, a kiszolgáló és a hálózat felügyeletét.

### **Felhasználók kezelése**

Felhasználók és csoportok létrehozása, jogosultságaik meghatározása, hozzáférés felügyeleti listák kezelése.

### **Egyesített keresések**

Egyesített keresési sablonok létrehozása és használata a Content Manager rendszer tartalmának visszakereséséhez.

### **Információbányászat**

Információk kinyerése a dokumentumokból, dokumentumok és keresési eredmények kategorizálása.

### **Web bejárás**

Webes tartalom keresése és importálása az IBM Web Crawler felhasználásával.

### **Szöveg keresés**

Dokumentumok indexelése és keresése az IBM DB2 TIE vagy az IBM szöveges keresőmotorjával (csak a Content Manager 7.1 és korábbi változatai esetén).

### **Kép keresés**

A Content Manager 7.1 (és korábbi) változatainak használata kép keresések végrehajtására.

### **Munkafolyamat kezelés**

A vállalat munkafolyamatainak irányítása az EIP munkafolyamat eszközeinek segítségével.

---

## Az adminisztrátorok szaktudása

A végrehajtott feladatok alapján az adminisztrátoroknak a következőkkel kell tisztában lenniük:

- Felhasználói hozzáférés biztonsági protokolljai.

- Windows NT, Windows XP, Windows 2000, AIX vagy Solaris operációs rendszerek kezelése.
- Hálózati adminisztráció.
- Content Manager rendszer tartalomkiszolgálóin alkalmazott adatmodellek.
- Adatbázis adminisztráció.
- A tartalom és a keresési feltételek ismerete a keresési sablonok létrehozásához.
- Információbányászati módszerek és eszközök.
- A munkafolyamat tervezés alapelvei.
- Az EIP munkafolyamat szolgáltatásaival támogatni kívánt üzleti folyamatok.

---

## Az üzleti elemzők vagy folyamatmodellezők szaktudása

A kézikönyvben az üzleti elemzők és folyamatmodellezők is találhatnak alapvető információkat az EIP munkafolyamatok meghatározásával és modellezésével kapcsolatban.

Az Enterprise Information Portal folyamat tervező használatához képesnek kell lennie az alábbiakra:

- A vállalat üzleti folyamataiban alkalmazott személyi követelmények, programok és adatszerkezetek megértése.
- A cég munkafolyamataival vagy üzleti folyamataival kapcsolatos döntések meghozatala.

---

## További információforrások

A termékcsomag teljes információkészletet biztosít a rendszer tervezésével, telepítésével, felügyeletével és használatával kapcsolatban. A termékdokumentáció és a támogatás az Interneten is elérhető.

## Termékcsomagban található információk

A termékcsomag tartalmaz egy Információs központot, illetve minden kiadvány megtalálható benne PDF formátumban.

### Az Információs központ

A termékcsomag tartalmaz egy Információs központot, amely a termék telepítésének részeként telepíthető. Az információs központ telepítésével kapcsolatban további információkat a *Planning and Installing Your Content Management System* című kiadványban talál.

Az Információs központ Content Manager, Enterprise Information Portal és IBM Content Manager VideoCharger for Multiplatforms dokumentációkat tartalmaz. A témakör alapú információk termékenként és feladatonként (például adminisztráció) vannak szervezve. A navigációs lehetőségek és tárgymutatók mellett keresési szolgáltatás is rendelkezésre áll.

### PDF kiadványok

A PDF fájlokat a csomagban található Adobe Acrobat Reader megfelelő operációs rendszerre készült változatával tekintheti meg. Ha az Acrobat Reader még nincs telepítve, akkor töltsse le azt az Adobe webhelyéről: [www.adobe.com](http://www.adobe.com).

Az IBM Content Manager for Multiplatforms csomagban található Content Manager kiadványokat az 1. táblázat: oldalszám: vii sorolja fel.



### 1. táblázat: Content Manager kiadványok

Fájlnev	Cím	Kiadvány száma
install	<i>Planning and Installing Your Content Management System<sup>1</sup></i>	GC27-1332-01
migrate	<i>Áttérés a Content Manager 8. változatára</i>	SC22-5301-01
sysadmin	<i>System Administration Guide</i>	SC27-1335-01

Az IBM Content Manager for Multiplatforms megrendelésekor az IBM Enterprise Information Portal for Multiplatforms is szállításra kerül. Ettől függetlenül az IBM Enterprise Information Portal for Multiplatforms önállóan is megrendelhető. A termékhez tartozó Enterprise Information Portal kiadványokat a 2. táblázat sorolja fel.

### 2. táblázat: Enterprise Information Portal kiadványok

Fájlnev	Cím	Kiadvány száma
apgwork	<i>Application Programming Guide for Windows<sup>1</sup></i>	SC27-1347-01
ecliinst	<i>Az e-Ügyfél telepítése, beállítása és kezelése</i>	SC22-5302-02
eipinst	<i>IBM Information Integrator for Content tervezése és telepítése</i>	GC22-5303-01
eipmanag	<i>IBM Information Integrator for Content kezelése</i>	SC22-5304-01
messcode	<i>Üzenetek és kódok<sup>2</sup></i>	SC22-5305-01

#### Megjegyzések:

1. Az *Application Programming Guide for Windows* című kiadvány a Content Manager és az Enterprise Information Portal alkalmazások programozásáról is tartalmaz információkat.
2. Az *Üzenetek és kódok* című kiadvány a Content Manager és az Enterprise Information Portal üzeneteit is tartalmazza.

## Támogatás az Interneten

A terméktámogatás az Interneten is elérhető. Kattintson a termék alábbi webhelyein a **Support** hivatkozásra.

[www.ibm.com/software/data/cm/](http://www.ibm.com/software/data/cm/)

[www.ibm.com/software/data/eip/](http://www.ibm.com/software/data/eip/)

A termékhez tartozó dokumentáció elektronikus változatban érkezik. Az Internetes dokumentációk eléréséhez kattintson a termék webhelyének **Library** hivatkozására.

Az Interneten emellett elérhető egy HTML alapú dokumentációs felület is, az online dokumentáció. Itt jelenleg az API referenciainformációk találhatók. Az online dokumentáció elérésével kapcsolatban nézze meg az Enterprise Information Portal Library webhelyét.

## Megjegyzések küldése

A visszajelzések segítenek bennünket a minőségi információk biztosításában. Várjuk megjegyzéseit erről, vagy bármely másik Content Manager vagy Enterprise Information Portal dokumentációról. A megjegyzéseket az alábbiak szerint juttathatja el hozzánk:

- Küldje el megjegyzéseit a weben. Használja az IBM adatkezelés online megjegyzések űrlapját a következő címen:

[www.ibm.com/software/data/rcf](http://www.ibm.com/software/data/rcf)

Ezen az oldalon beírhatja és elküldheti a megjegyzéseit.

- Küldje el a megjegyzéseit e-mailben a [comments@vnet.ibm.com](mailto:comments@vnet.ibm.com) címre. Ne feledje el megadni a termék nevét és verziószámát, valamint a könyv nevét és termékszámát. Ha a megjegyzések egy adott szövegrészre vonatkoznak, akkor adja meg a szöveg helyét is (például a fejezet és a szakasz címét, illetve a táblázatszámot, oldalszámot vagy a sugótémakör címét).

## Az EIP 8.2 változatának újdonságai

A termékben az alábbi változások történtek:

### Sun Solaris támogatása

A kapcsolatok, szolgáltatások és adatbázisok telepíthetők Solaris rendszerekre.

### Közös rendszeradminisztráció

Azonos ügyfélalkalmazás használható a Content Manager és az Enterprise Information Portal adminisztrációjára.

### Új kapcsolatok

- A Content Manager 8.1 kapcsolat lehetővé teszi a Content Manager v8 erőteljes dokumentumtárolási képességeinek kihasználását.
- Az új C++ Extended Search 3.7 kapcsolat változat AIX alatt is fut.

### Tökéletesített kapcsolatok

- A paraméteres szöveg keresések az egyesített rétegen és közvetlen Extended Search kapcsolaton keresztül is támogatottak.
- Funkcionális bővítések és teljesítménynövekedés az OnDemand kapcsolatnál, beleértve a következőket:
  - Változások az OnDemand DDO szerkezetében.
  - Aszinkron keresés támogatása

### Új Információbányászat szolgáltatások

- Jellemző kivonatolás
- Fürtözés
- Nyelvazonosítás

### IBM Web Crawler

Az IBM Web Crawler lehetővé teszi a webes és Lotus Notes adatbázisokban tárolt információk bejárását és összegzését.

### Munkafolyamat továbbfejlesztések

A munkafolyamat szolgáltatások teljes mértékben támogatottak AIX és Solaris operációs rendszereken. A folyamat tervező, az API-k és a JavaBeans komponensek még jobb működést és használhatóságot biztosítanak.

### Információs központ

A böngésző alapú Információs központ tartalmazza az Content Manager, Enterprise Information Portal és IBM Content Manager VideoCharger for Multiplatforms dokumentációkat. A témakör alapú információk termékenként és feladatonként (például adminisztráció) vannak szervezve. A navigálási lehetőségek és tárgymutatók mellett keresési szolgáltatás is rendelkezésre áll.

### Kisegítő lehetőségek

A kisegítő lehetőségek biztosítják a fizikai kihívásokkal küzdő, például mozgásában vagy látásában korlátozott felhasználóknak a szoftvertermékek sikeres használatát. A termék lényegesebb kisegítő lehetőségei:

- A termék minden funkciója elérhető a billentyűzet segítségével is, egér használata nélkül.
- Kiterjesztett megjelenítési tulajdonságok támogatása.

- Látható és hallható riasztások lehetősége.
- Kompatibilitás a segítő technológiákkal.
- Kompatibilitás az operációs rendszer kisegítő lehetőségeivel.
- Elérhető dokumentációs formátumok.



---

## 1. fejezet Az Enterprise Information Portal bemutatása

A sok papírt igénylő vállalatok – például a biztosítási cégek és gazdasági intézmények – nagy mennyiségű üzleti adatot kezelnek. Számos iparágban felmerült az igény az üzleti információk hozzáféréseinek és kezelésének vállalati szintű megoldására.

*A tartalomkiszolgálók olyan szoftverrendszerek, amelyek multimédiás adatokat, üzleti űrlapokat, dokumentumokat, adatokat, valamint a tartalom feldolgozásához szükséges metaadatokat tartalmazzák. Ha nincs lehetőség a különféle tartalomkiszolgálók hatékony összekapcsolására, akkor a vállalkozások időt és pénzt pazarolnak az információk többszörös tárolásával és az alkalmazottak kiképzésével többféle keresés végrehajtására.*

Az Enterprise Information Portal (EIP) a legújabb technológiákkal a vállalkozás összes erőforrását a munkaállomás munkaasztalára varázsolja. Az EIP a különböző tartalomkiszolgálók egyetlen ügyfélen keresztül megoldott összekapcsolásával maximalizálja a vállalkozás információs és multimédia tőkéjének értékét. Az EIP ügyféllel a felhasználók gyorsan és párhuzamosan férhetnek hozzá az összes csatlakoztatott tartalomkiszolgálóhoz. A felhasználók ezen kívül információbányászati műveleteket és összetett kereséseket is végrehajthatnak a web vagy egy intranet tartalomkiszolgálóin. Emellett elvégezhetik az üzleti tevékenységekkel kapcsolatos folyamatokat.

Az EIP lehetővé teszi az alkalmazások testreszabását az adott vállalat igényeinek megfelelően. Az EIP példák a programozók munkaasztal és web alapú alkalmazásokat írhatnak.

Ez a rész az EIP terméket tekinti át. Egy kitalált biztosítótársaság, az XYZ Biztosító példáján keresztül bemutatja az EIP szolgáltatásait és funkcióit.

---

### Ügyfélinformációk keresése

Az XYZ biztosítónak (XYZ) sok fotóból, keresetből, kötvényből, kárbecslői feljegyzésből, szakértői jelentésből és egyéb üzleti dokumentumból álló gyűjteménye van.

Az XYZ Biztosító minden kiküldött emlékeztetőt, valamint orvosi és kárbecslési elektronikus űrlapot Lotus Domino.Doc fájl tárolókban tárol. Hosszútávon az XYZ minden kötvény bejelentést, megjegyzést és számlát egy Content Manager OnDemand kiszolgálón tárol, ahol a felhasználók gyorsan hozzáférhetnek az adatokhoz. Az XYZ minden igénylési űrlapot, fotót és a kötvénytulajdonosoktól kapott levelet egy Content Manager for iSeries rendszermappában tárol. Az XYZ Biztosító a szakértői jelentéseket egy DB2 Universal Database (DB2 UDB) Data Warehouse Center Information Catalog Manager adatbázisban tárolja. Az XYZ Biztosító a vállalati médiatulajdont, például a nagy felbontású képeket egy Content Manager rendszeren tárolja hirdetési célból, valamint hogy megoszthassa a nyilvánossággal és az új vállalati osztályokkal. Ezekon felül az XYZ bizonyos információkat, például a vállalati eljárásokat az intraneten tárolja.

### Az igény

A kereseteket, az ügyfélhívásokat és az általános kötvénytulajdonosi szolgáltatásokat nem lehet egy kiszolgáló tartalmából kezelni, mivel az alkalmazottaknak minden ügyfélinformációra szükségük van. Az ügyfélszolgálat biztosítása érdekében az alkalmazottaknak számos tartalomkiszolgálóhoz van szükségük párhuzamos hozzáférésre. Az XYZ Biztosítónak egy olyan megoldásra van szüksége, amely az információk kereséséhez és

lekérdezéséhez összekapcsolja a tartalomkiszolgálókat és a vállalati intranetet. A biztosító ki szeretné terjeszteni a munkafolyamat feldolgozás használatát is.

A hivatalnokoktól kezdve az ügyintézőkön át az ügynökökig különböző alkalmazottaknak van szükségük a dokumentumokra. Az XYZ Biztosítónak egyes elemekhez korlátoznia kell a hozzáférést, míg másokhoz korlátlan hozzáférést kell biztosítania. Az XYZ Biztosítónak a betanítási idő és költség csökkentése miatt egy könnyen használható felületre van szüksége.

## A megoldás

Az XYZ Biztosító bevezeti az EIP rendszert, amelynek átfogó keresési technológiái lehetővé teszik a kiszolgálók csatlakoztatását, és az adatok keresését az összes kiszolgáló adatai között. Ha az XYZ telefonközpontjának kezelője egy hívást kap, akkor egy egyesített kereséssel megjelenítheti a kötvénytulajdonossal kapcsolatos valamennyi szükséges információt.

Az XYZ Biztosító az EIP információbányászat szolgáltatásának segítségével keresi ki és jeleníti meg az információkat a vállalati intranetről. Ezek használatát kívánják kibővíteni a munkafolyamat feldolgozásban is.

---

## Áttekintés

Az EIP egy átfogó termék. Összetevői úgy működnek együtt, hogy az adott vállalkozásra szabott, egyedi megoldást nyújtanak. A többszintű architektúrára összpontosítva az EIP tartalmaz egy adminisztrációs programot a keresések kezeléséhez, egy ügyfelet a keresések futtatásához, és kapcsolatokat a különböző tartalomkiszolgálók (például IBM Content Manager, Content Manager ImagePlus for OS/390, Content Manager OnDemand, Lotus Domino.Doc, DB2 Universal Database, DB2 DataJoiner és DB2 Data Warehouse Center Information Catalog Manager) összefogásához. Az EIP kapcsolati eszközkészletek és példák segítségével a további tartalomkiszolgálókhoz kiegészítő kapcsolatok írhatók.

Az EIP architektúra lehetővé teszi, hogy az ügyfélalkalmazások több tartalomkiszolgálón futtassanak kereséseket. A keresésekhez az ügyfelek az EIP adminisztrátorok által meghatározott keresési sablonokat használják.

A keresési sablonok segítségével az ügyfelek *egyesített kereséseket* futtatnak, amelyek párhuzamosan keresnek több olyan tartalomkiszolgálón, amelyek eredeti attribútumai le vannak képezve keresési sablonban használt egyesített attribútumokra. Az EIP keresési sablonjai olyan egyesített attribútumokra hivatkozó keresési feltételeket tartalmaznak, amelyek az egyes tartalomkiszolgálók eredeti attribútumaira vannak leképezve. A keresési sablonokat az EIP adminisztrátor hozza létre. Az EIP biztosítja a különféle tartalomkiszolgálókon tárolt adatok eléréséhez és kereséséhez szükséges kapcsolatokat. A tartalomkiszolgálók ezután képesek az adatobjektumok visszaadására az ügyfélnek.

Az EIP architektúra előnyei:

- Egyetlen lekérdezéssel férhet hozzá több különböző tartalomkiszolgálóhoz, amelyek támogatják az e-üzleti tranzakciókat és az ügyfélszolgálati alkalmazásokat.
- Információbányászat több tartalomkiszolgálón, beleértve a webkiszolgálókat is.
- A különböző tartalomkiszolgálók adatainak elérése munkafolyamatok segítségével.
- Olyan ügyfélalkalmazások fejlesztésének a támogatása, amelyek az ügyfélalkalmazások, az indexek és az adatok elkülönítése miatt függetlenek az adatok helyétől a tartalomkiszolgálókon.

## Enterprise Information Portal összetevők bemutatása

Ez a szakasz írja le az EIP összetevőit és a telepítési lehetőségeket.

Az összetevőket és a kompatibilis operációs rendszereket a 3. táblázat: oldalszám: 3 sorolja fel.

3. táblázat: EIP összetevők operációs rendszer kompatibilitása

Összetevő	Windows	AIX	Solaris	Megjegyzések
Adminisztrációs adatbázis	igen	igen	igen	Az adatbázis tartalmazza a folyamat tervező funkcionalitást.
Adminisztrációs ügyfél	igen	nem	nem	Az ügyfél Windows, AIX és Solaris operációs rendszerek adatbázisaihoz tud csatlakozni.
Kapcsolatok	igen	igen	igen	
Információbányászat	igen	igen	igen	
IBM Web Crawler	igen	igen	igen	
Szöveg keresési ügyfél	igen	igen	igen	
Kép keresési ügyfél	igen	igen	igen	
Kapcsolati eszközkészlet és példák	igen	igen	igen	<ul style="list-style-type: none"> <li>A Windows változat tartalmazza a forráskódot a példa ügyfél lefordításához. A példa ügyfél kódja AIX alatt nem kerül telepítésre.</li> <li>A munkafolyamat példák és API-k az egyesített kapcsolat példáival kerülnek telepítésre.</li> </ul>
Megjelenítő	igen	nem	nem	OnDemand ügyfél és megjelenítő.
Információs központ	igen	igen	igen	

## Adminisztráció

Az adminisztrációs összetevő az adminisztrációs adatbázis és az adminisztrációs ügyfél részösszetevőket tartalmazza. Az adminisztrációs adatbázis telepítésekor telepítésre kerül a munkafolyamat szolgáltatás is.

**Adminisztrációs adatbázis:** Az adminisztrációs adatbázis egy DB2 adatbázis, amely az EIP felhasználók és csoportok információit, a jogosultsági szintek, jelszavak, felhasználói azonosítók adatait és egyéb információkat tárolja. Emellett biztosítja a munkafolyamat és választhatóan az információbányászat szolgáltatásokat is. Több adatbázist is telepíthet. Mindegyik adatbázis biztosítja az EIP munkafolyamat szolgáltatásokat. Ha rendelkezik Content Manager v8 rendszerrel, akkor az EIP adminisztrációs adatbázist megoszthatja a Content Manager könyvtárkiszolgáló adatbázisával. Az adatbázis megosztása azért lehetséges, mert a könyvtárkiszolgáló adatbázis az EIP által igényelt összes információt tartalmazza.

**Adminisztrációs ügyfél:** Az Adminisztrációs ügyfél csak Windows munkaállomásokra telepíthető. Több ügyfelet is telepíthet. Ha rendelkezik Content Manager v8 rendszerrel, akkor az EIP és a Content Manager v8 azonos ügyfélből adminisztrálható.

Az ügyfél kezelőfelületével az adminisztrátor az alábbi feladatokat végezheti el:

- Tartalomkiszolgálók meghatározása az egyesített keresésekhez.

- Eredeti példányok és attribútumok beazonosítása a tartalomkiszolgálókon, és ezek leképezése egyesített példányokra.
- A felhasználó által megadott tartalomkiszolgálók eredeti példányait és attribútumait tartalmazó nyilvántartás karbantartása.
- Keresési sablonok létrehozása.
- Felhasználók és csoportok azonosítása és kezelése.
- Jogosultságok és jogosultságkészletek hozzárendelése a felhasználókhoz.
- Hozzáférés megadása a keresési sablonokhoz és a felhasználók által a keresésből származó adatokon végrehajtható műveletek feltételeinek megadása.
- Üzleti munkafolyamatok tervezése és adminisztrálása.

## Kapcsolatok

A kapcsolatok biztosítják az EIP ügyfelek, a tartalomkiszolgálók és az adminisztrációs adatbázis közötti kommunikációs illesztőket. A tartalomkiszolgálók kapcsolatai lehetővé teszik az EIP számára a bejelentkezést a kiszolgálóra, az információk keresését és visszaadását az adminisztrációs vagy a végfelhasználó ügyfelek számára. Az adminisztrációs ügyfelet az egyesített kapcsolat csatlakoztatja az adminisztrációs adatbázishoz.

Az EIP az alábbi kapcsolatokat tartalmazza:

- Egyesített kapcsolat, amely az EIP ügyfelet az adminisztrációs adatbázishoz csatlakoztatja.
- DB2 Universal Database 7.1 relációs adatbázis kapcsolat, JDBC illesztőprogram 1.3 változat (csak Java), ODBC 3.0 (csak C++), DataJoiner 2.1.1.
- Content Manager kapcsolat a Content Manager 7.1 kiszolgálóhoz.
- Content Manager kapcsolat a Content Manager 8.2 kiszolgálóhoz.
- Content Manager OnDemand kapcsolat a Content Manager OnDemand 7.1 kiszolgálóhoz.
- Content Manager for VisualInfo/400 4.3 és 5.1 változatok.
- Content Manager ImagePlus for OS/390 kapcsolat az ImagePlus/390 Folder Application Facility 3.1 és az Image Plus/390 ODM 3.1 termékhez
- Lotus Domino.Doc kapcsolat a Domino.Doc 3.0a és a Desktop Enabler 3.0a termékhez.
- Extended Search kapcsolat a 3.7 változathoz.
- Information Catalog Manager kapcsolat a DB2 Universal Database Visual Warehouse 5.2 és a DB2 Universal Database 7.2 termékhez.

## Szolgáltatások

Az EIP négy választható szolgáltatással rendelkezik.

### Információbányászat

Az Információbányászat nyelvi szolgáltatásai lehetővé teszik a rejtett információk megkeresését a tartalomkiszolgálók szöveges dokumentumaiban. A szöveges dokumentumok feldolgozása közben metaadatokat hoz létre, amelyek összesíthetők, kategorizálhatók és kereshetők. Az információbányászat egyik előfeltétele a WebSphere alkalmazáskiszolgáló 4.0 változata. Az említett funkciók mellett lehetőség van a dokumentumok csoportosítására, a dokumentumok nyelvének meghatározására, és bizonyos jellemzők, például vállalat- vagy személynevek kinyerésére.

### Kép keresési ügyfél

Biztosítja a Content Manager v7 tartalomkiszolgálók által nyújtott kép keresési funkciók elérését és felügyeletét.

### Szöveg keresési ügyfél

Biztosítja a Szöveg keresés kiszolgálók által nyújtott szöveg keresési funkciók elérését és felügyeletét.



## IBM Web Crawler

A Web Crawler egy Java alapú tartalom bejáró és kereső. A Web crawler intraneten, extraneten, az Interneten, Lotus Notes adatbázisokban, Domino kiszolgálókon, helyi fájlrendszerekben és FTP gyűjteményekben végezhet bejárást.

A Web Crawler különféle típusú tartalmakból képes metaadatokat és szöveget bányászni. A HTML tartalom bányászható például URL, cím, szövegtörzs, legutóbbi módosítás dátuma és metajelölők, például szerző, kulcsszavak és leírás alapján. Egy adott tartalomtípushoz többféle előre beállított bejáróból is választhat. A rendszer a tartalmat és/vagy a bányászott metaadatokat helyi lemezre menti.

## Tartalom megjelenítő

Az OnDemand megjelenítő telepítésekor a telepítő az OnDemand ügyfelet és az OnDemand kiszolgálókról visszaadott dokumentumok megjelenítéséhez szükséges további fájlokat telepíti.

## Kapcsolati eszközkészletek és példák

Az EIP termékben található egy példaprogramokat is tartalmazó kapcsolati eszközkészlet, amellyel különböző EIP funkciók működését tesztelheti.

- Tartalomkiszolgálók csatlakoztatása és szétkapcsolása
- SQL és egyéb példa lekérdezések végrehajtása a tartalomkiszolgálókon
- Tartalomkiszolgáló MIME típusainak és egyéb információinak meghatározása

**Windows kapcsolati eszközkészlet:** A kapcsolati eszközkészlet Windows kiszolgálókra telepítéséhez a Fejlesztői munkaállomás számítógéptípust kell kiválasztani. Ezután válassza ki a Kapcsolati eszközkészlet és a Példák összetevőt. A példaprogramok az összes kapcsolathoz telepíthetők, illetve kiválaszthatók a telepített kapcsolatoknak megfelelő egyedi példák is.

Windows kiszolgálókon a kapcsolati eszközkészlet példaprogramok szervezése a következő:

```
c:\CMBROOT\SAMPLES\activex\xx  
c:\CMBROOT\SAMPLES\cpp\xx  
c:\CMBROOT\SAMPLES\java\xx  
c:\CMBROOT\SAMPLES\jsp\xx  
c:\CMBROOT\SAMPLES\server\xx
```

ahol az *xx* az egyes példaprogramoknak megfelelő kapcsolat könyvtára, például *db2*, *od*, *dl*, *stb*.

AIX kiszolgálókon a példaprogramok szervezése a következő:

```
/usr/lpp/cmb/samples/cpp/icm  
/usr/lpp/cmb/samples/java/xx  
/usr/lpp/cmb/samples/jsp/xx  
/usr/lpp/cmb/samples/server/exi t
```

ahol az *xx* egy könyvtárnév, például *beans*, *servlets*, *stb*.

Solaris kiszolgálókon a példaprogramok szervezése a következő:

```
/opt/IBMcmb/samples/java/xx  
/opt/IBMcmb/samples/jsp/xx  
/opt/IBMcmb/samples/server/exi t
```

ahol az *xx* egy könyvtárnév, például *beans*, *servlets*, *stb*.

| A példaprogramokhoz dokumentáció is tartozik, amely leírja a programokat és megadja a  
| példakód számára szükséges kiszolgáló beállításokat (környezeti beállítások, memória, stb.).

### **Információs központ**

Az Információs központ összetevő az Enterprise Information Portal Információs központját tartalmazza. Az Információs központ a Enterprise Information Portal könyvtár web alapú, kereshető kiadása.

---

## 2. fejezet Adminisztrációs ügyfél bemutatása

Az adminisztrációs ügyfél biztosítja az EIP adminisztrációs adatbázis és az EIP rendszeradminisztrátor kapcsolatát. Ez a szakasz írja le az ügyfél által az EIP rendszer kezeléséhez biztosított új funkciókat.

Bizonyos szolgáltatások és funkciók, például a kiszolgálómeghatározások és a felhasználókezelés a bal oldalon található ikonok segítségével érhetők el. A további funkciók az Eszközök menüben találhatók.

---

### Az EIP Első lépések használata a rendszeradminisztrációs ügyfél megértéséhez

Az *Első lépések* modul az EIP minden telepítésének része. Az *Első lépések* minta adatok segítségével népesíti be az objektumokat, így a gyakorlást nem valós adatokon kell végeznie. Az *Első lépések* segítségével felfedezheti a kiszolgálómeghatározásokat, a felhasználókat és csoportokat, illetve a további szolgáltatásokat, melyek segítségével alapszintű betekintést nyer az adminisztrációs ügyfél működésébe és használatába.

---

### Az EIP adminisztrálása

A rendszeradminisztrátor az alábbi tevékenységeket végezheti el az adminisztrációs ügyfél segítségével:

- Tartalomkiszolgálók meghatározása
- Felhasználók és csoportok kezelése
- Jogszükségletek és hozzáférési szintek kezelése
- Egyesített keresési sablonok létrehozása
- Egyesített példányok létrehozása.
- Résztartományok létrehozása, ha az adminisztrációs tartományok engedélyezettek.
- Munkafolyamatok kezelése, ha a munkafolyamat szolgáltatások engedélyezettek.
- Egyesített szövegpéldány létrehozása a Content Manager 7. változatában.

---

### Felhasználók és csoportok kezelése

A felhasználókat felhasználói azonosítók és jogosultságok létrehozásával jogosíthatja fel a tartalomkiszolgálókon tárolt dokumentumok keresésére és kezelésére. A rendszeren tárolt adatok hozzáféréseinek korlátozásához létre kell hozni és hozzá kell rendelni a felhasználókhoz a megfelelő jogosultságokat.

---

### Az adminisztráció ügyfél eszközeinek használata

Ez a szakasz írja le az adminisztrációs ügyfél által biztosított eszközöket.

#### LDAP konfiguráció

Az elemre kattintáskor az EIP megjelenít egy négy lapból álló ablakot:

- LDAP lap - Lehetővé teszi az LDAP kiszolgálók adatainak importálását és az LDAP felhasználó importálás és/vagy hitelesítés engedélyezését.
- Kiszolgáló lap - LDAP kiszolgáló meghatározásához szükséges információk megadása, például hosztnév, felhasználói név, hivatkozási típus, stb.
- Hitelesítés lap - itt találhatók a Védett socket réteg (SSL) beállítási információk.

- Továbbiak lap - maximális rekordszámmal és a kiszolgáló időkorláttal kapcsolatos beállítások.

### **Felhasználói leképezés beállítása**

Lehetővé teszi az alapértelmezésben engedélyezett Felhasználó leképezés letiltását.

### **Egyesített felhasználói leképezés szerkesztő**

Az egyesített felhasználó szerkesztő megjeleníti a felhasználók listáját, és lehetővé teszi a felhasználók leképezését különféle tartalomkiszolgálókra.

### **Keresési sablon megjelenítő**

A keresési sablon megjelenítő részletes információkat biztosít az összes keresési sablonról. A megjelenítő három lehetőséget biztosít a keresési sablon részletek megtekintésére:

- Társított leképezések (alapértelmezett) - részletek az egyesített példányokról és a keresési sablonról
- Keresési sablon - részletek az alapértelmezett operátorokról, alapértelmezett értékekről, stb.
- Eredmények megjelenítése - részletek a megjelenítési névről, megjelenítési szélességről, a feltételek sorrendjéről, stb.

### **Kiszolgáló tároló megjelenítő**

Megjeleníti a kijelölt kiszolgálók tárolóit.

### **Napló megjelenítő**

A napló megjelenítő segítségével tekintheti meg a kiszolgáló tároló frissítése során keletkezett naplót. A naplóban láthatók az üzenetek, amelyek az új és a korábbi tárolók közötti különbségeket jelzik.

### **Szolgáltatások**

Válassza a Szolgáltatások elemet a munkafolyamatok engedélyezéséhez, vagy válassza ki az Információbányászat szolgáltatást.

### **Adminisztrációs tartományok**

Válassza ezt a beállítást az adminisztrációs tartományok engedélyezéséhez. Az adminisztrációs tartományok az engedélyezés után többé nem tilthatók le.

### **MIME típus szerkesztő**

A MIME típus szerkesztő az alábbi információkat sorolja fel valamennyi tartalomkiszolgálóról:

- Tartalomosztály
- Fájlkiterjesztés
- Relációs adatbázis (RDB) oszlop
- MIME típus

A MIME típus szerkesztőben megjelenő tartalomkiszolgáló nevek rövidítve vannak, és megegyeznek az új tartalomkiszolgálók beállításakor megjelenő tartalomkiszolgáló listában szereplő névvel. **Tipp:** a DL a Content Manager 7.1 tartalomkiszolgálók rövidítése. A V4 rövidítés a Content Manager for AS/400 tartalomkiszolgálókat jelöli.

A MIME típus szerkesztőben lehetséges az alapértelmezett információk hozzáadása, eltávolítása és szerkesztése.

### **MIME - alkalmazás szerkesztő**

A MIME - alkalmazás szerkesztő segítségével adhatja hozzá, törölheti vagy módosíthatja az öt alapértelmezett MIME - alkalmazás társítást. A MIME - alkalmazás szerkesztőben beállított értékek a végfelhasználók által használt megjelenítőkre vannak hatással.

### Kiszolgálótípus meghatározás

Ezzel az eszközzel határozhat meg saját kiszolgálókat.

### DB2 azonosító/jelszó módosítása

Ezzel a lehetőséggel cserélheti le a csatlakozási jogosultságokkal rendelkező DB2 felhasználói azonosítót és jelszót. A felhasználói azonosítónak semmi köze az adminisztrátori azonosítóhoz.

---

## Jogosultságok bemutatása

Ez a szakasz tárgyalja az Enterprise Information Portal jogosultságait. A Jogosultság ikon kibontásával négy jogosultsághoz férhet hozzá. **Tipp:** Mivel a Content Manager v8 és az EIP v8 egyazon ügyfélből felügyelhető, az ügyfél mindkét oldalon minden jogosultságot megjelenít.

## Jogosultságok

Az adminisztrációs ügyfél előre meghatározott jogosultságokat, jogosultsági csoportokat és jogosultságkészleteket tartalmaz. A jogosultságok adnak felhatalmazást a rendszer felhasználóinak, például az adminisztrátoroknak és végfelhasználóknak bizonyos objektumok egy adott módon történő kezelésére.

### Jogosultságok

Az EIP számos jogosultságot biztosít. A jogosultság egy adott objektum egy adott módon végrehajtott elérésére vonatkozó jog. A végfelhasználókhoz hozzárendelhető például az ItemAdd és ItemDelete jogosultság, amelyek birtokában az ügyfelek létrehozhatnak és törölhetnek elemeket a tartalomkiszolgálókon. A jogosultságok megtekintéséhez bontsa ki a Hitelesítés ikont, és kattintson duplán a Jogosultságok elemre. Jogosultságok létrehozásához kattintson a jobb egérgombbal a Jogosultságok elemre, majd válassza az előugró menü Új menüpontját.

### Jogosultsági csoportok

Az EIP biztosít néhány alapértelmezett jogosultsági csoportot. A jogosultsági csoportok kapcsolódó jogosultságok gyűjteményei. Az EIP adminisztráció nevű jogosultsági csoporthoz például hozzárendelhetők az EIP rendszer adminisztrálásával kapcsolatos öt alábbi jogosultság:

- EIPAdminServer
- EIPAdminEntity
- EIPAdminTextEntity
- EIPAdminTemplate
- EIPAdminInfoMining

A jogosultsági csoportokhoz előre meghatározott jogosultságok megjelenítéséhez vagy módosításához bontsa ki a Hitelesítés ikont, majd kattintson duplán a jogosultsági csoport nevén. Jogosultsági csoportok létrehozásához kattintson a jobb egérgombbal a Jogosultsági csoportok elemen, majd válassza az előugró menü Új menüpontját.

### Jogosultságkészletek

Az EIP számos alapértelmezett jogosultságkészletet biztosít. A jogosultságkészletek felhasználói szerepeket meghatározó jogosultságok gyűjteményei. A ClientUserCreateAndDelete nevű jogosultságkészlet például az ügyfél végfelhasználói szerepekhez köthető 17 jogosultságot tartalmazza, például a Delete (elem törlése) és az ItemAdd (elem hozzáadása), stb. jogosultságokat. Ha létrehoz egy felhasználói azonosítót, amelyhez hozzárendeli a ClientUserCreateAndDelete jogosultságkészletet, akkor a felhasználó a tartalomkiszolgálóra bejelentkezve a jogosultságkészletben meghatározott 17 felhasználói szerep bármelyikét elvégezheti. A jogosultságkészletek megtekintéséhez vagy módosításához bontsa ki a Hitelesítés

kategóriát, kattintson a Jogosultságkészletek elemre, majd kattintson duplán egy jogosultságkészlet nevére. Jogosultságkészletek létrehozásához kattintson a jobb egérgombbal a Jogosultságkészletek elemen, majd válassza az előugró menü Új menüpontját.

A felhasználók és csoportok felügyelete során a jogosultságkészleteket hozzárendelheti a felhasználókhoz és/vagy felhasználói csoportokhoz. Ha egy jogosultságkészletet felhasználói csoportnak oszt ki, akkor a csoport valamennyi tagja végrehajthatja a jogosultságkészletben meghatározott felhasználói szerepeket.

---

## Termék nézetek és adatbázisok átkapcsolása

Ha a vállalati megoldásnak a Content Manager és az Enterprise Information Portal is része, akkor a rendszeradminisztrációs ügyfél segítségével mindkét termék kezelhető. A korábbi kiadásokban ha mindkét termék telepítve volt, akkor két ügyfelet kellett megnyitnia. Az egyik ügyfél nézetről a másikra váltás kényelmes megoldást biztosít az információk módosításához, illetve gyors elérést nyújt mindkét termékhez.

Ha kijelentkezés nélkül kíván átkapcsolni az egyik termékről a másikra, akkor a rendszeradminisztráció főablakából válassza ki a bal oldali ablakrész feletti menüből a megfelelő termék nevét.

Egyesített adatbázisok közötti átkapcsoláshoz kattintson duplán a megfelelő egyesített adatbázis ikonjára a bal oldali ablakrészben.

Az ügyfélből különböző adatbázisok is kezelhetők kilépés és az új adatbázisba való bejelentkezés nélkül. Az adminisztrációs ügyfél a cmbds.ini fájlban felsorolt valamennyi adminisztrációs adatbázishoz megjelenít egy ikont. Másik adatbázisra átkapcsoláshoz kattintson a megfelelő ikonra. Ha az új adatbázis az aktuális bejelentkezéshez megadottól eltérő felhasználói azonosítót használ, akkor a program felszólítja egy felhasználói azonosító és jelszó megadására.

---

## Az adminisztrációs ügyfél javításai és bővítései

Az EIP 8.2 változata jelentős fejlesztéseket hozott az EIP adminisztrációs ügyfélben, egyebek között a következőket:

### Tökéletesített varázslók és párbeszédablakok

Az új párbeszédablakok egyszerűbbé teszik a felhasználók kezelését. Az egyesített példányok és keresési sablonok létrehozását új varázslók könnyítik. A felhasználók ettől függetlenül használhatják az EIP 7.1 változatában megszokott párbeszédablakokat is.

### Megosztott adminisztrációs ügyfél

Az EIP 8.2 és a Content Manager 8.2 változatának azonos rendszerre telepítésekor a két termék azonos adminisztrációs ügyfelet használ. Ha mindkét terméket azonos személy adminisztrálja, akkor elég egyszer is bejelentkezni, a két alkalmazás utána már egyszerűen átkapcsolható. Az adminisztrációs adatbázisok között is átválthat kijelentkezés nélkül.

### Tartomány adminisztrátorok

Lehetőség van olyan tartomány adminisztrátorok létrehozására, akik csak egy adott tartomány vonatkozásában jogosultak adminisztrátori teendők elvégzésére.

### Egyszeres bejelentkezés és LDAP támogatás

Az EIP jelenlegi változata Windows Active Directory és az LDAP kiszolgálók alkalmazásával lehetővé teszi a felhasználók egyszeri bejelentkezését több tartalomkiszolgálóra.

---

## Adminisztrációs ügyfél csatlakoztatása helyi adminisztrációs adatbázishoz

Ha az adminisztrációs adatbázist az adminisztrációs ügyféllel megegyező számítógépre telepíti, akkor a helyi ügyfél és a kiszolgáló csatlakozásához szükséges információk már megtalálhatók az adatbázis kapcsolati információkat tároló cmbds.ini fájlban. Nem kell telepítés utáni konfigurációs lépéseket végrehajtani, hanem azonnal csatlakozhat.

**Követelmény:** Ha további helyi adatbázisokat hoz létre az EIP adatbázis telepítő segédprogramjával, akkor a cmbds.ini fájlban módosítani kell a szükséges információkat, mielőtt használhatná az új adatbázist.

1. Kattintson a **Start → Programok → Enterprise Information Portal for Multiplatforms 8.2 → Adminisztráció** menüpontra.
2. A Kiszolgáló mező legördülő listájában válassza ki a helyi adatbázist.
3. Írja be az adminisztrátori felhasználói azonosítót és a jelszót, majd nyomja meg az OK gombot.
4. Megjelenik a rendszeradminisztrációs ügyfél. **Tipp:** Ha használta az EIP Első lépéseket, akkor az ügyfél bal oldalán a minta adatbázisok láthatók.

---

## Adminisztrációs ügyfél csatlakoztatása távoli adminisztrációs adatbázishoz

Az EIP adminisztrációs ügyfél kétféleképpen csatlakoztatható távoli AIX, Windows vagy Solaris adatbázishoz:

- Csatlakozás RMI kiszolgálón keresztül (további információkat az *Information Integrator for Content tervezése és telepítése* című kiadványban talál).
- Kapcsolat meghatározása az adatbázis katalogizálásával a DB2 konfigurációs segédletben, illetve a kiszolgáló kapcsolati paramétereinek meghatározásával az EIP kiszolgáló konfigurációs segédprogramban. A segédprogram bemásolja a csatlakozáshoz szükséges információkat, például az adatbázis sémájának nevét, az álnevet és operációs rendszert a cmbds.ini fájlba. A rendszeradminisztrációs ügyfél indításakor a választható kiszolgálókat a cmbds.ini fájl bejegyzései határozzák meg.

**Követelmény:** Minden távoli adatbázist külön-külön kell katalogizálni. A távoli adatbázisoknak benne kell lenniük a cmbds.ini fájlban ahhoz, hogy csatlakozhasson hozzájuk az adminisztrációs ügyféllel.

**Tipp:** A haladó felhasználók kihagyhatják a Kiszolgáló konfigurációs segédprogramot, és közvetlenül is szerkeszthetik a cmbds.ini fájlt. A cmbds.ini alapértelmezett elérési útja a C:\Program Files\IBM\CMgmt könyvtár.

**Fontos:** Ha a termék telepítését végző felhasználó már meghatározta a csatlakoztatni kívánt távoli adatbázis katalogizálási értékeit, akkor ezen az adatbázison nem kell elvégezni a DB2 CCA lépéseket. Ha azonban a telepítést végző nem adta meg az adatbázis katalogizálási értékeit, vagy egy újabb távoli adatbázishoz kíván csatlakozni, akkor használnia kell a DB2 CCA segédprogramot, a cmbds.ini fájlt pedig frissíteni kell a további adatbázis kapcsolati paramétereivel.

### 1. lépés - Távoli adatbázis katalogizálása a DB2 konfigurációs segédlettel

A DB2 konfigurációs segédlet (CCA) katalogizálja a távoli EIP adatbázist a DB2 kiszolgálón. A távoli adatbázis katalogizálásához ismernie kell a távoli kiszolgáló hosztnevét, az adatbázis nevét és az adatbázis példány portszámát, emellett meg kell határoznia egy álnevet a távoli adatbázishoz.



Az 1a - 1f lépések az adatbázisnév, a sémanév és a kapcsolati portszám megkeresését írják le. Az adminisztrációs ügyfél és a távoli adatbázis között kapcsolat beállításához ismernie kell a neveket és a csatlakozási portszámokat.

1. Távoli adatbázis csatlakozási információk megkeresése:
  - a. Jelentkezzen be a távoli AIX, Windows vagy Solaris kiszolgálóra egy DB2 adminisztrációs jogosultságokkal rendelkező felhasználói azonosítóval.
  - b. Írja be a `db2 list db directory` parancsot.
  - c. Válassza ki az adminisztrációs adatbázis nevét, amelyhez csatlakozni szeretne. Mivel a különböző példányok eltérő portszámokkal rendelkezhetnek, fel kell jegyezni az adatbázist futtató DB2 példány nevét.
  - d. Írja be a `db2 connect to <adatbázis> user <felhasználói_azonosító> using <jelszó>` parancsot.
  - e. Írja be a `db2 list tables` parancsot, és jegyezze fel az adatbázis sémájának nevét (erre a konfigurációs segédprogramnak van szükség).
  - f. Jegyezze fel a távoli adminisztrációs adatbázishoz tartozó csatlakozási portszámot.
    - Windows platformon:
      - 1) Nyissa meg a DB2 Irányítóközpontot a távoli Windows kiszolgálón.
      - 2) Kattintson a jobb egérgombbal a helyi számítógép valamelyik példányára.
      - 3) Válassza az előugró menü Kommunikáció beállítása... menüpontját. .
      - 4) Kattintson a TCP/IP beállítás melletti Tulajdonságok gombra. Ebben az ablakban jelenik meg a portszám.
    - AIX és Solaris platformon:
      - 1) Írja be a `cd /usr/etc` parancsot.
      - 2) Írja be a `cat services` parancsot.
      - 3) Görgesse végig a szolgáltatások listáját, amíg meg nem találja a távoli adatbázis példányának portszámát. Ha például az adatbázis a `db2inst1` példányon található, akkor a portszám lehet `50000`.
      - 4)
2. A DB2 konfigurációs segédlettel katalogizálja a távoli adatbázist. További információkat a DB2 CCA súgójában talál.
  - a. Jelentkezzen be a Windows kiszolgálóra, amelyen az adminisztrációs ügyfél telepítve van. A bejelentkezést teljes DB2ADM jogosultságokkal rendelkező felhasználóként kell megejteni.
  - b. A Start → Programok menüben keresse ki a DB2 konfigurációs segédletet.
  - c. Kövesse a DB2 konfigurációs segédlet útmutatásait a távoli adatbázis katalogizálásához és a kapcsolat ellenőrzéséhez.
  - d. Ha a DB2 CCA által végzett csatlakozási teszt sikerült, akkor kövesse a “2. lépés - Kiszolgáló konfigurációs segédprogram használata” helyen leírtakat, vagy módosítsa közvetlenül a távoli adatbázis kapcsolatra vonatkozó csatlakozási paramétereket a `cmbds.ini` fájlban.

## 2. lépés - Kiszolgáló konfigurációs segédprogram használata

A kiszolgáló konfigurációs segédprogram megkérdezi a távoli adatbázisra vonatkozó csatlakozási információkat (portszám, hosztnév, stb.), majd tárolja a megadott adatokat a `cmbds.ini` fájlban.

1. Kattintson a **Start → Programok → IBM Enterprise Information Portal for Multiplatforms → Kiszolgáló konfigurációs segédprogram** menüpontra.



2. Töltse ki a mezőket (lásd: 4. táblázat:).

4. táblázat: Kiszolgáló konfigurációs segédprogram

Mező	Információk	Megjegyzések
Kiszolgáló	Válassza ki az adatbázis típusát: Content Manager vagy EIP.	A kiszolgáló az adatbázis típusát jelenti, nem az adatbázisnak otthont adó kiszolgáló nevét. <b>Tipp:</b> Csak akkor kezelheti mindkét adatbázist egyazon adminisztrációs ügyféllel, ha a számítógépre az EIP és Content Manager adminisztrációs ügyfél is telepítve van.
Kiszolgáló neve	A használni kívánt adatbázis álneve. Követelmény: A DB2 CCA-ban megadottal megegyező álnevet kell használni.	Az álnev biztosítja a távoli adatbázis egyedi azonosítását a munkaállomáson. Az álnevek legfeljebb 8 karakterből állhatnak. Ha például a távoli adatbázis neve ICMNLSDB, akkor egy lehetséges álnev a TAVOLI1.
Séma neve	Írja be a távoli adatbázis létrehozásakor hozzárendelt sémanevet.	Az EIP és Content Manager adatbázisok alapértelmezett sémaneve ICMADMIN.
Hosznév	Írja be a távoli adatbázisnak otthont adó számítógép nevét.	A távoli adatbázis számítógépének megadása történhet teljes képzésű hosznévvel és IP címmel is.
Operációs rendszer	Válasszon ki egy operációs rendszert a legördülő listából.	A választható lehetőségek az AIX, a Sun Solaris és a Windows. Az OS/390 beállítás nem működik az EIP 8.2 változatában.
Portszám	Írja be a távoli adatbázishoz rendelt port számát.	A Windows, AIX vagy Solaris alatt futó EIP és Content Manager adatbázisok alapértelmezett csatlakozási portja az 50000.
Távoli adatbázis neve	Írja be a távoli adatbázis nevét. Használjon nagybetűket.	Az EIP és Content Manager adatbázisok alapértelmezett neve ICMNLSDB.
Csomópontnév	Írja be a távoli EIP vagy Content Manager adatbázis csomópontnevét.	A csomópontnév a távoli adatbázishoz hozzárendelt egyedi név, a távoli adatbázishoz rendelt álnevhez hasonlóan. Windows, AIX vagy Solaris kiszolgálón található adatbázis csomópontnevének megkeresése: a. Nyisson meg egy DB2 parancssort. b. A db2=> parancssorba írja be a LIST NODE DIRECTORY parancsot. c. A DB2 megjeleníti a távoli kiszolgálón található vagy meghatározott adatbázisok csomópontneveit és egyéb adatait.
Egyszeri bejelentkezés engedélyezése	Válassza ki a beállítást, ha az egyszeri bejelentkezés engedélyezett volt az adatbázis telepítésekor.	Az alapértelmezett beállítás a tiltott.

4. táblázat: Kiszolgáló konfigurációs segédprogram (Folytatás)

Mező	Információk	Megjegyzések
Biztonsági lehetőségek	Válassza ki az Ügyfél hitelesítés beállítását, ha ez engedélyezve volt az adatbázis létrehozása során is.	Az alapértelmezett beállítás a Kiszolgáló.

3. Kattintson az OK gombra.

### 3. lépés - Távoli adatbázis kapcsolatának ellenőrzése

1. Jelentkezzen be a Windows kiszolgálóra, amelyen az adminisztrációs ügyfél telepítve van.
2. Kattintson a **Start → Programok → Enterprise Information Portal for Multiplatforms 8.2 → Adminisztráció** menüpontra.
3. A Kiszolgáló mező legördülő listájában válassza ki a távoli adatbázis álnévét. A név megegyezik a Kiszolgáló konfigurációs segédprogramban és a DB2 konfigurációs segédletben beállított álnévvel.
4. Írja be a távoli adatbázishoz tartozó felhasználói azonosítót és jelszót.
5. Kattintson az OK gombra. Megjelenik az adminisztrációs ügyfél.

## Dokumentumtípusok meghatározása

Az EIP bizonyos dokumentumtípusokhoz rendelkezik megjelenítő támogatással. Ha a kiszolgálón meghatározza a dokumentumtípusokat, akkor a dokumentumokat a létrehozó alkalmazásban nyithatja meg. Ha például Lotus Word Pro dokumentumokat tárol a Content Manager kiszolgálón, akkor az EIP környezetet beállíthatja úgy, hogy az .lwp kiterjesztésű dokumentumokat az ügyfél dokumentum megjelenítője helyett a Lotus Word Pro alkalmazással nyissa meg.

Dokumentumtípus meghatározásához módosítsa a CMBROOT (általában az x:\Program Files\IBM\CMgmt) könyvtárban található cmbcc2mime.ini fájlt. Az egyéni MIME meghatározásokról a fájlban talál információkat. A fájl lefordítja a tartalomosztályokat MIME-típus folyammá, így a tartalomkiszolgálók adatfolyamai olvashatóvá válnak az ügyfelek számára.

**Fontos:** Amikor egy alkalmazást a MIME típus alapján indít el, akkor csak az alapobjektum jelenik meg. A dokumentumhoz készült leírások nem jelennek meg. Ha a dokumentum több részből áll, akkor csak az első rész jelenik meg. A MIME típusnak mindkét fájlban meg kell egyeznie.

### A kiszolgáló MIME típus fájljának (cmbcc2mime.ini) módosítása

Kiszolgáló MIME típusok hozzáadásakor győződjön meg róla, hogy a hozzáadott dokumentumtípus egy olyan MIME típus, amely ehhez a fájlhoz lett létrehozva. További információkat az <ftp://ftp.isi.edu/in-notes/iana/assignments/media-types> fájlban talál.

Ha új értékeket kíván hozzáadni a cmbcc2mime.ini fájlhoz, akkor tegye a következőket:

1. Nyissa meg a cmbcc2mime.ini fájlt egy szövegszerkesztőben.
2. Az egyéni értékekhez használja a következő formátumot:
  - A tartalomosztály 4096-tal kezdődik
  - A tartalomosztály értéket az egyenlőségjel (=) követi

- Az egyenlőségjelet a MIME típus követi. Ha ez nem a tartalomosztály szabványos MIME típusa, akkor tegye a következőket:
  - a. A MIME típusok egy típusból és egy altípusból állnak. Az érvényes típusok a következők: application, text, image, model, message, audio és video
  - b. A típust egy osztásjel (/) követi
  - c. Az altípus létrehozásához az dokumentumhoz használt jelsort az x- jelsornak kell megelőznie, például:  
         x-dokumentumosztály (4096=application/x-dokumentumosztály)
- Ha szükséges, minden új MIME típusnál ismételje meg a következő lépéseket: 2b és 2c.

**Tipp:** Az OnDemand a tartalomosztályok számértékei helyett a fájlok kiterjesztéseit képezi le MIME típus adatfolyamokra.



---

## 3. fejezet Az EIP adminisztrációs ügyfél szolgáltatásainak használata

Ez a szakasz írja le az EIP adminisztrátorok néhány általános feladatát.

---

### Egyesített keresés létrehozása

Az egyesített keresések olyan lekérdezések, amelyeket az ügyfélalkalmazások adnak ki, és egyszerre egynél több tartalomkiszolgálón is futhatnak. Az EIP biztosítja az egyesített keresésekhez használható keresési sablonok létrehozásához szükséges eszközöket. Mivel minden egyes tartalomkiszolgáló másként tárolja és szervezi az információkat, a keresési sablonnak tükröznie kell a kiszolgálók közötti különbségeket. A keresési sablon a tartalomkiszolgálók tartalmának kereséséhez az egyesített példányokat és ezek egyesített attribútumait eredeti attribútumokra képezi le.

Az egyesített keresések létrehozása a következőkből áll:

- Tartalomkiszolgálók kapcsolatainak meghatározása az EIP kapcsolatokkal.
- Egyesített példányok létrehozása.
  - Egyesített példányok meghatározása.
  - Egyesített attribútumok létrehozása.
  - Egyesített attribútumok leképezése az eredeti attribútumokra.
  - Paraméterek hozzárendelése.
- Keresési sablonok létrehozása.
  - Keresési sablon meghatározása.
  - Keresési feltételek meghatározása.
  - Sablon beállításainak meghatározása.
  - Ügyfél felhasználók hozzáféréseinek beállítása.

Az EIP 8.2 változata két varázslót biztosít az egyesített példányok és keresési sablonok egyszerűbb létrehozásához. Az egyesített példány varázsló tartalmaz egy szűrhető kiszolgáló tárolót, amely leegyszerűsíti az eredeti attribútumok keresését. Emellett az egyesített attribútumokat alapértelmezett paraméterekkel látja el, lecsökkentve a téves beállítások lehetőségét. A keresési sablon varázsló segítséget nyújt új keresési feltételek létrehozásához. Emellett segít a keresési feltételek és eredmények megjelenésének tervezésében is. Még azt is megtekintheti, hogy a keresési sablon hogy nézne ki az ügyfélalkalmazásban. Emellett az egyesített példányok és keresési sablonok az EIP 7.1 változatában megszokott párbeszédablakok segítségével is létrehozhatók.

Az EIP online súgóiban valamennyi varázsló, párbeszédablak és mező le van dokumentálva.

---

### Kiszolgálók meghatározása

A kiszolgálókat meg kell határozni, mielőtt csatlakozni lehetne hozzájuk, és műveleteket lehetne végrehajtani rajtuk. Kattintson a jobb egérgombbal a Kiszolgáló ikonra, majd válassza az előugró menü Új menüpontját az EIP által támogatott összes kapcsolat megjelenítéséhez. Kiszolgálók meghatározása előtt a kapcsolatokról rendelkezni kell bizonyos alapvető információkkal:

- Milyen kapcsolatok kerültek kiválasztásra a telepítéskor? A telepített kapcsolatokat a cmbscs.ini konfigurációs fájl sorolja fel. Windows számítógépeken az alapértelmezett

elérési út az `x:\Program Files\IBM\CMgmt`. AIX vagy Solaris rendszereken a `cmbcs.ini` fájlok helyéről érdeklődjön az adminisztrátornál.

- A telepítéskor helyi vagy távoli kapcsolatok lettek megadva? A `cmbcs.ini` fájl tartalmazza a helyi vagy távoli kapcsolati típusokat.
- Ha a rendszer RMI kiszolgáló használatára van beállítva, akkor az RMI kiszolgáló elindult? Ha el kívánja indítani az RMI-t a helyi RMI kiszolgálón, akkor válassza a **Start → Programok → IBM Enterprise Information Portal for Multiplatforms 8.2 → RMI kiszolgálók indítása** menüpontot. Ha a rendszer távoli RMI kiszolgálót használ, akkor a telepített RMI kapcsolatokkal rendelkező távoli kiszolgálóra vonatkozó információk a `cmbsvclient.ini` fájlban találhatók. További információkat az RMI kiszolgáló adminisztrátorától tudhat meg.
- Ha az EIP telepítését végző felhasználó telepítette a CM for AS/400 kapcsolatot, akkor milyen információk kerültek bele a `frnolint.tbl` hálózati táblába? Az AS/400 `frnolint.tbl` fájl a `%CMBROOT%` könyvtárban található.
- A relációs adatbázisokat tároló távoli tartalomkiszolgálók (például Content Manager v8, DB2, DataJoiner és Information Catalog) beállításakor az adatbázis katalogizálását arról a munkaállomásról kell végeznie, amelyről az ügyfelet használja.

A kiszolgálók meghatározásának általános lépései a következők:

1. Kattintson a jobb egérgombbal a **Kiszolgálók** elemre, majd válassza az előugró menü **Új** menüpontját.
2. Válasszon ki egy kiszolgálót a listából. Megjelenik az Új kiszolgáló ablak.
3. Az Általános lap Kiszolgáló neve mezőjébe írja be a kiszolgáló nevét és leírását. Bizonyos kiszolgálótípusok esetén csak az adatbázis nevét kell megadni. Más kiszolgálóknál az adatbázist tároló kiszolgáló teljes képzésű nevét meg kell adni.
4. Ha szükséges, akkor adja meg az inicializálási paramétereket. Bizonyos kiszolgálók inicializálási paramétereket, például csatlakozási karaktersorozatot és konfigurációs karaktersorozatot igényelnek. Más kiszolgálóknál csak az adatbázis neve szükséges.
5. Kattintson a Kiszolgáló kapcsolat tesztelése gombra. Az EIP bejelentkezik a megadott felhasználói azonosítóval és jelszóval bizonyos kiszolgálókra az adminisztrációs ügyfél elindításához. Ha a kiszolgáló eltérő felhasználói azonosítót és jelszót igényel, akkor az EIP felszólítja egy érvényes felhasználói azonosító és jelszó megadására.

**Tipp:** Az előre meghatározott tartalomkiszolgálóktól eltérő típusú kiszolgálókat is beállíthat, ilyenkor azonban meg kell adni a Java vagy C++ kapcsolati osztályokat, valamint az új kiszolgálótípus kiszolgálómeghatározási osztályát. A kiszolgáló tároló futtatásához szükség van egy Java kapcsolatra is. A tartalomkiszolgálók hozzáadására vonatkozó útmutatásokat a *Munkaállomás alkalmazás programozói kézikönyv* és az online API kézikönyv kiadványokban találja.

Ha a tartalomkiszolgáló beállítása nem járt sikerrel, akkor nézze meg az *Üzenetek és kódok* kiadványban a kapott hibaüzenetre vagy problémás helyzetre vonatkozó hibaelhárítási információkat.

Érdemes lehet továbbá konzultálni a kiszolgáló adminisztrátorával.

## Kiszolgálók meghatározásának irányelvei

Ez a szakasz tárgyalja a kiszolgálók kezdeti meghatározásának irányvonalait.

### Csatlakozás DB2 (relációs) adatbázisokhoz

Ez a szakasz a DB2, DataJoiner, JDBC, ODBC, Information Catalog illetve a Content Manager v7 és v8 kiszolgálókra vonatkozik.

- **Fontos:** A kiszolgáló meghatározása előtt valamennyi DB2 adatbázist katalogizálni kell. Az adatbázisok katalogizálására a DB2 CCA vagy a DB2 parancssor használható. További információkat a DB2 adminisztrátortól szerezhet.
- Az Általános lap Kiszolgáló neve mezőjébe írja be a csatlakoztatni kívánt adatbázis nevét. A kiszolgálónév meghatározásához használjon nagybetűket.
- DB2, DataJoiner, JDBC, ODBC és Information Catalog meghatározásakor kattintson az Inicializálási paraméterek lapra, és írja be a csatlakoztatni kívánt adatbázis tábláihoz társított sémanevet, például SCHEMA=ICMADMIN.
- Content Manager 7.1 vagy 8.2 kiszolgáló meghatározásakor csak az adatbázis nevét kell megadni. Az Inicializálási paraméterek lap alapértelmezett beállításait ne módosítsa.
- Content Manager 7.1 kiszolgáló meghatározásakor rendelkeznie kell egy frnlint.tbl nevű hálózati tábla fájlal az x:\CMBROOT helyi könyvtárban. A hálózati tábla tartalmazza az EIP által a Content Manager 7.1 könyvtárkiszolgáló eléréséhez megkövetelt információkat, például a hosztnevet, portszámot és kiszolgálótípus azonosítást. Ha több Content Manager 7.1 kiszolgálóval rendelkezik, akkor minden kiszolgálóhoz külön bejegyzést kell létrehozni a frnlint.tbl fájlban a kiszolgáló meghatározása előtt.
- DB2 DataJoiner kiszolgálóra csatlakozáskor győződjön meg róla, hogy a DB2 Universal Database kiszolgálón az Enterprise Information Portal számára megadott hitelesítési módszer a kiszolgáló.
- A DataJoiner 2.1 változatára csatlakozáshoz le kell tölteni egy kötés programot a DataJoiner webhelyről, és a DataJoiner kiszolgáló meghatározása előtt el kell végezni a DataJoiner adatbázis kötését.

## Csatlakozás szöveg keresési kiszolgálóhoz

Szöveg keresési kiszolgáló meghatározásához először meg kell határozni a Content Manager 7.1 kiszolgálót, amelyhez a szöveg keresési kiszolgáló társítva van.

A szöveg keresési kiszolgáló nevét a "Társított Content Manager 7.1 kiszolgáló nevének kiválasztása" mezőbe kell beírni. A mező a Társított kiszolgáló lapon található.

A Content Manager 7.1 és szöveg keresési kiszolgálónak futnia kell, mielőtt az EIP csatlakozhatna hozzájuk.

## Csatlakozás több Content Manager for AS/400 kiszolgálóhoz

Ha egynél több AS/400 kiszolgálót használ, akkor a további kiszolgálókat a hálózati táblában kell meghatározni. A (frnlint.tbl) hálózati tábla az x:\<cmbroot> könyvtárban található. Az új kiszolgálónak adja meg a nevét, kapcsolattípusát (például TCP/IP), hosztnevét és kiszolgálótípusát. Az első kiszolgálónál a telepíté adja meg a kiszolgáló, hosztnév és portszám értékeket a frnlint.tbl létrehozásához.

A frnlint.tbl fájlban tárolt információk az alábbiakhoz hasonlítanak:

```
/* VI/400 hálózati tábla */
SERVER: VI400 REMOTE TCPIP
        HOSTNAME = vi400
        PORT      = 29000
        SERVER_TYPE = FRNLS400
```

## Extended Search kapcsolat beállítása

Az Extended Search kiszolgáló meghatározásához megadott információkat két tényező határozza meg:

- A webkiszolgáló típusa, amelyen az Extended Search kiszolgáló telepítve van: Domino, WebSphere vagy IIS.
- Az Extended Search kiszolgálónak otthont adó webkiszolgáló által használt portszám.

Extended Search kapcsolat meghatározásakor tegye a következőket:

1. Az Általános lap Kiszolgáló neve mezőjében adja meg az Extended Search kiszolgálónak otthont adó webkiszolgáló teljes képzésű hosztnevét.
2. Az inicializálási paraméterek lapon adja meg a 80-as portszámot, ha az Extended Search telepítése az alapértelmezett webkiszolgáló port beállításával történt.
3. Az alkalmazás azonosítója mezőben adja meg a Demo értéket. A nevet a megjelenő formában kell megadni.
4. A Jelszó mezőben adja meg a Demo értéket.
5. A további paraméterek mezőben:
  - a. Ne módosítsa a két pontosvesszőt ha az Extended Search Domino webkiszolgálón került telepítésre, és a webkiszolgáló valamint az Extended Search kiszolgáló telepítése alapértelmezett portszám beállításokkal történt.
  - b. Az egyéni beállításokkal telepített Extended Search kiszolgálók további paramétereinek módosításával kapcsolatos tudnivalókat az alábbi szakasz tartalmazza.

Ha az Extended Search kapcsolatot WebSphere alkalmazáskiszolgálóhoz állította be, ha az Extended Search portszáma nem 6001, vagy ha az Extended Search kiszolgáló neve nem azonos a webkiszolgáló nevével, akkor a kapcsolatot be kell állítania az ES kiszolgáló kisalkalmazás, az ES portszám és az ES kiszolgálónév megtalálásához.

Ha az Extended Search kiszolgálót WebSphere alkalmazáskiszolgálón állította be, ha az Extended Search portszám nem 6001, vagy ha az Extended Search kiszolgáló neve nem egyezik meg a webkiszolgáló nevével, akkor az alábbi lépések elvégzésével hozhat létre egy ideiglenes konfigurációs fájlt, például `desclient.cfg` néven:

Állítsa a könyvtárat arra a könyvtárra, amelyben az alkalmazások vagy példák találhatók. Hozzon létre egy konfigurációs fájlt például `desclient.cfg` néven. Ez a fájl nem képezi az Enterprise Information Portal részét.

A `desclient.cfg` fájlhoz adja hozzá a `DESHOSTNAME=(ES_hosztnev)` sort.

Ha az Extended Search Domino alkalmazáskiszolgálón van beállítva, akkor adja meg a `DESREQURI=/lotuskms/ESAdmin` utasítást. Ha az Extended Search WebSphere alkalmazáskiszolgálón található, akkor állítsa be a `DESREQURI=/servlet/ESAdmin` értéket, ahol a *servlet* a HTTP kiszolgálónak az Extended Search kapcsolatot tartalmazó könyvtára.

Ha az alkalmazáskiszolgáló WebSphere, akkor a `DESREQURI` paraméternek a `/lotuskms/ESAdmin` helyett a `/servlet/ESAdmin` értéket kell felvennie.

Ha az ES forrásokat egyszerűsített/teljes ügyfelekről szándékozik keresni, akkor adjon hozzá egy további

`"DESCFGPATH=<desclient.cfg abszolút elérési útja>"`

paramétert az adminisztrációs ügyfél kiszolgáló meghatározási párbeszédablakában.

Ha futtatni kívánja az ES példákat, akkor a `desclient.cfg` abszolút elérési útját parancssori paraméterként adja át.

1. példa:

```
TConnectDES es.stl.ibm.com user password
PORT=80;DESAPPID=Demo;DESAPPPW=password;DESCFGPATH
=<desclient.cfg abszolút elérési útja>;
```

2. példa:



```
java TConnectDES es.stl.ibm.com user password  
PORT=80;DESAPPID=Demo;DESAPPPW=password;DESCFGPATH  
=<desclient.cfg abszolút elérési útja>;
```

### **Information Catalog kiszolgáló meghatározása**

A kiszolgáló meghatározása előtt katalogizálni kell az Information Catalog kiszolgálót. A Kiszolgáló neve mezőben adja meg a kiszolgáló nevét, például **SAMPLE1**. Az inicializálási paraméterek lapon írja be a **SCHEMA=<SAMPLE1-hez társított séma neve>** értéket.

### **OnDemand kiszolgáló meghatározása**

Az OnDemand kiszolgáló meghatározása előtt az OnDemand kiszolgálónak és a könyvtárkiszolgáló démonnak futnia kell. A kiszolgáló és a démon futásának ellenőrzéséhez pingelje meg a kiszolgálót, mielőtt beállítaná az EIP környezetben.

Az Általános lapon adja meg az OnDemand kiszolgáló teljes képzésű hosztnevét.

Az Inicializálási paraméterek lapon adja meg az OnDemand kiszolgáló telepítése során beállított portszámot. Ha az OnDemand telepítésekor a 0 alapértelmezett portszám került megadásra, akkor a Portszám mezőben adja meg a 0 értéket. Ha a telepítés során más portszám lett megadva, akkor adja meg a portszámot előtte egy # karakterrel. Például: #5000.

Ha az OnDemand kiszolgáló az OS/400 V4 kiadását futtató AS/400 rendszeren került telepítésre, akkor a További paraméterek mezőben meg kell adni a **STATECONNECT=#1** paramétert.

Ha az OnDemand kiszolgáló az operációs rendszer 2.1 kiadását futtató S/390 rendszeren került telepítésre, akkor adja meg az OnDemand telepítésekor megadott egyéni portszámot.

Az OnDemand megköveteli a socket fenntartását a kapcsolat során.

### **Szöveg keresési kiszolgáló meghatározása**

Szöveg keresési kiszolgáló meghatározásához először meg kell határozni a Content Manager v7 kiszolgálót, amelyhez a szöveg keresési kiszolgáló társítva van.

Az Általános lap Kiszolgáló neve mezőjében adja meg a szöveg keresési kiszolgáló nevét. A Társított kiszolgáló lap legördülő listájából válassza ki a szöveg keresési kiszolgálóhoz tartozó Content Manager 7.1 kiszolgálót.

A Content Manager 7.1 és szöveg keresési kiszolgálónak futnia kell, mielőtt az EIP csatlakozhatna hozzájuk.

### **Domino.Doc kiszolgáló meghatározása**

A Kiszolgáló neve mezőben adja meg a Domino.Doc kiszolgáló nevét és könyvtárnevét. Például: oakley/DominoDoc1/Lib.nsf.

Helyi kapcsolatok használatakor a Domino.Doc Desktop Enabler terméket telepíteni kell a munkaállomásra, amelyen az EIP ügyfél is telepítve van. RMI használata esetén a Domino.Doc Desktop Enabler terméket az RMI kiszolgálón kell telepíteni. A Domino.Doc Desktop Enabler termék verziószámának meg kell egyeznie a Domino.Doc kiszolgáló verziószámával.

Az Inicializálási paraméterek lapon ne módosítsa a két pontosvesszőt.

### **ImagePlus for OS/390 kiszolgáló meghatározása**

ImagePlus for OS/390 kiszolgáló esetén a csatlakozáshoz az alábbi paraméterek szükségesek. Az alábbi lista példa értékeket tartalmaz:

- FAF portszám: 3061
- FAF alkalmazásazonosító: 01
- FAF protokoll: 4000
- FAF IP cím: 9.67.43.83
- Objektum terjesztéskezelő CICS: 4000
- Objektum terjesztéskezelő IP cím: 9.67.43.83
- Objektum terjesztéskezelő portszám: 3082
- Objektum terjesztéskezelő terminálazonosító: *hagyja üresen a mezőt*
- További paraméterek: *FAFSITE=CS61;;*

## Nyomkövetés használata Content Manager ImagePlus for OS/390 kiszolgálókkal

A nyomkövetés segíthet megoldani a problémákat, ha nem tud csatlakozni a Content Manager ImagePlus for OS/390 server kiszolgálóhoz. Ha telepíti a Content Manager ImagePlus for OS/390 kapcsolatot, akkor az ImagePlus for OS/390 nyomkövetés bekapcsolásához módosítsa a cmbroot könyvtárban található *eypapi.ini* fájlt.

Az *eypapi.ini* fájl a következő sorokat tartalmazza:

```
; IPFAF fájllok tárolási útvonala
; (Meg KELL adni a befejező '\' jelet)
; -- alapértelmezés a <gyökérkönyvtár>\
;
IPFAFPath=d:\cmbroot
; Naplózási jelző (EYPLhnn.LOG fájllok)
; -- alapértelmezésben a naplózás KI van kapcsolva (0)
; -- 0 Minden naplózás KIKAPCSOLVA
; -- 1 Naplófájlok létrehozása, hibahelyzetek naplózása
; -- 2 Naplófájlok létrehozása, minden helyzet naplózása
;
Logging = 0

;-----
;
; API-k segítségével létrehozott FAF paramétertípusok naplózása
; -- alapértelmezésben a naplózás KI van kapcsolva (0)
; -- 0 Paramétertípusok nem kerülnek naplózásra
; -- 1 FAF paramétertípusok naplózása
;
FafTypeLogs = 0
```

### IPFAFPath

Meghatározza a naplófájlok könyvtárát. A naplófájlok neve:

*EYPhhnn.LOG*

ahol a *hhnn* a naplófájl létrehozásának hónapja és napja.

### Logging

Meghatározza, mikor kerül sor naplófájl létrehozására.

- 0** Nincs naplózás. Az alapértelmezett beállítás a 0.
- 1** A létrejött naplófájlok csak hibákat tartalmaznak.
- 2** A létrejött naplófájlok minden helyzetet tartalmaznak.

### FafTypeLogs

Megadja az API-k segítségével létrehozott FAF paramétertípusok naplózását.

- 0** A paramétertípusok nem kerülnek naplózásra. Ez az alapértelmezett beállítás.

## OnDemand kapcsolatok kezelése: TCP/IP hangolás és socketek

A Windows egy ismert problémája befolyásolhatja az OnDemand kiszolgálók kapcsolatának teljesítményét. Az OnDemand kiszolgálón végzett ismétlődő keresések és lekérdezések kiadásakor sok Windows socket megnyitására és bezárására kerül sor. Két alapértelmezett Windows beállítás hatással lehet az EIP és az OnDemand kiszolgáló közötti nagymértékű forgalomra.

- Amikor egy alkalmazás bezár egy Windows socketet, a Windows a socket portját TIME\_WAIT állapotba helyezi 240 másodpercre, amely idő alatt a port nem használható fel újra.
- A Windows az alkalmazások által megnyitható portok számát 5000-re korlátozza.

A lehetséges problémák elkerüléséhez módosítsa az időkorlát várakozási idő vagy a portszám paraméter értékét a Windows rendszerleíró adatbázis szerkesztőjében.

- Az időkorlát értékét 240 másodpercről állítsa egy alacsonyabb értékre (a lehetséges értékek 30 és 300 másodperc között vannak). A kulcs neve HKEY\_Local\_Machine\System\CurrentControlSet\services\Tcpip\Parameters\TcpTimedWaitDelay .
- A maximális portszámnak adjon meg egy magasabb értéket (az érvényes tartomány 5000 és 65534 között van). A kulcs neve HKEY\_Local\_Machine\System\CurrentControlSet\services\Tcpip\Parameters\MaxUserPort.

A TcpTimedWaitDelay és MaxUserPort paraméterekkel kapcsolatban további információkat a Windows dokumentációból szerezhet.

## Extended Search kapcsolat kezelése

Ez a szakasz írja le az Extended Search kapcsolat változásait az EIP 8.2 változatában.

A felhasználó által megadott területi beállítások a DESLOCALE kulcsban átadott területi beállítás értékek esetén támogatottak. Az értékpárt a parancssorban lehet megadni, ha az Extended Search kapcsolatot közvetlenül hívja meg. Az értéket az Extended Search tulajdonságainak **További** paraméterei között állíthatja be.

**Tipp:** Az Extended Search kiszolgáló szoftver az EIP 8.2 termékkel együtt érkezik.

---

## Egyesített példányok létrehozása

A tartalomkiszolgálók kapcsolatainak meghatározása után az egyesített keresések létrehozásának következő lépése a keresési sablonok építőköveit jelentő egyesített példányok létrehozása. Ez a szakasz magyarázza el az egyesített példányokat és az Egyesített példány létrehozása varázsló használatát.

## Egyesített példányok – Áttekintés

A legtöbb esetben az ügyfélalkalmazások felhasználói nem kiszolgálónként keresik az információkat. Ehelyett inkább egyetlen egyesített keresést kívánnak végrehajtani. A keresési sablonok lehetővé teszik az ügyfélalkalmazások felhasználóinak a keresések egyetlen lekérdezésbe kötegelését. Az ügyfélalkalmazásokban használt keresési sablonokat az EIP adminisztrátora hozza létre. A keresési sablon létrehozása előtt először létre kell hozni az egyesített példányokat, amelyek az egyesített attribútumai a tartalomkiszolgálók eredeti attribútumaira vannak leképezve.

A DB2 például az információkat táblákban tárolja, amelyek oszlopai jelentik a táblákban tárolt információk attribútumait. Az **Ügyfél statisztika** tábla tartalmazhatja például a **Név**, **Kötvényszám**, **Cím**, **Telefonszám** és **Foglalkozás** oszlopokat.

A Content Manager táblák és oszlopok helyett elemeket és attribútumokat alkalmaz. A DB2 táblában tárolt információk ebben az esetben például egy `Ügyfél_info` példányban tárolhatók. Ennek attribútumai például az `Ügyfél_név`, `Számla`, `Cím`, `Telefon` és `Munkahely` lehetnek. Bár mindkét esetben azonos információkról van szó, a tárolásuk és azonosításuk módja eltérő.

Az EIP megoldja az azonos információkat tároló sokféle tartalomkiszolgáló számításba vételével kapcsolatos problémát. Az egyesített példányok pont ezeket az információkat követik nyomon. Az egyesített példányok valójában nem tárolnak adatokat, csak az egyes tartalomkiszolgálók által alkalmazott adattárolási módszerekre vonatkozó metaadatokat. Az egyesített példányok létrehozásakor az összes attribútumot leképezheti a tartalomkiszolgálók megfelelő lekérdezni kívánt eredeti attribútumaira.

A fenti példánál maradva létrehozhat például egy `Kötvény_info` egyesített példány, amelynek egyesített attribútumai a következők: `Kötvény_név`, `Kötvény_szám`, `Cím` és `Foglalkozás`. Ezután az egyesített attribútumok leképezhetők a megfelelő eredeti attribútumokra.

Az EIP elő tud állítani egy kiszolgáló tárolót, amely ezeket az információkat tartalmazza. Ezeket az információkat a kiszolgáló tároló tartalmazza, az Egyesített példányok létrehozása varázsló pedig lehetőséget ad a kiszolgáló tároló lekérdezésére és tartalomkiszolgálók szerinti szűrésére. A szűrés csak az Egyesített példányok létrehozása varázsló használatakor áll rendelkezésre. A szűrés nem választható az egyesített példányok kézi (varázsló nélküli) létrehozása esetén. A kiszolgáló tároló létrehozása után megkezdhető az egyesített attribútumok leképezése az eredeti attribútumokra.

Az egyesített attribútumokat nem elég leképezni az eredeti attribútumokra. Minden eredeti attribútum rendelkezhet különféle tulajdonságokkal is. Az attribútumok (1) nullképesek, (2) lekérdezhetők, (3) frissíthetők és (4) szövegesen kereshetők lehetnek. A kijelölt adattípustól függően lehetősége lehet az adat hosszának, pontosságának, maximális vagy minimális értékének újraosztályozására.

A tulajdonságok meghatározásakor nem állíthatja ezeket szigorúbbra az egyesített attribútumra leképezett eredeti attribútum által már megadott tulajdonságoknál. A varázsló által biztosított alapértelmezett tulajdonságok megfelelnek ezeknek a feltételeknek. Ha az alapértelmezett egyesített attribútum alapértelmezett tulajdonságainak testreszabása után vissza kívánja állítani a varázsló által javasolt alapértelmezett tulajdonságokat, akkor kiválaszthatja az alapértelmezett beállításokat.

Összefoglalásképp: az egyesített attribútumok az egyes tartalomkiszolgálók megfelelő eredeti attribútumaira kerülnek leképezésre. Minden egyesített attribútum tulajdonságai felveszik az eredeti attribútumok tulajdonságait. Az egyesített példányok létrehozása után már tudni lehet, hogy megvannak a tartalomkiszolgálókon tárolt információkhoz vezető útvonalak. Ezután hozhatók létre az egyes lekérdezések keresési sablonjai az egyesített példányok felhasználásával.

## Az Egyesített példány létrehozása varázsló használatával

Az egyesített példány létrehozási varázsló az EIP 8.2 változatának újdonsága. Bár az egyesített példányok létrehozására az EIP 7.1 és korábbi változatában megszokott párbeszédablakokat is felhasználhatja, a varázsló egyszerűbbé teszi az egyesített példányok létrehozását.

Egyesített példány létrehozásához a varázslóban tegye a következőket:

1. **Egyesített példány meghatározása:** Adja meg az egyesített példány nevét és leírását. Emellett azt is beállíthatja, hogy a példány szövegesen kereshető-e.
2. **Egyesített attribútumok meghatározása:** Nevezze el és adja meg az egyesített attribútumokat.

3. Az **Egyesített attribútumok leképezése**: Adja meg az egyesített attribútumok és az eredeti attribútumok leképezéseit. A varázsló lehetőséget ad a kiszolgáló tároló lekérdezésére, a leképezni kívánt eredeti attribútumok kiválasztására, és a leképezések későbbi módosítására is.
4. **Tulajdonságok meghatározása**: Adja meg az egyes egyesített attribútumok tulajdonságait. Elfogadhatja az alapértelmezett beállításokat, vagy testreszabhatja a tulajdonságokat.
5. Az **Egyesített példány megerősítése**: Tekintse át az egyesített példányok megadott beállításait. A beállítások módosításához visszatérhet a korábbi panelekhez. Ha elkészült, akkor kattintson a **Befejezés** gombra.

A fenti lépések megfelelnek a varázsló használatának. A varázsló használatát az EIP online súgója írja le részletesebben.

---

## Egyesített szövegindexek létrehozása

A Szöveg keresőmotor a Content Manager 7.1 kiszolgálókba integrálható, így lehetőségessé válik a Content Manager környezetben tárolt szöveges információk automatikus indexelése, keresése és visszaadása. A felhasználók szavak vagy kifejezések alapján találhatják meg a dokumentumokat. A szöveg keresési kiszolgáló az egybyte-os és duplabyte-os karakterkészleteket is támogatja.

Content Manager kiszolgálók (7.1 változat) és a Szöveg keresőmotor használatakor létrehozható egy egyesített szövegindex. Az egyesített szövegindex ezután leképezhető a Content Manager szöveg keresési kiszolgálókon található Content Manager szövegindexre.

Az egyesített szöveg keresési index létrehozásakor engedélyezhető az összetett keresés, amely eredeti szövegindexekre és eredeti attribútumokra is irányulhat. Ha egy szövegindexen engedélyezi a kombinált kereséseket, akkor az indexet le kell képezni egy egyesített példányra. Az egyesített példány és az egyesített attribútumok által leképezett eredeti attribútumokat a szöveg keresési kiszolgálók szöveg keresési indexeire kell leképezni.

---

## Keresési sablonok létrehozása

Az egyesített példány létrehozása után készen áll a keresési sablonok létrehozására. Ne feledje el, hogy a keresési sablon az egyesített példányt használja a tartalom tárolási helyének meghatározásához. A keresési sablon létrehozásakor meg kell határozni, hogy pontosan mit kíván keresni, mit kíván tenni a keresés eredményeivel, és ki jogosult a sablon használatára. Bár minden sablonhoz csak egy egyesített példányt használhat fel, az egyes egyesített példányokat több sablonban is felhasználhatja. Keresési feltételként az egyesített példány attribútumait tetszőleges kombinációban felhasználhatja. A keresési sablonok létrehozása az alábbi lépések végrehajtásából áll a Keresési sablon varázslóban:

1. Keresési sablon meghatározása.
2. Keresési feltételek meghatározása.
3. Keresés beállításainak meghatározása.
4. Hozzáférési jogosultságok hozzárendelése.

A fenti lépések megfelelnek a Keresési sablon varázsló lépéseinek. A keresési sablon létrehozási folyamat befejezéséről az EIP online súgójában talál részletes információkat.

## Keresési sablon meghatározása

A varázsló az elindítás után megkéri egy keresési sablon megadására. Legyen felkészülve a következőkre:

- Név és leírás megadása a keresési sablonnak.

- Keresési sablon egyesített példányának kiválasztása. **Korlátozás:** Keresési sablononként csak egy egyesített példány használható.
- Egyesített szövegindex kiválasztása, ha van ilyen.

**Tipp:** Az egyesített szövegindex jelölőnégyzete csak akkor alkalmazható, ha a Content Manager 7.1 vagy korábbi változatának szöveg keresőmotorját is használja. Ha a szöveg kereséshez a DB2 TIE szolgáltatását használja, akkor ez egy paraméteres keresés, amely a keresési sablonban így is állítható be.

## Keresési feltételek meghatározása

A keresési sablon meghatározása után a varázsló a következő tevékenységekre kéri meg:

1. Keresés típusának (attribútum vagy dokumentum) kiválasztása. A dokumentum csak akkor áll rendelkezésre, ha az előző lépésben egy egyesített szövegindex kiválasztására került sor.
2. Keresési feltétel elnevezése.
3. Egyesített attribútum kiválasztása.
4. Rendelkezésre álló műveletek kiválasztása.
5. Alapértelmezett keresési karaktersorozat megadása (csak dokumentum kereséseknél).

A varázsló egy legördülő listában felajánlja a megadott egyesített példány valamennyi egyesített attribútumát. Ezek az attribútumok lesznek a keresési sablon keresési feltételei. A varázsló emellett megjeleníti a rendelkezésre álló műveletek listáját.

**Tipp:** A sablonban egynél több keresési feltételt is létrehozhat, illetve a sablonból törölhet is meglévő feltételeket.

## Keresési beállítások meghatározása

A párbeszédablak lehetővé teszi a keresés alapértelmezett beállításainak, és a megjelenő kezdeti értékek beállítását. Minden beállítás rendelkezik egy módosítható alapértelmezett értékkel. A beállítások módosításához kattintson a beállításnak megfelelő gombra.

Az alapértelmezett beállítások ablak a következőket teszi lehetővé:

- Annak beállítása, hogy mi történjen abban az esetben, ha az ügyfélalkalmazás által használni kívánt keresési sablon nem érhető el.
- Paraméteres keresések helyettesítő karakterének meghatározása.
- Keresés eredményeit tároló mappa nevének meghatározása.
- Az összes (ÉS) vagy bármelyik (VAGY) feltétel használatának kiválasztása.

A feltétel beállításai ablak lehetővé teszi a keresési feltételek sorrendjének meghatározását, a keresés eredményoszlopainak kiválasztását, és az oszlopok fejlécének, szélességének illetve sorrendjének meghatározását.

A Megjelenő érték beállítások ablak lehetővé teszi a keresési eredményekben szereplő értékek megjelenésének módosítását. Ha például a Nap attribútum értéke az egyik kiszolgálón Hétfő, egy másikon viszont H, akkor megadhatja, hogy a keresés eredményeiben mindig a Hétfő jelenik meg.

## Jogosultságok hozzárendelése

Miután beállította, hogy a keresési sablon hol (egyesített példányok) illetve mit (keresési feltétel) keressen, és hogyan jelenítse meg az eredményeket (beállítások), meg kell határoznia azt is, hogy ki férhet hozzá a keresési sablonokhoz.

A Keresési sablon varázsló Jogosultságok hozzárendelése ablakában adható meg a sablon elérése a meglévő felhasználóknak vagy felhasználói csoportoknak.

A felhasználók hozzáférése a keresési sablonokhoz nem jelenti automatikusan azt, hogy a sablonban meghatározott tartalomkiszolgálókhoz is hozzáférnek. A felhasználóknak meg kell felelniük minden egyes tartalomkiszolgáló egyéni biztonsági követelményének. Mielőtt a felhasználóknak jogosultságot biztosítana egy keresési sablonhoz, a hozzáférés felügyeleti listák és a felhasználókezelési funkciók segítségével győződjön meg róla, hogy a felhasználók rendelkeznek a megfelelő jogosultságokkal.

Ha a varázslót használja a felhasználók vagy felhasználói csoportok kikeresésére, akkor az EIP csak azokat a felhasználókat adja vissza, akik hozzáférnek a kért tartalomkiszolgálókhoz.





---

## 4. fejezet Felhasználói hozzáférés kezelése

A felhasználók nem férhetnek hozzá az EIP rendszerhez felhasználói azonosító, jelszó és jogosultságkészlet nélkül. A felhasználók létrehozása, és jogosultságaik meghatározása előtt azonban meg kell határozni, hogy ki fog hozzáférni a rendszerhez, és a feladataik milyen jogosultságokat igényelnek. Az objektumok törlésének engedélyezése például nem túl jó ötlet, mivel a felhasználók többsége nincs feltétlenül tisztában az objektumok törlésének hatásaival. Másrészt viszont az se jó megközelítés, ha a felhasználók nem tudják elvégezni a munkájukhoz szükséges feladatokat. Ennek megfelelően a felhasználói jogosultságok hozzárendelése előtt meg kell határozni az egyes munkakörökben végzett feladatok típusát.

Amikor a felhasználók létrehoznak egy objektumot az EIP rendszerben, meg kell határozniuk, hogy a többi felhasználó milyen hozzáféréssel rendelkezzen az objektumhoz. Az objektumot létrehozó felhasználóknak meg kell adniuk, hogy ki férhet hozzá az objektumhoz, és milyen műveleteket végezhet el rajta. Ezt a meghatározást nevezzük az EIP rendszerben hozzáférés felügyeleti listának, vagy ACL-nek.

---

### Felhasználói azonosítók és jelszavak létrehozása

Ha a rendszeradminisztrátorban létrehozott felhasználói azonosítót DB2 hitelesítéshez is fel kívánja használni, akkor a felhasználói azonosítónak követnie kell a DB2 elnevezési szabályait. A DB2 elnevezési szabályok a felső szintű adminisztrátorokra és a csatlakozási felhasználói azonosítókra is vonatkoznak. Nem használhatók a következő kifejezések:

- USERS
- ADMINS
- GUESTS
- PUBLIC
- LOCAL
- Az SQL referenciában megadott bármely fenntartott SQL kifejezés.

A felhasználói azonosító nem kezdődhet a következő karakterekkel:

- SQL
- SYS
- IBM

A következő karakterek használhatók:

- A – Z **Korlátozás:** Bizonyos operációs rendszerek lehetővé teszik kis- és nagybetűk figyelembe vételét a felhasználói azonosítókban és jelszavakban. A kis- és nagybetűk megkülönböztetéséről az operációs rendszer dokumentációjában talál információkat.
- 0 – 9
- #
- \$

**Korlátozás:** A felhasználói azonosítók hossza nem lehet nagyobb 30 karakternél.

---

## DB2 adminisztrátori jogosultság

A rendszeradminisztrációs ügyfélbe való bejelentkezéskor kétféle hitelesítési szint van: egy az adatbázis szintjén, és egy a termék szintjén. Az adminisztrátorok kétféle besorolással rendelkezhetnek az adminisztrációs tartományok engedélyezésekor: felső szintű adminisztrátorok és aladminisztrátorok. Általában csak a felső szintű adminisztrátorok férnek hozzá a rendszeradminisztrációs ügyfélhez.

A felső szintű adminisztrátoroknak DB2 adminisztrátori jogosultságokkal is kell rendelkezniük. A felhasználói azonosítót ezért az operációs rendszeren db2admin jogosultságokkal kell létrehozni. Az operációs rendszeren megadott jelszó kerül felhasználásra a DB2 csatlakozáshoz és a könyvtárkiszolgálóra való bejelentkezéshez. A könyvtárkiszolgálón megadott jelszó nem kerül felhasználásra. Content Manager jogosultságok: A felhasználói azonosító a könyvtárkiszolgálón teljes Content Manager felhatalmazással ("AllPrivs") rendelkezik, hogy minden adminisztrációs tevékenységet végre tudjon hajtani.

Az aladminisztrátorok számára nem szükségesek DB2 jogosultságok. Az aladminisztrátorok csak a könyvtárkiszolgáló bizonyos részeit felügyelik, ennek megfelelően kétféleképpen jelentkezhetnek be a rendszeradminisztrációs ügyfélbe:

- Ha a felhasználói azonosító egy operációs rendszeren megadott azonosító, akkor a felhasználó ezzel csatlakozik a DB2 adatbázishoz, és jelentkezik be a könyvtárkiszolgálóra.
- Ha a felhasználói azonosító nem operációs rendszer azonosító, akkor a felhasználó a cmbfedenv.ini (Enterprise Information Portal) vagy a cmbicmenv.ini (Content Manager) fájlban tárolt titkosított felhasználói azonosító/jelszó párral csatlakozik a DB2 adatbázishoz, illetve a Bejelentkezés ablakban megadott azonosítóval és jelszóval jelentkezik be a könyvtárkiszolgálóra.

A könyvtárkiszolgáló bejelentkezésről további információkat a következő szakaszban talál.

Az aladminisztrátoroknak és EIP jogosultságokkal is rendelkezniük kell. Tartomány adminisztrátori jogosultság szükséges a résztartományokkal kapcsolatos valamennyi adminisztrációs tevékenységhez.

## Csatlakozás a DB2 adatbázishoz INI fájlok használatával

Az INI fájl minden bejegyzése egy könyvtárkiszolgáló nevéből, illetve egy titkosított felhasználói azonosító/jelszó párból áll a DB2 adatbázisra csatlakozáshoz. Ezt a titkosított felhasználói azonosítót (csatlakozási felhasználói azonosító) és jelszót a termék telepítésekor kell megadni. A csatlakozási felhasználói azonosítónak és a rendszeradminisztrátori azonosítónak különbözőnek kell lennie. Az Enterprise Information Portal a cmbfedenv.ini fájlt, a Content Manager a cmbicmenv.ini fájlt használja a DB2 csatlakozáshoz. Az alapértelmezett csatlakozási felhasználói azonosító az ICMCONCT. A telepítés során a könyvtárkiszolgáló és erőforráskezelő jelszavai három helyen találhatók meg: A cmbicmenv.ini fájl tartalmazza a könyvtárkiszolgáló elérésére szolgáló felhasználói azonosítót és jelszót. A könyvtárkiszolgálót és az erőforráskezelőt tartalmazó adatbázis hozzáférését az operációs rendszer határozza meg. Az ICMRM.properties fájl tartalmazza az erőforráskezelő felhasználói azonosítót és jelszót.

Az INI fájl használatakor, vagyis ha a felhasználói azonosító nem operációs rendszer azonosító, akkor az INI fájl felhasználói azonosítójának és a csatlakozási felhasználói azonosítónak is léteznie kell a könyvtárkiszolgálón.

A csatlakozási felhasználói azonosítót a könyvtárkiszolgálón és az operációs rendszeren kell megadni. Meg kell adni neki a UserDB2Connect jogosultságot. Az INI fájlban tárolt

csatlakozási felhasználói azonosító és jelszó cseréjéhez válassza az adminisztrációs ügyfél ablakban az **Eszközök → Adatbázis azonosító/jelszó cseréje** menüpontot.

## Könyvtárkiszolgáló és erőforráskezelő rendszeradminisztrátori jelszó módosítása

Ha módosítani kívánja az erőforráskezelő jelszót, akkor módosítani kell a könyvtárkiszolgáló által az erőforráskezelőre való bejelentkezéshez használt jelszót, és az erőforráskezelő rendszeradminisztrátori jelszavát is. **Fontos:** Az erőforráskezelő rendszeradminisztrátori és könyvtárkiszolgáló jelszavának cseréjekor be kell tartani az alábbi sorrendet:

1. Jelentkezzen be a rendszeradminisztrációs ügyfélbe.
2. Bontsa ki az erőforráskezelő fát.
3. Kattintson a módosítani kívánt erőforráskezelőre, és bontsa ki a fáját.
4. Kattintson a kiszolgálómeghatározások elemre, majd válassza a Tulajdonságok elemet. Megjelenik a kiszolgáló panel ablak.
5. Módosítsa a jelszót a Jelszó mezőben.
6. Kattintson az OK gombra.
7. Kattintson a jobb egérgombbal a (2. lépésben) kibontott erőforráskezelőre, majd válassza az előugró menü Tulajdonságok menüpontját. Megjelenik az Erőforráskezelő tulajdonságai ablak.
8. Módosítsa a jelszót a Jelszó mezőben, majd kattintson az OK gombra.

## Adatbázis hozzáférési jelszavak módosítása

Ha módosítani kell az adatbázis hozzáférési jelszavakat, akkor módosítani kell az az operációs rendszer jelszót és az ICMRM.properties fájlt is, hogy az erőforráskezelő azonosíthassa az új jelszót.

Az adatbázis csatlakozás operációs rendszer jelszavának cseréjéhez tegye a következőket:

1. Az operációs rendszertől függően keresse meg a Felhasználók és jelszavak segédprogramot.
2. Kattintson az ICMRM bejegyzésre.
3. Kattintson a Jelszó beállítása elemre.
4. Adja meg az új jelszót.

Az ICMRM.properties fájl módosításához tegye a következőket:

1. Nyissa meg az ICMRM.properties fájlt. A fájl alapértelmezett helye az X:\WebSphere\AppServer\installedApps\icrmr.ear\icrmr.war\WEB-INF\classes\com\ibm\mm\icrmr\ICMRM.properties, ahol az X a Content Manager telepítési meghajtója.
2. Módosítsa a DBPassword bejegyzést, hogy az megegyezzen az operációs rendszer jelszóval.
3. Mentse az ICMRM.properties fájlt.

Az adatbázis jelszó módosítása után az adatbázis újra kell indítani, hagyja, hogy kidobjon két-három hibát, amíg nem állítja magát alaphelyzetbe.

A jelszavak módosítására vonatkozó részletes útmutatásokat, illetve a rendszeradminisztrációs ügyfél erőforráskezelőre vonatkozó mezőinek leírását a rendszeradminisztráció online súgójában találja.

---

## Felhasználók importálása LDAP címtárból

Az LDAP a rendszerenkénti felhasználókezelés helyett lehetővé teszi a felhasználói azonosítók vállalati szintű kezelését. Az EIP háromféle LDAP technológia használatára képes: IBM Directory (korábbi nevén IBM SecureWay Directory), Windows 2000 Active Directory és Lotus Domino Directory Notes Address Book (NAB). A felhasználói jelszó az LDAP kiszolgálón található. Amikor egy felhasználó bejelentkezik az Enterprise Information Portal kiszolgálóra, akkor a rendszer hitelesíti a felhasználói azonosítót és a jelszót, továbbá ellenőrzi a felhasználó jogosultságait az EIP adatbázisban tárolt felhasználói profilban. Elképzelhető, hogy az LDAP engedélyezve lett az EIP telepítése során. Ha az LDAP nem lett engedélyezve a telepítés során, akkor ezt később is engedélyezheti.

Az LDAP engedélyezéséhez kattintson a **Start → Programok → EIP for Multplatforms → LDAP felhasználói azonosító importálás ütemező** menüpontra, majd indítsa el a rendszeradminisztrációs ügyfelet. Jelenítse meg az LDAP konfiguráció ablakot (Eszközök → LDAP konfiguráció). Válassza ki az LDAP felhasználó importálás és hitelesítés jelölőnégyzetet, majd adja meg az LDAP kiszolgálóra vonatkozó információkat a Kiszolgáló lapon.

Az LDAP engedélyezése után a felhasználók importálása az Új felhasználó ablak LDAP gombjával lehetséges. Ez lehetővé teszi a felhasználók szelektív importálását az LDAP kiszolgálóról az EIP rendszerbe. Ennek alternatívájaként a felhasználókat csoportosan is importálhatja az LDAP felhasználói azonosító importálás ütemezője segédprogrammal. A bejelentkezés során a könyvtárkiszolgáló automatikusan csatlakozik az LDAP kiszolgálóhoz a felhasználó hitelesítése érdekében. Ha az LDAP kiszolgáló bármely okból nem képes a felhasználó jelszavának ellenőrzésére, akkor a hitelesítés meghiúsul.

Az LDAP kiszolgáló módosításához kattintson a rendszeradminisztrációs ügyfél főablakában az **Eszközök → LDAP konfiguráció** menüpontra. Az aktuális LDAP kiszolgálót a **Start → Programok → IBM Enterprise Information Portal for Multiplatforms 8.2 → LDAP felhasználói azonosító importálás ütemező** menüponttal elérhető LDAP felhasználói nyilvántartás importálási segédprogrammal is módosíthatja. Az LDAP tervezéséről további információkat a *Planning and Installing Your Content Management System* című kiadványban talál. Az LDAP kiszolgálónak a rendszeradminisztrációs ügyfélben végzett beállításáról az ügyfél online súgójában talál további információkat.

Az LDAP tervezéséről további információkat a *Planning and Installing Your Content Management System* című kiadványban talál. Az LDAP használatának megvalósítását a rendszeradminisztrációs ügyfél online súgója tárgyalja.

---

## Jogosultságok bemutatása

Az adminisztrációs ügyfélben jogosultsági csoportok, jogosultságkészletek és egyéni jogosultságok rendelhetők hozzá. Összevont EIP/Content Manager rendszer felügyelete esetén a jogosultságok közősek a kétféle ügyfélben. Az ügyfél beépített jogosultságai a felügyelet leegyszerűsítését szolgálják.

### Jogosultsági csoport

A jogosultsági csoportok felhasználói feladatok gyűjteményei, amelyek segítséget nyújtanak az adminisztrátoroknak az új jogosultságkészletek létrehozásakor vagy a szerepek használatakor a Jogosultságkészlet párbeszédablakban.

### Jogosultságkészletek

A jogosultságkészletek felhasználói szerepek gyűjteményei.

### Jogosultság

A jogosultság egy felhasználói tevékenységet képvisel. Például:

**1. példa - jogosultságok:** Tegyük fel, hogy a ClientScan és ClientImport jogosultságokat csak azon felhasználók csoportjának kívánja megadni, akik az ügyfelet jellemzően csak dokumentumok beolvasására és importálására használják. Ha a feladatot általában több felhasználó végzik, akkor létrehozható egy felhasználói azonosító (például user1). Ezután a ClientScan és ClientImport jogosultságok társíthatók a user1 felhasználóhoz. Ezután hozzárendelheti a user1 felhasználót a group1 csoporthoz. A Content Manager rendszerbe user1 felhasználóként bejelentkező végfelhasználók csak a dokumentumok beolvasására és importálására lesznek képesek.

**2. példa - jogosultsági csoportok:** Egy csoportnyi tapasztalt végfelhasználónak hozzá kell férnie a tipikus ügyfél feladatok mindegyikéhez. Ehhez létre kell hozni egy felhasználói azonosítót (például user2). Ezután a user2 felhasználót hozzá kell rendelni egy csoporthoz (például group2). Ezután kiosztható a ClientTaskAll jogosultsági csoport a user1 felhasználónak. A Content Manager rendszerbe user2 felhasználóként bejelentkező végfelhasználók a ClientTaskAll jogosultsági csoportba tartozó összes feladat végrehajtására képesek.

**3. példa - Jogosultságkészletek:** A felhasználók egy csoportjának csak olvasási hozzáférésre van szüksége. Ehhez létre kell hozni egy felhasználói azonosítót (például user3). Ezután a user3 felhasználót hozzá kell rendelni egy csoporthoz (például group3). Ezután kiosztható a ClientUserReadOnly jogosultságkészlet a user3 felhasználónak. A Content Manager rendszerbe user3 felhasználóként bejelentkező végfelhasználók csak a ClientUserReadOnly jogosultságkészletbe tartozó feladatok végrehajtására lesznek képesek.

---

## Jogosultságkészletek létrehozása

Az EIP rendszerkonfiguráció tervezésekor meg kell határozni, hogy ki férjen hozzá a rendszerhez, és a felhasználók hogyan használhatják a rendszer objektumait. Az EIP rendszer a hozzáférést jogosultságokon keresztül kezeli.

A jogosultság egy adott objektum egy adott módon végrehajtott elérésére vonatkozó jog. A jogosultságokba egyebek között a rendszeren tárolt objektumok létrehozása, törlése és kiválasztása tartozik bele. A felhasználókhoz hozzárendelt jogosultságok csoportját jogosultságkészletnek hívjuk.

A hozzáférések kezeléséhez kapcsolódó első feladat a felhasználók jogosultságkészleteinek létrehozása. A *jogosultságkészlet* azonosítja a felhasználó által elvégezhető feladatokat vagy tevékenységeket. A jogosultságkészletek csoportosítják a jogosultságokat, és beállíthatók az egyes felhasználótípusokhoz. Megadható például, hogy néhány adminisztrátor kezelje a dokumentumtovábbítási kiszolgálót, néhány másik pedig egy tartományt. Amikor egy adminisztrátor bejelentkezik, az EIP ellenőrzi az adminisztrátor jogosultságkészletét.

A rendszeradminisztrációs ügyfél több előre meghatározott jogosultsággal is rendelkezik, amelyek jogosultságkészletekbe csoportosíthatók. Ezután az egyes felhasználókhoz a jogosultságkészlet rendelhető hozzá. A jogosultságkészletek nem rendelhetők felhasználói csoportokhoz.

## Jogosultsági csoportok létrehozása

A jogosultsági csoportok olyanok a jogosultságoknál, mint a felhasználói csoportok a felhasználók esetén. A jogosultsági csoportok segítségével a hasonló jogosultságok csoportokba szervezhetők, így könnyedén megtalálhatók a jogosultságkészlethez hozzáadni kívánt jogosultságok. Ha például van két olyan jogosultság, amivel a rendszernek gyakorlatilag minden felhasználója rendelkezik, akkor ahelyett, hogy a jogosultságkészletek létrehozásakor minden egyes alkalommal végigkeresné a jogosultságok listáját, inkább hozzárendeli a jogosultságokat egy AlapJogok nevű jogosultsági csoporthoz.

## Jogosultságkészlet felhasználókhoz rendelése

A rendszeradminisztrációs ügyfél több előre meghatározott jogosultsággal is rendelkezik, amelyek jogosultságkészletekbe csoportosíthatók. Ezután az egyes felhasználókhoz a jogosultságkészlet rendelhető hozzá. A jogosultságkészletek nem rendelhetők felhasználói csoportokhoz.

Jogosultság nevek létrehozhatók, maguk a jogosultságok azonban nem. A rendszeradminisztrációs ügyfélben nem meghatározott jogosultságok létrehozásával kapcsolatban a rendszerprogramozóval kell konzultálnia.

Használhatja az EIP beépített jogosultságkészleteit, vagy létrehozhat saját jogosultságkészleteket is.

## Adományozható jogosultságkészlet felhasználói azonosítóhoz rendelése

Annak érdekében, hogy a felhasználók ne hozhassanak létre saját maguknál több jogosultsággal rendelkező felhasználókat, az EIP bevezette az adományozható jogosultságkészletek használatát. A felhasználók adományozható jogosultságkészletének meghatározásával megadható, hogy a felhasználók milyen jogosultságok adományozására legyenek képesek. Megadható például egy felhasználói azonosítónak néhány rendszeradminisztrációs jogosultság egy tartomány kezeléséhez. Biztosítani kívánja viszont, hogy a felhasználó ne tudjon további felhasználókat létrehozni. Ennek megfelelően a felhasználói azonosító létrehozásakor az adományozható jogosultságkészlet a NoPrivs lesz. Ennek következményeként a felhasználó kezelheti a tartományt, de nem tud felhasználói azonosítókat létrehozni a tartományban.

## Felhasználók erőforráskezelőkhöz rendelése

Ha a felhasználóknak engedélyezni kívánja egy adott erőforráskezelő elérését, akkor az erőforráskezelőt hozzá kell rendelni egy olyan tartományhoz, amelyet a felhasználók elérhetnek. Az erőforráskezelők tartományhoz rendeléséről további részleteket az "Objektumkiszolgáló tartományhoz rendelése" oldalszám: 37 szakasz tartalmaz.

## Felhasználók gyűjteményekhez rendelése

Ha a felhasználóknak engedélyezni kívánja bizonyos gyűjtemények elérését, akkor az erőforráskezelő megfelelő gyűjteményeit hozzá kell rendelni egy olyan tartományhoz, amelyet a felhasználók elérhetnek. A gyűjtemények tartományhoz rendeléséről további információkat a "Gyűjtemény tartományhoz rendelése" oldalszám: 37 szakasz tartalmaz.

---

## Felhasználói csoportok létrehozása

Az azonos munkakörökben dolgozó felhasználóknak gyakran kell azonos tevékenységeket végezniük, ennek megfelelően azonos objektumok elérésére van szükségük a rendszeren. A közös hozzáférési igényekkel rendelkező felhasználók felhasználói csoportokba szervezhetők. A felhasználói csoportok azonban nem ágyazhatók egymásba.

A felhasználói csoportok a hasonló feladatokat végző felhasználók egyszerű kényelmi csoportosításai. A felhasználói csoportokhoz nem rendelhető jogosultságkészlet. A csoport minden felhasználója saját jogosultságkészlettel rendelkezik. A felhasználói csoport csak egyszerűbbé teszi a rendszeren tárolt objektumok hozzáférés felügyeleti listáinak létrehozását.

Ha a tartományok engedélyezettek, akkor a felhasználói azonosítók csoporthoz rendelése előtt ellenőrizze, hogy a felhasználói csoport egy adott tartományhoz tartozik, vagy a NYILVÁNOS tartományban található (a tartományokról további információkat a "Tartományok felügyelete" oldalszám: 36 szakaszban talál). Győződjön meg róla, hogy a



felhasználói csoport ahhoz a tartományhoz tartozik, amelyben a felhasználót is tudni szeretné. Ha egy felhasználói azonosítót kifejezetten egy tartományon belül kíván létrehozni, akkor kattintson a Felhasználói csoport ablak **Új felhasználó** gombjára. Ekkor a létrehozott felhasználó a megfelelő felhasználói csoportba kerül, így biztosítva, hogy a felhasználó is ahhoz a tartományhoz fog tartozni.

---

## Hozzáférés felügyeleti listák létrehozása

A felhasználóknak meg kell adni a feladataik elvégzéséhez szükséges jogosultságokat. Az objektumok mellett rendelkezhetnek bizonyos hozzáférési megszorításokkal.

A hozzáférés felügyeleti listák legalább egy egyéni felhasználói azonosítóból vagy felhasználói csoportból és a hozzájuk tartozó jogosultságokból állnak. A hozzáférés felügyeleti listák segítségével felügyelhető a felhasználók hozzáférése az EIP rendszer objektumaihoz. A hozzáférés felügyeleti listához rendelhető objektumok a következők: felhasználók által tárolt adatok, elemtípusok, elemtípus részhalmazok, feladatlisták és folyamatok.

A jogosultságkészletek az egyéni felhasználók maximális képességeit határozzák meg, míg a hozzáférés felügyeleti listák a felhasználók objektumokra vonatkozó jogosultságait korlátozzák. Attól, hogy egy ACL olyan jogosultságot határoz meg, amely nem szerepel a felhasználó jogosultságkészletében, a felhasználó még nem fog rendelkezni az adott jogosultsággal. Csak az adott jogosultsággal rendelkező felhasználók használhatják jogosultságokat az objektumokon. Az ACL csak korlátozza a felhasználói hozzáférést, többletjogosultságokat nem biztosít. A hozzáférés felügyeleti listák egy újabb biztonsági szintet nyújtanak a rendszer kezeléséhez.

## Jogosultságkészlet hozzáférés felügyeleti listához rendelése

A hozzáférés felügyeleti listákhoz rendelt valamennyi felhasználónak szüksége van egy jogosultságkészletre. A felhasználói azonosító és a jogosultságkészlet határozza meg, hogy milyen felhasználók hogyan férhetnek hozzá az objektumhoz.

A felhasználók csak akkor érhetik el az objektumokat, ha szerepelnek a megfelelő hozzáférés felügyeleti listában. Felhasználó vagy felhasználói csoport hozzáférés felügyeleti listához adása érdekében jelöljön ki egy felhasználói azonosítót és egy jogosultságkészletet, majd kattintson a **Hozzáadás** gombra. A Hozzáférés felügyeleti lista ablakban az ACL minden felhasználója és csoportja megjelenik. A táblázatot a felhasználói azonosítók és csoportok hozzáadásával és eltávolításával módosíthatja. A hozzáférés felügyeleti listák létrehozásáról és módosításáról további információkat a rendszeradminisztrációs ügyfél online súgóiban talál.

---

## Tartományok létrehozása

A tartomány az adminisztrációs adatbázis olyan része, amelyet egy vagy több adminisztrátor kezel. A tartományok felhasználói azonosítókból, felhasználói csoportokból, hozzáférés felügyeleti listákból, erőforráskezelőkből és SMS gyűjteményekből állnak. A tartományok a felhasználók számára nem láthatók, ezért ezek elnevezése csak az azokat felügyelő rendszeradminisztrátorok számára bír jelentéssel. A felhasználók nem tudják, hogy az adminisztrációs adatbázis csak egy részére vannak korlátozva, vagyis csak az adott tartomány elemeiről rendelkeznek ismeretekkel.

A tartományok az adminisztrációs és felhasználói hozzáférést az adminisztrációs adatbázis bizonyos részeire korlátozzák. Az adminisztrációs adatbázisban teljes jogkörrel rendelkező adminisztrátorok korlátozott adminisztrátori jogosultságokat biztosíthatnak más adminisztrátorok számára. Egy teljes jogkörrel bíró adminisztrátor, a felsőszintű

adminisztrátor az adminisztrációs adatbázis valamennyi részéhez hozzáfér, míg a korlátozott jogkörű aladminisztrátorok csak annak bizonyos részeihez.

A tartományok korlátozzák a tartomány adminisztrátorok hozzáférését a hozzáférés felügyeleti listákhoz. Csak az adminisztrátorok hozhatnak létre hozzáférés felügyeleti listákat, amelyekben a tartomány adminisztrátorok hozzáadhatnak és eltávolíthatnak felhasználókat és csoportokat. A tartomány adminisztrátorok nem hozhatnak létre, módosíthatnak vagy törölhetnek hozzáférés felügyeleti listákat.

Az aladminisztrátorok és a felsőszintű adminisztrátorok rendelkezhetnek közös felelősségi körrel, de csak egy adott tartományra vonatkozóan. A tartományok létrehozásával és az adminisztrátorok kijelölésével a felsőszintű adminisztrátorok a teljes rendszerre koncentrálhatnak, miközben az egyes tartományok felhasználókezelési és egyéb feladatait az aladminisztrátorok végzik.

A tartományok engedélyezése előtt gondolja át a következőket:

- A tartományok nem tilthatók le.
- Az Erőforráskezelők, gyűjtemények, felhasználói azonosítók és felhasználói csoportok egyszerre csak egy tartományhoz tartozhatnak.
- A jogosultságkészletek és hozzáférés felügyeleti listák egyszerre több tartományhoz is tartozhatnak.
- A NYILVÁNOS (megosztott) tartomány kivételével a tartományok között nem lehetnek átfedések.
- A felsőszintű adminisztrációs tartományban létrehozott objektumok nem helyezhetők át, függetlenül attól, hogy a rendszer vagy a felhasználó állította ezeket elő.

A tartományok engedélyezéséhez válassza az **Eszközök → Adminisztrációs tartományok** menüpontot, majd az **Adminisztrációs tartományok engedélyezése** beállítást. A tartományok hatályba léptetéséhez újra kell indítani a rendszeradminisztrációs ügyfelet. Az adminisztrációs adatbázis tartományok kezelésére beállításáról további részleteket a rendszeradminisztrációs ügyfél online súgója tartalmaz.

## Tartományok felügyelete

A jogosultságkészlettől függően az adminisztrátorok a teljes adminisztrációs adatbázist vagy egy megadott tartományt felügyelhetnek. Az adminisztrációs adatbázison teljes hozzáféréssel rendelkező adminisztrátorok a felsőszintű adminisztrátorok. Az aladminisztrátorok egy adott tartományhoz férnek hozzá teljes jogosultsággal.

Mindkét adminisztrátornak lehetősége van objektumok, köztük felhasználók és gyűjtemények létrehozására, visszakeresésére, frissítésére és törlésére a megfelelő tartományokban. Az aladminisztrátorok az objektumok megjelenítésére és lekérdezésére csak a saját tartományukban és a NYILVÁNOS tartományban jogosultak.

## Tartományok elérése

Az aladminisztrátorok nem módosíthatják az objektumok tartományát. Ettől függetlenül elérhetik a saját tartományuk és a NYILVÁNOS tartomány minden objektumát.

A felsőszintű adminisztrátorok az adminisztrációs adatbázis valamennyi tartományához hozzáférnek. Ők hozhatnak létre objektumokat, és rendelkeznek hozzá azokat a tartományhoz. Bizonyos objektumokat, például jogosultságkészleteket és hozzáférési felügyeleti listákat csak ők hozhatnak létre.

Az aladminisztrátorok csak a saját tartományukban jogosultak objektumok létrehozására, visszakeresésére, frissítésére és törlésére.



## Felhasználó tartományhoz rendelése

A felhasználói azonosítók létrehozásakor az azonosító hozzárendelhető egy tartományhoz, vagy meghagyható az alapértelmezett tartományban. A felhasználói azonosító tartományát a későbbiekben a felhasználó tulajdonságai segítségével módosíthatja.

A felhasználói azonosítók egyszerre csak egy tartományt érhetnek el. A felhasználók nem adhatók hozzá a megosztott NYILVÁNOS tartományhoz.

Csak a felsőszintű adminisztrátorok rendelkeznek a tartományok létrehozásához és a felhasználók tartományhoz rendeléséhez szükséges jogosultságokkal. A tartományoknak lehet egynél több adminisztrátora is, egyedül a felsőszintű adminisztrátor képes azonban ezen adminisztrátorok kijelölésére a rendszeradminisztrációs jogosultságok biztosításával. Az Új felhasználó vagy a Felhasználó tulajdonságai ablak **Adományozható jogosultságkészlet** mezője jelzi a tartomány aladminisztrátorainak jogosultságait.

## Felhasználói csoport tartományhoz rendelése

A felhasználói csoport tartományhoz rendelése megváltoztatja a csoport felhasználói azonosítói számára kijelölt tartományt. A felhasználói azonosítók egyszerre csak egy tartományt érhetnek el. Ennek megfelelően a csoport felhasználói azonosítói a csoporttal együtt átkerülnek az új tartományba.

A felhasználói csoportok egyszerre csak egy tartományban lehetnek. A felhasználói csoportok a megosztott NYILVÁNOS tartományhoz is hozzárendelhetők.

## Jogosultságkészlet tartományhoz rendelése

A tartományokhoz hozzáadott valamennyi felhasználói azonosítónak rendelkeznie kell jogosultságkészlettel is. Ha nem tartalmazza a társított jogosultságkészleteket, akkor a felhasználók nem fogják tudni végrehajtani a feladataikat. A jogosultságkészletek tárolására legalkalmasabb a NYILVÁNOS tartomány, mert így minden felhasználó számára elérhetők lesznek.

## Objektumkiszolgáló tartományhoz rendelése

A felhasználói hozzáférés korlátozható bizonyos erőforráskezelőkre ezeknek egy adott tartományba helyezésével. Amikor az adminisztrációs adatbázisban új erőforráskezelőt hoz létre, akkor lehetősége van egy tartomány kiválasztására.

Alapértelmezésben minden erőforráskezelő a NYILVÁNOS tartományban található. Ha egy erőforráskezelőhöz nem kíván mindenkinek hozzáférést biztosítani, akkor rendelje hozzá egy tartományhoz. Ha nem jelenik meg a tartomány, amelyhez az erőforráskezelőt hozzá kívánja rendelni, attól még beállíthatja az erőforráskezelőt, majd később létrehozhatja a szükséges tartományt. A megfelelő tartomány létrehozása után nyissa meg az erőforráskezelő tulajdonságait, és válassza ki a tartományt.

## Gyűjtemény tartományhoz rendelése

A felhasználói hozzáférés korlátozható egy erőforráskezelő bizonyos gyűjteményeire ezeknek egy adott tartományba helyezésével. Ha az erőforráskezelő a NYILVÁNOS tartományban található, akkor a gyűjteményei tetszőleges más tartományhoz hozzárendelhetők. Ha viszont az erőforráskezelő már egy adott tartományhoz tartozik, akkor a gyűjtemény nem rendelhető hozzá ettől eltérő tartományhoz, még akkor sem, ha a gyűjteményt a NYILVÁNOS tartományba kívánja helyezni.

A felhasználónak el kell érnie az erőforráskezelőt a rajta tárolt gyűjtemények eléréséhez, ennek megfelelően nem korlátozhatja az erőforráskezelő elérését a rajta tárolt gyűjtemények korlátozása nélkül.

## Felhasználó áthelyezése egyik tartományból a másikba

Bizonyos esetekben szükség lehet a felhasználók áthelyezésére az egyik tartományból egy másikba. Fontolja meg, hogy a Felhasználó meghatározása ablak **Leírás** mezőjében megadja a felhasználó csoportjait. Ez lényegesen könnyebbé teszi a feladatot.

**Fontos:** A feladat rendkívül időigényes, és helytelen végrehajtása problémákat okozhat a rendszer elérésekor. A felhasználók tartományának módosításához felsőszintű adminisztrátornak kell lennie.

Figyelmesen tegye a következőket:

1. Keresse meg az összes csoportot, amelynek a felhasználó tagja.
2. Az összes ilyen csoportot helyezze át a NYILVÁNOS tartományba, vagy távolítsa el belőlük a felhasználót.
3. Helyezze át a felhasználóhoz társított valamennyi erőforráskezelőt a NYILVÁNOS tartományba, majd helyezze át a NYILVÁNOS tartományba a cél tartományba áthelyezni kívánt valamennyi erőforráskezelő minden gyűjteményét.
4. Hozza létre (*ne helyezze át*) a cél tartományban az összes olyan jogosultságkészletet, amely társítva van a felhasználóhoz, és még nem létezik a cél tartományban.
5. Hozza létre (*ne helyezze át*) a cél tartományban az összes olyan hozzáférés felügyeleti listát, amely társítva van a felhasználóhoz, és még nem létezik a cél tartományban.
6. Helyezze át a felhasználót a cél tartományba a felhasználó Tulajdonságok ablakának megnyitásával, és a tartomány módosításával.
7. **Nem kötelező:** A NYILVÁNOS tartományból áthelyezheti a cél tartományba a lépés 1, lépés 2 és lépés 3 helyeken áthelyezett csoportokat és erőforráskezelőket, ez azonban csak akkor kivitelezhető, ha az áthelyezett csoportokhoz és erőforráskezelőkhöz a forrástartomány egyik csoportja sincs társítva. Ellenkező esetben a csoportoknak és az erőforráskezelőknek a NYILVÁNOS tartományban kell maradniuk ahhoz, hogy több tartomány felhasználói használhassák azokat.

**Emlékeztető:** A felhasználó sohasem lehet tagja a NYILVÁNOS tartománynak. A felhasználók nem oszthatók meg.

## Felhasználói csoport áthelyezése egyik tartományból a másikba

**Fontos:** A feladat helytelen végrehajtása problémákat okozhat a rendszer elérésekor. A felhasználói csoportok tartományának módosításához felsőszintű adminisztrátornak kell lennie.

Felhasználói csoport egyik tartományból másikba helyezéséhez tegye a következőket:

- Ha a felhasználói csoport üres, akkor törölje a csoportot a jelenlegi tartományból, majd hozza létre ismét, és rendelje hozzá a cél tartományhoz.
- Ha a felhasználói csoport nem üres, akkor tegye a következőket:
  1. Keresse meg a csoporthoz tartozó valamennyi felhasználót.
  2. Törölje a csoportot a jelenlegi tartományban. Ezzel a benne található összes felhasználó is törlődik.
  3. Hozza létre ismét a csoportot, és rendelje hozzá a cél tartományhoz.
  4. Adja hozzá az összes felhasználót az újonnan létrehozott csoporthoz.

## **Erőforráskezelő áthelyezése egyik tartományból a másikba**

Az erőforráskezelők tartományának módosításához felsőszintű adminisztrátornak kell lennie.

Erőforráskezelő egyik tartományból másikba helyezéséhez tegye a következőket:

- Ha az erőforráskezelő nem tartalmaz gyűjteményeket, akkor az áthelyezéshez nyissa meg a tulajdonságait, és módosítsa a tartományát.
- Ha az erőforráskezelő tartalmaz gyűjteményeket, akkor tegye a következőket:
  1. Helyezze át az erőforráskezelőt a NYILVÁNOS tartományba.
  2. Helyezze át a gyűjteményeket a cél tartományba a Tulajdonságok ablak megnyitásával és az új tartomány kijelölésével.
  3. Az erőforráskezelő áthelyezéshez nyissa meg a tulajdonságait, és módosítsa a tartományát.

## **Gyűjtemény áthelyezése egyik tartományból a másikba**

A gyűjtemények tartományának módosításához felsőszintű adminisztrátornak kell lennie.

Gyűjtemény egyik tartományból másikba helyezéséhez tegye a következőket:

1. Keresse ki az erőforráskezelőt, amelyhez a gyűjtemény tartozik.
2. Helyezze át a társított erőforráskezelőt a NYILVÁNOS tartományba.
3. Helyezze át a gyűjteményt a cél tartományba a Tulajdonságok ablak megnyitásával és az új tartomány kijelölésével.
4. Az erőforráskezelő áthelyezéshez nyissa meg a tulajdonságait, és módosítsa a tartományát.

## **Jogosultságkészlet áthelyezése egyik tartományból a másikba**

Mivel a jogosultságkészletek több tartománynak is részei lehetnek, ezeket úgy is hozzá lehet adni a cél tartományhoz, hogy a forrástartományból nem kerülnek eltávolításra.

## **Hozzáférés felügyeleti lista áthelyezése egyik tartományból a másikba**

Mivel a hozzáférés felügyeleti listák több tartománynak is részei lehetnek, ezeket úgy is hozzá lehet adni a cél tartományhoz, hogy a forrástartományból nem kerülnek eltávolításra.



---

## 5. fejezet Információbányászat kezelése

Ez a szakasz az Információbányászat bemutatásával, és az üzleti környezetben való felhasználásával kezdődik. Ezt követik az Információbányászat első lépései, az Információbányászat alapelvei, az Információstrukturálási eszköz, majd befejezésül néhány teljesítményhangolási megjegyzés.

---

### Mi az Információbányászat?

Az Enterprise Information Portal Információbányászat vezető technológiája az információkinyerés és -elemzés legnagyobb részének automatizálásával alacsony költségbefektetés árán segíti a vállalati felhasználókat a lényeges információk egyszerű elérésében.

Az Információbányászat számára az első kihívást a szervezetlen szövegekben található információknak a számítógép számára elérhetővé tétele jelenti. A korlátlan lehetőségekkel bíró természetes nyelven megfogalmazott tényszerű ismeretek teljes értelmezése még a mai technológiától is messze áll. Ennek ellenére a mintafelismerési technikákat és heurisztikus algoritmusokat alkalmazó eszközök már ma is képesek arra, hogy tetszőleges szövegből értékes információkat nyerjenek ki. A kinyert információk fontosnak tekintett szavaktól, például nevektől, intézményektől vagy helyszínektől a dokumentumok teljes összegzéséig terjedhetnek.

Az Információbányászat azonban többre hivatott az egyes dokumentumokban található információmorzsák kinyerésénél. Hatalmas dokumentumgyűjtemények kezelésekor jön be a képbe az Információbányászat. A "bányászat" kifejezést a tudás feltérképezésének folyamatára, vagyis az önálló dokumentumokban található kódolt információk azonosítására, kinyerésére és metaadatként tárolására; illetve az ilyen jellegzetességeknek a dokumentumgyűjteményben vett eloszlásának elemzésére, más szóval az érdekes jelenségek, minták vagy irányzatok felismerésére is alkalmazzák.

### Az Enterprise Information Portal Információbányászat szolgáltatásai

Az Enterprise Information Portal Információbányászat szolgáltatás infrastruktúrát biztosít az egyéni dokumentumokra vagy dokumentumgyűjteményekre vonatkozó információk előállítására és karbantartására. Ezeket az információkat nevezzük metaadatoknak. Példák a dokumentum tartalmára jellemző, és metaadatként tárolt információkra:

- Cím
- Kivonat vagy összegzés
- Nevek, megfogalmazások vagy kifejezések
- Kategóriák, amelybe a dokumentum beletartozik

Ami az Információbányászat szolgáltatást megkülönbözteti a hagyományos, a dokumentumokhoz metaadatokat hozzárendelő metaadat tárolóktól az, hogy az Információbányászat lehetőséget ad a metaadatok automatikus létrehozására még akkor is, ha ezek nem állnak kifejezetten rendelkezésre. A bányászati és visszakeresési algoritmusok képesek arra, hogy tényszerű információkat keressenek ki hatalmas dokumentumgyűjteményekből úgy, hogy a metaadatokat a keresési folyamat vezetőiként használják, vagy a metaadatokon statisztikai modelleket lefuttatva olyan érdekes relációkat találnak a dokumentumok között, amelyek nem tűnnének fel a gyűjtemény egyes dokumentumainak áttekintésekor.

Mivel a bányászati és visszakeresési műveletek az eredeti dokumentumtartalom helyett jól meghatározott metaadat halmazon futnak, a folyamatok sebessége látványosan megnövelhető a metaadatok dedikált helyen, az úgynevezett *Információbányászat adattárolóban* tárolásával. Ez lényegesen felgyorsítja a metaadatok elérését, mivel az alkalmazások egyetlen lerakútból férnek hozzá a metaadatokhoz ahelyett, hogy több más tartalomkiszolgálóhoz kellene ezekért fordulniuk. Mivel a metaadatok kritikusak a visszakeresés és navigáció szempontjából, továbbá gyakran használják fel őket a keresések eredményeinek szűkítésére is, több lépés is végrehajtható anélkül, hogy közvetlenül a tartalomkiszolgálók lekérdezésére lenne szükség.

A metaadat tárolók használatának másik előnye, hogy az adatok a hozzájuk tartozó tényleges dokumentumtól elkülönítve tárolódnak. Emellett a csak olvasható dokumentumok, például webes anyagok esetén a tartalom és a metaadatok azonos tárolóban tárolása nyilvánvalóan nem is lehetséges.

Az Enterprise Information Portal Információbányászat szolgáltatásai a metaadatok automatikus létrehozására az alábbi módszereket biztosítják:

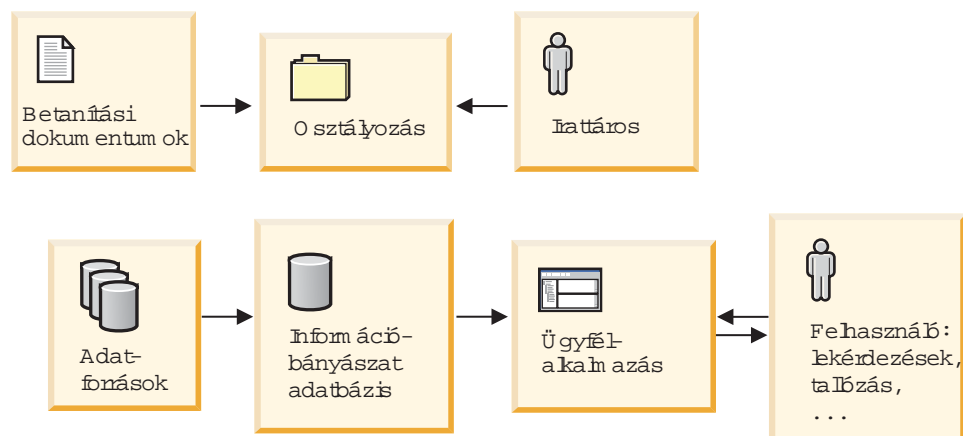
- A **Kategorizálás** a felhasználó által megadott osztályozásnak megfelelően hozzárendeli a dokumentumot legalább egy kategóriához. A kategorizálási összetevő egyik alkalmazása, az *Információstrukturálási eszköz* grafikus felhasználói felületet biztosít az osztályozások létrehozásához és karbantartásához.
- Az **Összefoglalás** kinyeri a dokumentumok legfontosabb mondatait, amely segít a felhasználónak annak meghatározásában, hogy el kell-e olvasnia a teljes dokumentumot. A felhasználó meghatározhatja a használni kívánt összefoglalások preferált méretét, így közvetlenül befolyásolhatja a kinyert metaadatok bonyolultsága és a dokumentum információmennyiségének egyensúlyát.
- A **Nyelv azonosítása** meghatározza, hogy a dokumentum milyen nyelven íródott. Ez hasznos előfeldolgozási lépés a további Információbányászat szolgáltatások alkalmazása előtt.
- Az **Információk kinyerése** automatikusan felismeri a szöveges dokumentumok lényeges szótári elemeit, például a neveket, szókapcsolatokat és kifejezéseket.
- A **Fürtözés** felosztja a dokumentumkészleteket hasonló csoportokra vagy fürtökre. A fürtök automatikusan kerülnek meghatározásra a dokumentumgyűjtemény alapján.

## Az Információbányászat szolgáltatások összetevői

Az Információbányászat alkalmazásai általában a következő feladatokat végzik el:

1. Az adatok tallózható és kereshető formába szervezése
2. Különböző adatforrások elérése
3. Előrejelzésekhez vagy irányvonalak megalkotásához szükséges adatok kiszűrése összetett keresési műveletekkel

Ezeket az Információbányászat feladatokat az 1. ábra: oldalszám: 43 mutatja be.



1. ábra: Az Információbányászat feladatai

Az Információbányászat funkcióinak kihasználásához a dokumentumokat tallózható és navigálható formába kell szervezni. Ezt a feladatot általában egy tudasszakértő végzi el. Az irattáros az Információstrukturálási eszköz (IST) segítségével meghatároz egy osztályozást: a dokumentumok adatainak téma szerinti hierarchikus besorolását. Az IST egy grafikus felhasználói felülettel rendelkező alkalmazás az osztályozások elkészítéséhez és karbantartásához. A kategóriák betanításra kerülnek, majd az osztályozás stabilnak minősítése után más adatforrások is elérhetők.

A dokumentumok tetszőleges EIP tartalomkiszolgálóról, illetve a web bejáró (csak JavaBeans megvalósításként érhető el) szolgáltatás segítségével az Internetről is importálhatók, majd kategóriákba sorolhatók.

A dokumentumokra alkalmazott szövegelemzési és metaadat létrehozási függvények csak felhasználói felület nélküli Java beanek és Java szolgáltatási API-k programozási illesztőjének szintjén érhetők el.

Az Információbányászat JavaBeans komponensek gyors alkalmazásfejlesztést biztosító szoftverösszetevők, és megfelelnek a JavaBeans szabványnak. A Java szolgáltatási API a teljes Információbányászat funkcionalitást egyedi építőkockákként biztosítja, amelyekből alkalmazások építhetők fel. Az ilyen alkalmazások létrehozását JavaBeans alapú példák és példa JSP oldalak segítik.

Dokumentum tartalmának azonosítása előfeltétel minden olyan Információbányászat művelethez, amely feldolgozza a dokumentumot. A feladat részlépései egyebek között a következők lehetnek:

1. Dokumentum kódlapjának azonosítása.
2. A feldolgozásra kerülő szövegrészek azonosítása, más szóval a jelölő információk vagy bináris adatok, például képek elkülönítése és figyelmen kívül hagyása.

Mivel a tartalomkiszolgálókon található dokumentumok szerkezete tetszőleges lehet, az Információbányászat szolgáltatások lehetőséget adnak egyedi modulok írására, amelyek a különféle dokumentumformátumok szöveges részeiből azonosítják és kiemelik a tényleges szövegrészeket. Az Információbányászat szolgáltatás részeként szállított alapértelmezett modul a gyakran használt dokumentumformátumok széles választékát felismeri. A támogatott formátumok listáját a 9. fejezet, "Dokumentumformátumok", oldalszám: 117 szakasz sorolja fel, a példa alapértelmezett modul használatára vonatkozó részletek az *Application Programming Guide* című kiadványban található.

A kijelölt dokumentumok metaadatainak létrehozása a dokumentum tartalmának feldolgozásával, és tudásbázisokon, például szótárakon vagy gyakorisági profilokon alapuló statisztikai módszerek vagy heurisztikus algoritmusok alkalmazásával történik.

Az Információbányászat alkalmazásprogram illesztők a következő műveleteket támogatják:

- Összegzés
- Kategorizálás
- Nyelvazonosítás
- Információkinyerés
- Fürtözés

A dokumentumok létrejött metaadatai az Információbányászat adattárolójában kerülnek tárolásra.

Az adattároló feltöltése után még egy lehetőség nyílik a dokumentumok kiválasztásának végrehajtására, mégpedig a dokumentumok kijelölése az adattárolóban található információk alapján. Az összetett keresés művelet a szöveges lekérdezéseket egy kategóriával kombinálja, így a keresés csak egy adott kategóriához tartozó dokumentumokra szorítkozik.

## **Az Információbányászat használata üzleti környezetben**

Az Információbányászat technológia megvalósítására képes szervezetek infrastruktúrája legalább a következő szerepekből áll:

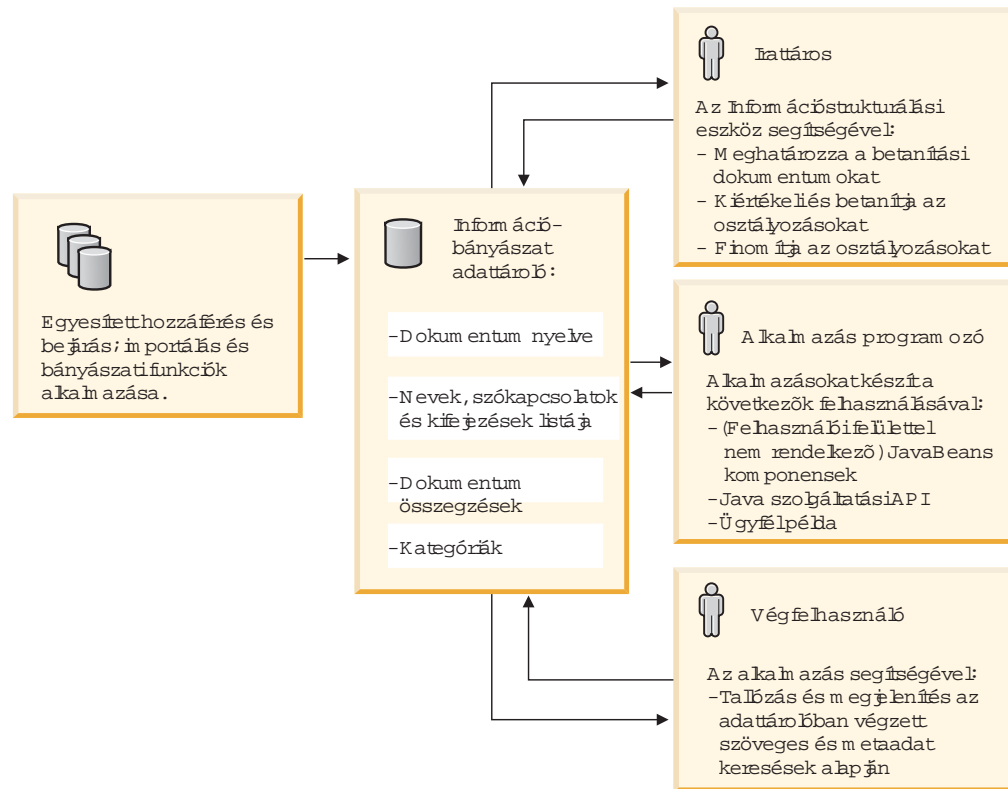
- Általános, nem feltétlenül az Információbányászat szolgáltatásra korlátozott IT rendszeradminisztrátor.
- Alkalmazásprogramozó.
- Irattáros vagy tudásszakértő.
- Az Információbányászatot használó emberek (végfelhasználók).

Az alkalmazás természetétől függően az előbb felsoroltakon kívül további egyedi szerepeket is találhat:

- Web tervező.
- Tervezőmérnök vagy konzultáns.

Ezeket az Információbányászat szerepeket és tevékenységeket a 2. ábra: oldalszám: 45 mutatja be.





2. ábra: Az Információbányászat szerepei és tevékenységei

A rendszeradminisztrátor állítja be a hardver- és szoftverkörnyezetet, és tartja karban a szükséges erőforrásokat, például a fájlrendszer területet és hozzáférési jogokat. A rendszeradminisztrátor telepíti a szükséges Enterprise Information Portal összetevőket és állítja be az aktuális forrásokat és az Enterprise Information Portal adminisztrációs alkalmazását, a különféle lerakatok elérésének megteremtéséhez. Ezen kívül a rendszeradminisztrátor végzi az Információbányászat adattároló adatbázisszintű felügyeletét.

Az alkalmazásprogramozó a JavaBeans komponensek vagy a szolgáltatási API felhasználásával elkészít egy alkalmazást. Néhány JavaBeans példát az *Application Programming Guide for Windows* című kiadványban, a szolgáltatási API leírását pedig az *Application Programming Guide* című kiadványban találja. Az alkalmazásprogramozó a feladatához segítségül hívhat egy web tervezőt is.

Az irattáros vagy tudásszakértő felelős a bányászathoz és visszakereséshez használt dokumentumgyűjtemények és erőforrások beállításáért és karbantartásáért. Az irattáros az Enterprise Information Portal adminisztrációs alkalmazása segítségével létrehozza a metaadat leképezéseket és keresési sablonokat, valamint az Információstrukturálási eszköz (lásd: "Osztályozás összeállítása" oldalszám: 61) segítségével meghatározza a katalógusokat és osztályozásokat. Az Információbányászat metaadat tárolójának feltöltése a különféle adatforrásokból vagy a webről származó dokumentumokkal általában szintén az irattáros feladata, aki ehhez az alkalmazásprogramozó által írt alkalmazásokat használja.

A végfelhasználók az alkalmazásprogramozó által készített alkalmazás felhasználásával, az irattáros vagy tudásszakértő által létrehozott és karbantartott erőforrások alapján elvégzik az információbányászati és -visszakeresési feladatokat. A végfelhasználók és az irattáros közötti munkamegosztástól függően a végfelhasználók is érintettek lehetnek a tartalomkiszolgálók dokumentumainak kiválasztásában és az Információbányászat metaadat tároló feltöltésében.

## Egy példa az Információbányászat felhasználására

Az Elektro Rt. elektronikai eszközöket gyárt a tömegpiac számára. Portfóliójában legalább öt különböző termék található az egyéni konfigurációk széles körű lehetőségeivel.

A marketing osztály rendelkezik az egyes eszközök alkalmazási területeinek vásárlói preferenciáival. Ezek a felhasználási profilok magyarázzák el, hogyan használják a vásárlók a termékeket és a különféle összeállítási lehetőségeket. Minden profil egy adott marketing, kibocsátási és relációkezelési stratégiához tartozik.

A szerviz osztály rendelkezik információkkal az eszközök alkatrészeiről, ezek összeszereléséről, az alkatrészek beszállítóiról és az egyes alkatrészek karbantarthatóságáról és megbízhatóságáról.

A partneri adminisztrációs osztály tárolja a viszonteladók és üzleti partnerek adatait. Emellett ők kezelik a különféle szerződéstípusok feltételeivel és kikötéseivel kapcsolatos jogi dokumentumokat is.

A vállalat felismerte, hogy bizonyos vásárlói körben az eladások rendkívüli mértékben lecsökkentek. A különféle elektronikai eszközök alkalmazási módszereinek versenye eltolódott, és az ügyfelek elvárásai megváltoztak az új technológiai fejlesztések hatására.

A változásokra való reagálás, és a saját célokra való kiaknázás céljával az Elektro Rt. felállít egy munkacsoportot egy olyan stratégia kidolgozására, amely segít visszahódítani az elvesztett piacot.

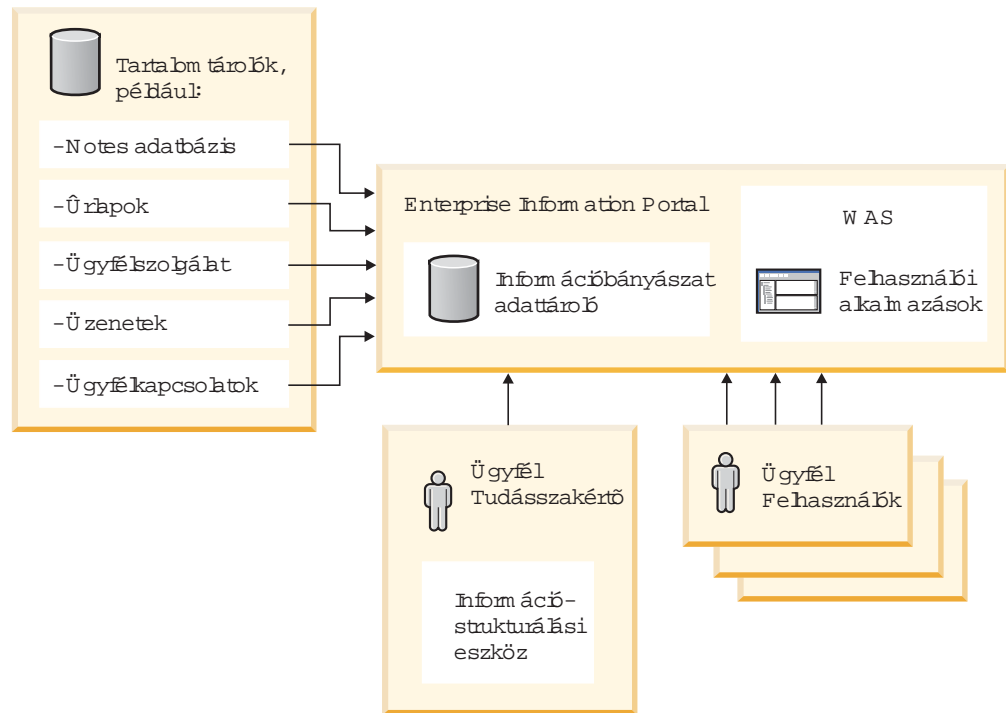
Az első lépés egy olyan információtechnológiai infrastruktúra megalkotása, amelynek segítségével az információk könnyen elérhetők, így a tervezők megfelelően informált döntéseket hozhatnak.

Ilyen információk például a következők:

- Versenytársak termékeivel, ezek képességeivel és áraival, illetve a vásárlói elfogadottságukkal kapcsolatos adatok.
- Az Elektro Rt. által kínált termékek erősségeinek és gyenge pontjainak felmérése a vásárló szemszögéből. Az ilyen termékek általános felhasználási területével kapcsolatos távlatok és irányvonalak.

Az információk különböző adatforrásokban, más-más hardver- és szoftverkörnyezetekben, eltérő szervezeti szinteken és különféle dokumentumformátumokban találhatók.

Az Információbányászat használatát a 3. ábra: oldalszám: 47 mutatja be.



3. ábra: Információbányászat példa

A munkacsoport elhatározza egy portál elkészítését, ahonnan minden ilyen forrás egy egyszerűsített ügyfélből (böngésző) érhető el. A megvalósításhoz az Enterprise Information Portal terméket választják, mivel valamennyi építőelemet biztosítja egy ilyen portál létrehozásához, sőt biztosítja az egyéni kapcsolatok elkészítésének lehetőségét a nem szabványos adatforrásokhoz is.

A következő lépések szükségesek:

1. Hardveres és szoftveres infrastruktúra összeállítása.
2. Adatforrások hozzáférési módszereinek meghatározása, a szükséges kapcsolatok beállítása, illetve a tényleges adatok leképezéseinek létrehozása.
3. Az adatok átszervezése úgy, hogy a tervezők szabadon tállózhassanak bennük.
4. Végfelhasználói alkalmazás létrehozása.

Az Információbányászat a harmadik lépésben jut szerephez. Az infrastruktúra felállítása, az adatforrások azonosítása, a megfelelő kapcsolatok kialakítása és a fontos leképezések létrehozása után a kívánt adatok egyetlen pontból érhetők el, és az adatok részhalmazai egy egyesített kereséssel határozhatók meg. A következő kérdés a szükséges adatok kiszűrése egy adott előrejelzés vagy irányvonal szempontjából, illetve az adatok olyan átszervezése, amely megfelelő a stratégiai tervezési folyamat szempontjából.

A munkacsoport kijelöl egy *tudásszakértőt*, aki stratégiai tervezési információk karbantartásáért, szervezéséért és frissítéséért felelős. A releváns információknak a hatalmas dokumentumkészletből való kinyerése érdekében a tudásszakértő megbeszélést folytat a korábbi stratégiai tervezőkkel az egyes folyamatok együttműködéséről és a jól bevált gyakorlatokról, emellett végigkeresi a vásárlói viszonyokat és terméktámogatási adatbázisokat.

Az Enterprise Information Portal keresési lehetőségeinek kihasználásával a dokumentumok könnyedén elérhetők a vásárló neve, címe, vagy az eszköz tulajdonságai alapján. A felhasználási profilok megalkotásához szükséges információk azonban a szöveg mögött

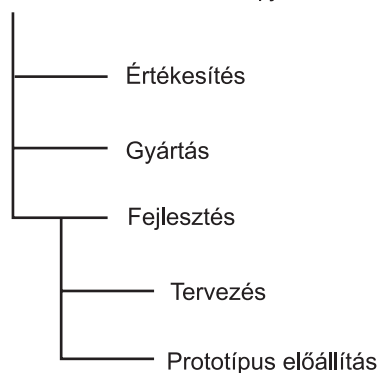
rejtőznek, és az információk megszerzésének egyedüli módja a dokumentum tartalmának intelligens elemzése az Enterprise Information Portal Információbányászat szolgáltatásainak segítségével.

A szolgáltatások által biztosított hasznos információk egyik típusa például a dokumentum tartalmának tematikai kategóriába sorolása, például *ez a dokumentum kéziszámitógépekről szól*. Az Információbányászat kategorizálási szolgáltatása a dokumentumokat a tartalmuk alapján sorolja kategóriákba. A kategóriákat egy osztályozásnak nevezett témakör szerinti hierarchia alkotja. A kifejezetten hozzáférhető és automatikusan előállított metaadatok is az Információbányászat szolgáltatás által fenntartott, katalógusoknak nevezett adattárolókban találhatók, amely meggyorsítja a hozzáférést és visszakeresést.

Az Információstrukturálási eszköz használatával a tudásszakértő meghatároz egy olyan katalógust, amely jól jellemzi, hogyan használják a vásárlók az eszközöket.

Egy katalógust a 4. ábra: oldalszám: 48 mutat be.

Adatok tervezése felhasználás alapján



4. ábra: Katalógus példa

A vásárlói adatbázison, illetve a használat módjának meghatározásához az eladási és terméktámogatási adatbázisokon végrehajtott egyesített keresés segítségével a tudásszakértő meghatározza az osztályozás kategóriáinak betanítási dokumentumait. Ez a lépés az osztályozás átszervezéséhez vezethet, mivel a releváns adatok tüzetesebb áttekintése során újabb kategóriák merülhetnek fel.

Miután az osztályozás kialakult, és minden kategóriában elegendő betanítási dokumentum gyűlt össze, a tudásszakértő az Információstrukturálási eszköz segítségével elvégzi az osztályozás betanítását. Az ennek eredményeképp létrejövő kategorizálási séma felhasználható a további dokumentumok hozzárendelésére a katalógusban található kategóriákhoz.

Eközben az információtechnológiai részleg programozói az Enterprise Information Portal egyszerűsített ügyfél, és a JavaBeans komponensek vagy a Java szolgáltatási API felhasználásával elkészítenek egy végfelhasználói alkalmazást a stratégiai tervezők számára.

Az alkalmazás több keresési sablonból áll, amelyek a stratégiai tervezésnek megfelelően kerültek beállításra. A sablonok segítségével a tervezők feltölthetik a katalógust a különféle háttérkiszolgálókról származó dokumentumokkal. A katalógus feltöltésekor az Információbányászat szolgáltatás automatikusan hozzárendel egy kategóriát a dokumentumokhoz. Az új használati profilok azonosításakor a tudásszakértő a tervezők által azonosított új dokumentumok segítségével megfelelő módon újraszervezi az osztályozást. A katalógus ismételt betanítása után az új eredmények eljutnak a tervezőkhöz.

A fenti példa illusztrálja, hogy az Információbányászat lehetőségeinek kombinálásával a vásárlói igényekre rámutatva a vállalatok a piac változásai közepette is versenyképesek tudnak maradni.

## Támogatott nyelvek és formátumok

A Enterprise Information Portal Információbányászat szolgáltatásai a következő nyelveket támogatják (lásd: 5. táblázat:):

5. táblázat: Támogatott nyelvek

Nyelv	Nyelv azonosítás	Információ kinyerés	Összegzés	Kategorizálás	Fürtözés
Angol	x	x	x	x	x
Német	x		x	x	x
Francia	x		x	x	x
Dán	x				
Finn	x				
Olasz	x		x	x	x
Norvég	x				
Portugál	x		x	x	x
Spanyol	x		x	x	x
Svéd	x				
Koreai	x		x	x	x
Japán	x	x	x	x	x
Kínai (hagyományos és egyszerűsített)	x		x	x	x

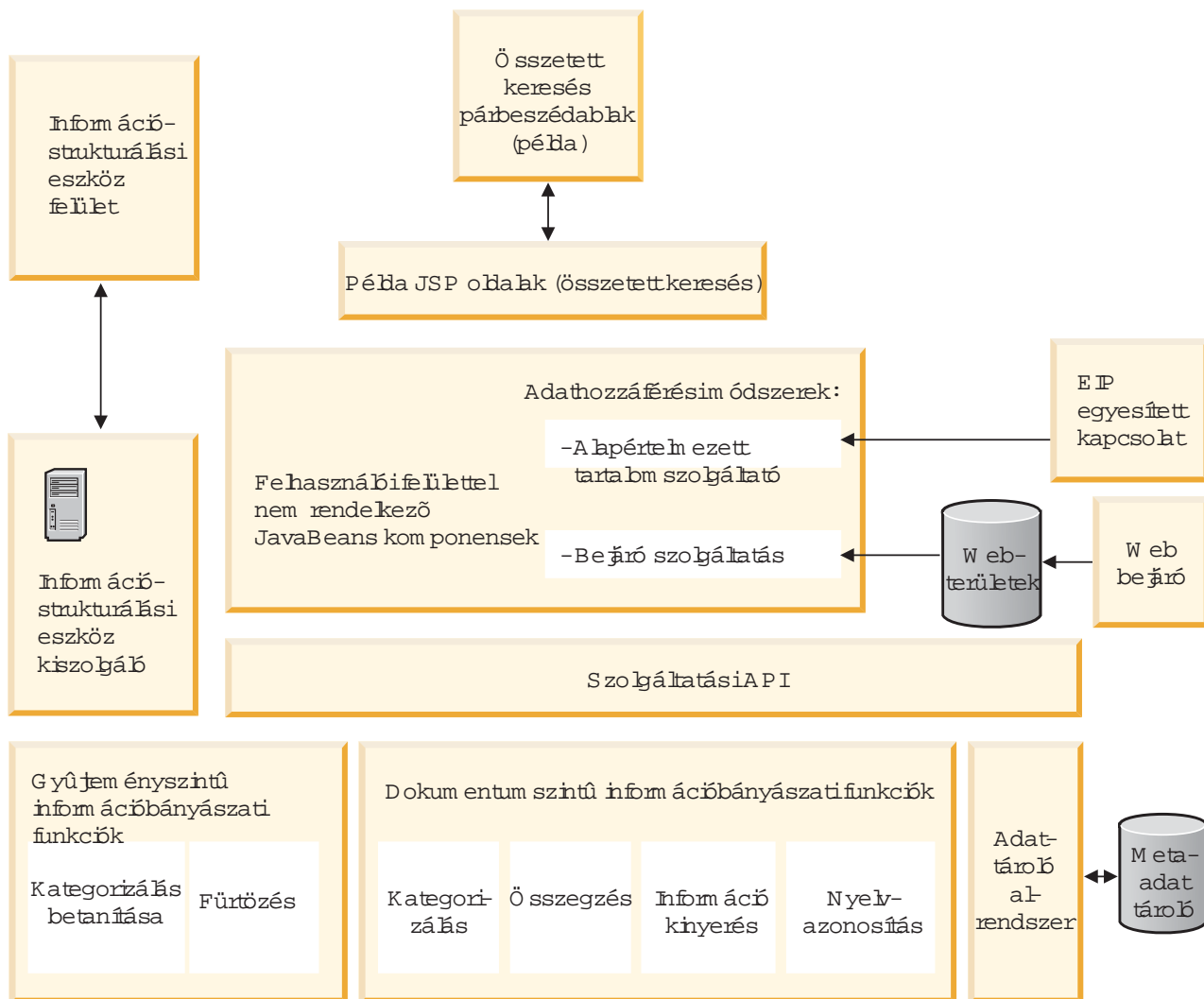
A támogatott dokumentumformátumokat a 9. fejezet, "Dokumentumformátumok", oldalszám: 117 szakasz sorolja fel.

## Alapelvek

A felhasznált információk mennyisége folyamatosan növekszik. A legtöbb szervezet egyre több és gyorsabban növekvő számú online dokumentummal rendelkezik, amelyek hatalmas potenciális értéket képviselnek. Az egyre kiélezettebb versenyben létfontosságúak például a vásárlói visszajelzések, a stratégiai információk, vagy az új és változó üzleti lehetőségek elemzései. Az Információbányászat szolgáltatások nagy mennyiségű online dokumentum kezelésére használható alkalmazásnak készült.

## A rendszer felépítése

Az Információbányászat rendszer felépítését az 5. ábra: oldalszám: 50 mutatja be.



5. ábra: Információbányászat rendszer architektúrája

Az ábra jobb oldalán látható keretek olyan összetevőket jelentenek, amelyeket az Információbányászat szolgáltatások felhasználnak ugyan, mégsem tekintendők a szolgáltatások részeként, ilyenek például:

- Egyesített kapcsolat (az OO API része)
- Web bejáró

Az Információbányászat funkcionalitása két különböző rétegből áll:

1. Java szolgáltatási API

Ez a réteg az Információbányászat funkcióit és metaadatait összefüggő JAVA alkalmazásprogram illesztőként biztosítja.

2. Felhasználói felülettel nem rendelkező JavaBeans komponensek

Ez a réteg használatra kész JavaBeans alapú összetevők felhasználásával a szabványos beanek eseménytípusait és megegyezéseit alkalmazza.

3. Példa Java Server Pages (JSP) oldalak.

Ez a réteg a felhasználói felülettel nem rendelkező JavaBeans komponenseket felhasználó mintakódból áll, és egy összetett keresési, vagyis kategória megszorítással rendelkező szöveg keresési alkalmazást szemléltet.

4. Információstrukturálási eszköz

Egy grafikus felhasználói felülettel rendelkező alkalmazás az osztályozások elkészítéséhez és karbantartásához.

## Az Információbányászat alapelvei

Az Információbányászat funkcióinak megértése és kihasználása érdekében ez a szakasz bemutatja a szolgáltatás alapelemeit.

Az Információbányászat szolgáltatás infrastruktúrát biztosít az egyéni dokumentumokra vagy dokumentumgyűjteményekre vonatkozó információk előállítására és karbantartására. Ezeket a dokumentumra vonatkozó információkat nevezzük a dokumentum **metaadatainak**.

A **könyvtár** az Információbányászat adatbázis tartalmának fogalmi meghatározása. A könyvtár katalógusok készleteiből áll.

A **katalógusok** szöveges dokumentumok metaadat tárolói, amelyek a következőket tartalmazzák:

- Egy **katalógus séma**, amely meghatározza, hogy az egyes dokumentumokhoz milyen jellemzők kerülnek tárolásra.
- Egy **osztályozás**, amely a **kategóriák** egy hierarchikus faszerkezete.
- A dokumentum betanítási eredményein alapuló **kategorizálási modell**, amely alapján a dokumentumok automatikusan kategóriákba sorolhatók. Ezt a modellt az **Információstrukturálási eszköz** állítja elő, itt lehet létrehozni és betanítani az osztályozásokat. Ez a modell jelenti a kategorizálási szolgáltatás bemenetét.

A séma határozza meg a katalógusban található dokumentumokkal kapcsolatban előállítható vagy tárolt attribútumok nevét és típusát. A séma előre meghatározott, és a következő attribútumokat tartalmazza:

- IKF\_CONTENT - karaktersorozat típus
- IKF\_TITLE - karaktersorozat típus
- IKF\_AUTHOR - karaktersorozat típus
- IKF\_CATEGORIES - karaktersorozat típus
- IKF\_SUMMARY - karaktersorozat típus
- IKF\_LANGUAGE - karaktersorozat típus
- IKF\_FEATURES - karaktersorozat típus
- IKF\_COMMENTS - karaktersorozat típus
- IKF\_DATE - időbélyeg típus
- IKF\_IDNUMBER - egész típus

A katalógus a katalógus sémája alapján létrehoz egy **rekordot** az importált dokumentumokból kinyert vagy létrehozott információk tárolásához. A rekordok egyedi azonosítóval rendelkeznek, és egy sor név/érték párt tartalmaznak. Az Állandó objektumazonosítóként (PID) is ismert egyedi azonosító jelenti a hivatkozást a létrehozott rekordok és az eredeti dokumentumforrás között.

Egy példa rekordot a 6. ábra: oldalszám: 52 mutat be.

Rekord	
IKF_TITLE	"Madarak"
IKF_AUTHOR	"Kovács J."
IKF_SUMMARY	Madarak című könyv összefoglalása"
IKF_CATEGORIES	Madarak/Ízelt-lábúakat evők
IKF_DATE	2001/07/01

6. ábra: Példa rekord

Ha egy Információbányászat rekord értékeit a JavaBeans vagy a szolgáltatási API segítségével állítja be, akkor gondoskodjék róla, hogy az értékek a meghatározott méretkorlátokon belülre essenek. Ellenkező esetben `DKIKFSizeOutOfBoundsException` érkezik. A meghatározott korlátok a következők:

6. táblázat:

Kulcs	Maximális méret (byte)
IKF_CONTENT	209715
IKF_TITLE	2048
IKF_AUTHOR	2048
IKF_CATEGORIES	8192
IKF_SUMMARY	8192
IKF_LANGUAGE	8
IKF_FEATURES	524288
IKF_COMMENTS	8192

A rekord létrehozása után hozzárendelésre kerül egy kategóriához, majd az ennek megfelelő katalógusban kerül tárolásra. A megfelelő kategória kiválasztására általában a kategorizálás eredményei használhatók, de a választás történhet a rekord egy másik értékétől függően is. A rekordokat hozzá kell rendelni egy kategóriához, mivel ez foglalja magában a dokumentum tartalmának indexelését is, amely a szöveges keresésekhez szükséges. Minden katalógus egyetlen szövegindexszel rendelkezik, ami azt jelenti, hogy a keresés eredményei mindig automatikusan a katalógus keresési hatókörén belül maradnak.

Az **adattároló alrendszer** nevű összetevő tartja fenn az állandó adattároló elérését.

## Információbányászati eszközök

Az Információbányászat szolgáltatások online dokumentumok kezeléséhez biztosítanak eszközöket. Ezek közé az alábbiak tartoznak:

- Az Információstrukturalási eszköz hozza létre és tartja karban a katalógusokat.
- A Nyelvazonosítási szolgáltatás automatikusan felismeri, hogy a dokumentum milyen nyelven íródott.
- A Kategorizálási szolgáltatás automatikusan hozzárendeli a dokumentumokat az Információstrukturalási eszközzel korábban már meghatározott kategóriákhoz.
- Az Összegzési szolgáltatás a dokumentum szavainak és mondatainak elemzésével elkészíti a dokumentumok összegzését.



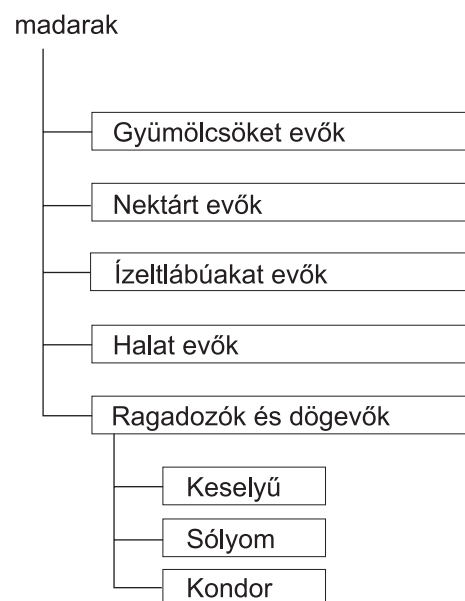
- Az Információkinyerési szolgáltatás automatikusan felismeri a szöveg fontosabb elemeit, anélkül, hogy szükség lenne tartományfüggő szótár létrehozására.
- A Fürtözési szolgáltatás a dokumentumok készleteit csoportokra vagy fürtökre bontja. Az azonos fürtben található dokumentumok közös jellemzőkkel rendelkeznek. A fürtök nem előre meghatározottak, ezek létrehozása automatikusan történik.
- Az Összetett keresés keresi a katalógusokban tárolt dokumentumok szövegeit az adott kategóriákra vonatkozóan.

## Az Információstrukturálási eszköz

Az Információstrukturálási eszköz egy web alapú alkalmazás, amely lehetőséget ad egy könyvtárnak nevezett katalóguskészlet létrehozásához és karbantartásához. A katalógusok tárolják a metaadatokat, és mindegyik katalógus egy osztályozáshoz tartozik, amely meghatározza a katalógusban található információk szervezését. Az osztályozás a kategóriáknak olyan hierarchikus felépítése, amely a dokumentumokat a téma szerinti tartalmuknak megfelelően osztályozza.

Az Információstrukturálási eszköz segítségével például az irattáros meghatározhat egy olyan katalógust, amely a madarak táplálkozási szokásait jelképezi.

Egy példa katalógust a 7. ábra: mutat be.



7. ábra: Példa katalógus

Egy jól megválasztott szerkezetű kategóriarendszer jelentősen megkönnyíti a releváns információk kikeresését az adattömegből. A kategóriák megválasztása úgy történik, hogy megfeleljenek egy adott dokumentumgyűjtemény tervezett felhasználásának, és a használat előtt ezeket be kell tanítani. Az Információstrukturálási eszköz segítségével létrehozott kategorizálási modellt a kategorizálási eszköz ezután felhasználhatja a dokumentumok kategóriákhoz rendelésének automatizálására.

Az Információstrukturálási eszköz funkciói közé az alábbi feladatok tartoznak:

- Katalógusok létrehozása, átnevezése és törlése.
- Katalógusok leírásának szerkesztése.
- Kategóriák létrehozása, átnevezése és törlése.
- Kategóriák betanítási dokumentumainak hozzáadása és eltávolítása.

- Betanítási dokumentumok tartalmának megjelenítése.
- Katalógus betanítási folyamat elindítása és leállítása.
- Visszajelzés szerzése a katalógusban található adatok betanításának minőségéről.

Az Információstrukturálási eszköz telepítéséről és használatáról további információkat az “Osztályozás összeállítása” oldalszám: 61 szakaszban talál.

## Nyelvazonosítás

A Nyelvazonosítási szolgáltatás egy adott nyelvkészletből kiválasztja azt, amelyen egy adott szöveges dokumentum legvalószínűbben íródott.

A Nyelvazonosítási szolgáltatás egy rangsorolt nyelvlistát, illetve az egyes dokumentumok megbízhatósági szintjét adja vissza. A nyelvet az ISO 639 számú szabványában megadott kétbetűs kód szerint adja meg. A megbízhatósági szint egy 0 (rossz) és 1 (jó) közötti tizedes szám, ez a mérték jelzi, hogy a dokumentum mennyire jól illik a nyelvhez. A nyelvazonosítási algoritmus egynyelvű dokumentumok nyelvének meghatározására készült. Ennek megfelelően a többnyelvű dokumentumok megbízhatósági érték szerinti rangsorolása nem garantálja a dokumentum helyes nyelvének meghatározását.

A felismerhető nyelvek a következők:

- Angol (EN)
- Német (DE)
- Francia (FR)
- Dán (DA)
- Finn (FI)
- Olasz (IT)
- Norvég/Bokmal (NB)
- Norvég/Nynorsk (NO)
- Portugál (PT)
- Spanyol (ES)
- Svéd (SV)
- Koreai (KO)
- Japán (JA)
- Egyszerűsített és hagyományos kínai (ZH)

Ezzel kapcsolatban egy tulajdonság beállítására van lehetőség:

- **maxResults** (csak a Java szolgáltatási API használatakor):

Az egyes dokumentumoknál megadható nyelvek maximális száma. Az érték 0, vagy egy annál nagyobb egész szám. Az alapértelmezett érték 1, ami azt jelenti, hogy a legjobb helyre rangsorolt eredmény kerül visszaadásra. Ha az érték 0, akkor valamennyi felismert nyelv visszaadásra kerül a rangsor szerint, kivéve azokat, amelyeknek megbízhatósági szintje 0,01 alatt van.

A nyelvazonosítás a többi Információbányászati szolgáltatás előfeldolgozási lépéseként használható. A nyelvazonosítás segítségével az információk kinyerése előtt megkereshető például az összes angol vagy japán dokumentum.

## Kategorizálás

A kategorizálás a dokumentumok kategóriákhoz rendelésének egy módja, amely a dokumentumokat az Információstrukturálási eszközzel létrehozott szervezeti modellbe szervezi.

Ennek megfelelően a kategorizálási szolgáltatás használata előtt létre kell hozni egy ilyen modellt egy osztályozás meghatározásával és betanításával az Információstrukturálási eszköz segítségével.

A kategorizálás eredményei a kategóriát és egy megbízhatósági értéket tartalmaznak, ez utóbbi jelzi, hogy a dokumentum mennyire illik bele a kategóriába. Ezek az eredmények minden dokumentumnál meghatározásra kerülnek. A lista a visszaadott rangsorolás szerint kerül rendezésre.

Ezzel kapcsolatban az alábbi tulajdonságok beállítására van lehetőség:

- **maxResults:**

Az egyes dokumentumok vonatkozásában visszaadott kategóriák maximális száma. Az alapértelmezett érték  $-1$ , ami azt jelenti, hogy minden kategória visszaadásra kerül. Az eredmények listája rangsorolva van.

- **minConfidence:**

A dokumentumokhoz rendelt megbízhatósági érték jelzi, mennyire illeszkedik bele a dokumentum az adott kategóriába. A minConfidence paraméter határozza meg a minimális értéket 0 (legkevésbé illeszkedő) és 1 (legjobban illeszkedő) között. Az alapértelmezett érték 0, ami azt jelenti, hogy a dokumentumhoz hozzárendelhető valamennyi kategória visszaadásra kerül. Az eredmények listája a jobban illeszkedőktől a kevésbé illeszkedők felé van rangsorolva.

- **catalogName:**

Ez határozza meg a kategorizálás során használt katalógust. A katalógusokat az Információstruktúrálási eszköz segítségével lehet létrehozni és betanítani a kategorizáláshoz.

A támogatott nyelveket a “Támogatott nyelvek és formátumok” oldalszám: 49 sorolja fel.

## Összegzés

A dokumentum összegzése a dokumentumokból kinyert azon mondatokat jelenti, amelyek jellemzik a dokumentum tartalmát. Az Összegzési eszköz egyebek között segíthet annak meghatározásában, hogy egy dokumentum releváns-e, tehát el kell olvasni teljesen, vagy amikor az összegzés egy lekérdezés eredményeként jelenik meg, akkor a segítségével meghatározható, hogy érdemes-e követni a dokumentum hivatkozását.

Az összegzés eredménye az összegzést egyetlen karaktersorozatként, illetve egy adatszerkezetként (mátrix) tartalmazza, ezek segítségével az alkalmazások kiválaszthatnak önálló mondatokat, és megállapíthatják, hogy ezek egymás után állnak-e.

Az Összegzési szolgáltatás többféleképpen is használható. A használat módja szabályozza, hogyan határozza meg a **maxLength** és a **ratio** érték az összegzés hosszát.

- **maxLength:**

Az összegzésben szereplő mondatok maximális száma. A létrejött összegzés nem lesz hosszabb a **maxLength** értéknél. Az alapértelmezett érték 3.

- **ratio:**

a mondatok száma a dokumentum teljes hosszára való tekintettel. A létrejött összegzés hosszát a dokumentum teljes hossza határozza meg. Az alapértelmezett érték 0,1.

- **mode:**

Megadja a **maxLength** és a **ratio** közti viszonyt az összegzés hosszának beállításához. A használható módok:

- **MODE\_LESS\_THAN\_MAXLENGTH:**

Az összegzés legfeljebb **maxLength** számú mondatból állhat. Ez az alapértelmezett mód.

- **MODE\_EQUALS\_RATIO:**

Az összegzés pontosan **ratio** számú mondatból áll. Az értéket a **ratio** és a dokumentumban található mondatok teljes számának szorzata adja meg.

- `MODE_EQUALS_RATIO_BUT_AT_MOST_MAXLENGTH`:

Az összegzés legalább **ratio** számú mondatból áll (a **ratio** és a dokumentumban szereplő összes mondat számának szorzata), de nem lehet hosszabb, mint **maxLength** mondat.

A támogatott nyelveket a “Támogatott nyelvek és formátumok” oldalszám: 49 sorolja fel.

## Információkinyerés

A dokumentumok elemzésének fontos feladata a dokumentum tartalmáról információkat nyújtó elemek kinyerése. Ezek az elemek az alábbiakra használhatók fel:

- Fontos információk jelzésével segíthetnek eldönteni, hogy egy dokumentum értékes-e egy adott szempontból.
- Kulcselemek megkeresése és tárolása a lekérdezések finomítása érdekében.
- Kapcsolódó dokumentumok összegyűjtési feltételeként.

Kulcselemek lehetnek például szótári szavak, vagyis szavak, nevek vagy több szóból álló kifejezések.

Angol nyelv esetén az Információkinyerési szolgáltatás normalizálja a talált kulcselemeket, és csoportosítja ezek előfordulásait, amennyiben ezek a szövegben azonos dologra utalnak, vagy azonos elvet fejeznek ki. Ha például egy dokumentumban előfordul a James J. Smith, a Mr. Smith, a James és a Smith név is, akkor a szolgáltatás ezeket azonos normalizált egyedre képezi le, mivel az összes azonos személyre utal. A szavak ragozott alakjai szintén egy normalizált egyedre kerülnek leképezésre, ilyen például a children és a child.

Japán nyelv esetén azonban a kulcselemek az előfordulásuknak megfelelően kerülnek kinyerésre. Az ISO8601 és ISO4217 szabványokban meghatározott dátum, idő és fizetőeszköz kifejezések kivételével nem történik normalizálás.

Az Információkinyerési szolgáltatás a dokumentumok következő szempontok szerinti elemzéseit teszi lehetővé:

- Egy- és többszavas szótári kifejezések, például "announcement", "value", "product cycle".
- Helyek, emberek és szervezetek nevei, például "Washington", "Clinton", "Data Management Academy".
- Rövidítések, például MB (Megabyte).
- Dátumok, pénzüsszegek és számok, például 11 Jan. 1958, 01/11/58, \$30, thirty pence, 4.5, 5000.

Háromféle kinyerhető információ típus adható meg:

- Név
- Szókapcsolat
- Kifejezés

A Java szolgáltatási API segítségével emellett azonosíthatja az alábbiakban felsorolt altípusokat is, és egy megbízhatósági érték segítségével meghatározhatja, mennyire illeszkedik az altípus a kinyert szótári elemre. A megbízhatósági érték 0 (rossz) és 1 (jó) között lehet. A név, szókapcsolat és kifejezés altípusai a következők lehetnek:

- **Név**
  - Hely, például Montreal vagy London
  - Személy, például Tim Brown
  - Szervezet, például Smith and Son

- Ismeretlen, például Smashing Pumpkins, Silicon Valley, CCTV (teljes forma nélküli rövidítés)
- Egyéb, például AIS Plan, ISO Conference, Internet, Privacy Act Officer, JCAHO Performance Report
- **Szókapcsolat**
  - Meg nem határozott kifejezés, például "entertainment conglomerate", "art world", "class variable", "source code", "data definition", "process improvement initiative"
- **Kifejezés**
  - Pozitív egész, például four, fifty, 70
  - Sorszám, például fourth, fiftieth
  - Százalék, például 12%, sixty percent
  - Dátum, például 07/28/98
  - Időpont, például 18 hrs, 4 o'clock
  - Pénzösszeg, például EUR 90, thirty pounds
  - Rövidítés, például NY

Az Információkinyerés csak angol és japán nyelvű dokumentumokon működik. A Nyelvazonosítás szolgáltatást előfeldolgozási lépésként felhasználva elkülönítheti a dokumentumgyűjtemény nem angol vagy japán nyelvű dokumentumait. Az Információkinyerési szolgáltatást más bányászati funkciókkal is kombinálhatja, például felhasználhatja az Összegzési szolgáltatás előfeldolgozási lépéseként, ha például csak azokat a dokumentumokat kívánja összegezni, amely Bush elnöki, és nem Texas állam kormányzói minőségében születtek.

## Fürtözés

A Fürtözés eszköz a dokumentumgyűjteményeket úgy rendezi el, hogy a hasonló dokumentumok azonos csoportba kerülnek, és az eltérő csoportokban (fürtökben) található dokumentumok a tartalmukra nézve különböznek. Ily módon a fürtözés felhasználható nagy számosságú dokumentumgyűjtemények kapcsolódó dokumentumainak csoportokra bontására. Hasonlóképpen felhasználható az Információstrukturalási eszközzel összeállított osztályozások segédleteként is a betanítási dokumentumok fürtözésével egy adott felhasználási területre vonatkozóan. A fürtözés emellett hasznos lehet a gyűjtemények hasonló, új irányzatokra vagy technológiákra mutató dokumentumainak megkeresésében, továbbá segíthet a többszörös vagy nagyon hasonló dokumentumok azonosításában is, amelyek viszont a versenytársak elemzéséhez nyújthatnak értékes információkat.

A fürtözés ismétlődő folyamat, amely a dokumentumokat úgy szervezi fürtökbe (csoportokba), hogy az azonos fürtben található dokumentumok tartalma a lehető leginkább hasonló legyen egymáshoz, míg a fürtök egymáshoz képest amennyire lehet eltérő dokumentumokat tartalmazzanak. A fürtözés a dokumentumgyűjteményt mint egész dolgozza fel, szemben a fenti más Információbányászat szolgáltatásokkal, például a kategorizálással vagy az összegzéssel, amelyek az egyes dokumentumok szintjén fejtik ki tevékenységüket. A fürtözés az egyes dokumentumokat jellemző tulajdonságok összehasonlításával, és ezeknek a tulajdonságoknak a csoportosításával végzi a feladatát.

A fürtözési fázis során nem adhatók hozzá új dokumentumok a dokumentumkészlethez.

Ezzel kapcsolatban az alábbi tulajdonságok beállítására van lehetőség:

- **maxClusterCount**  
A visszaadott fürtök maximális száma.
- **minClusterCount**  
A visszaadott fürtök minimális száma.

- **clusterFeatureCount**

A fűrtönként visszaadott címkék (kulcsszavak) száma.

Ezek az értékek nem kötelező érvényűek a fűrtözési szolgáltatás szempontjából, csak irányvonalakat jelentenek. A Fűrtözési szolgáltatás kimenete egy eredménylista.

A fűrtözés csak angol dokumentumokon működik. A Nyelvazonosítás szolgáltatást előfeldolgozási lépésként felhasználva elkülönítheti a dokumentumgyűjtemény nem angol nyelvű dokumentumait.

## Összetett keresés

A teljes Enterprise Information Portal tartalomkiszolgálón végzett szabványos Enterprise Information Portal kereséssel szemben az úgynevezett összetett keresés csak azokat a dokumentumokat keresi, amelyek azonosítói megtalálhatók az Információstruktúrási eszközzel létrehozott katalógusban. A keresés további szűkítéséhez az összetett keresés nemcsak a szövegre keres, hanem ezt a keresést csak adott kategóriák dokumentumain hajtja végre.

Ezzel kapcsolatban az alábbi paraméterek beállítására van lehetőség:

- **catalogName:**

Ez határozza meg a keresés során használt katalógust. A katalógusokat az Információstruktúrási eszköz segítségével lehet létrehozni és betanítani.

- **maxResults:**

Az egyes lekérdezésekben visszaadott keresési eredmények maximális száma. Az alapértelmezett érték 0, ami azt jelenti, hogy minden eredmény visszaadásra kerül.

Az alábbi lekérdezési típusok adhatók ki:

1. Egyszerű szöveges keresés. Ez a keresés a szöveges lekérdezésnek megfelelő összes dokumentumot visszaadja. Az eredménylista alkalmazhatóság szerint rendezett.
2. Egyszerű kategória keresés. Ez a keresés az adott kategóriához hozzárendelt összes dokumentumot visszaadja. Az eredmények sorrendje nem meghatározott.
3. Kombinált szöveges és kategória keresés. Ez a keresés az adott kategóriához hozzárendelt összes olyan dokumentumot visszaadja, amely megfelel a szöveges lekérdezésnek. Az eredménylista alkalmazhatóság szerint rendezett.

A rendszeren kiadott összetett keresések mindig egy adott katalógusra vonatkoznak. Ezt a katalógus keresési hatókörének hívjuk. A katalógusok közti keresés nem lehetséges, mivel a katalógusok az importált dokumentumok olyan nézetét adják meg, amelyet mindenképpen figyelembe kell venni.

A lekérdezési karaktersorozat Backus-Naur formulája (formális szintaxisa) a következő:

```
lekérdezési_karaktersorozat ::= kifejezés
kifejezés                    ::= ( kifejezés )
                           ::= egyedülálló_kifejezés
                           ::= összetett_kifejezés
egyedülálló_kifejezés      ::= kategória_kifejezés
                           ::= szöveg_keresési_kifejezés
                           ::= karaktersorozat_kifejezés
                           ::= szám_kifejezés
összetett_kifejezés        ::= kifejezés bináris_logikai_operátor kifejezés
                           ::= egyoperandusú_logikai_operátor egyedülálló_kif.
kategória_kifejezés        ::= ( DIKFCategory kat._op. kat._útv._érték )
szöveg_keresési_kifejezés  ::= ( "attribútumnév" CONTAINS szöveg_ker._érték )
karaktersorozat_kifejezés  ::= ( "attribútumnév" kar.sor._műv. kar.sor._ért.)
szám_kifejezés             ::= ( "attribútumnév" alapművelet számérték )
kétváltozós_logikai_művelet ::= AND | OR
egyváltozós_logikai_művelet ::= NOT
```

kategória_művelet	::= >=   =
karaktorsorozat_művelet	::= LIKE   alapművelet
alapművelet	::= >   <   <=   >=   !=   =
kategória_útvonal_érték	::= 'kategória útvonal'
szöveg_keresési_érték	::= "'karaktorsorozat'"
karaktorsorozat_érték	::= "karaktorsorozat"
számérték	::= "egész"   "tizedes szám"

- A lezárók, karaktorsorozatok és számok az általános kifejezéseket jelentik.
- Az "=" kategória\_művelet a keresési hatókört egy kategóriára szűkíti.
- A ">=" kategória\_művelet a keresést a kategóriára és ennek részkategóriáira bővíti.
- A CONTAINS karaktorsorozat keresés tartalmazhatja az ('\_') (egy karakter) és ('%') (tetszőleges számú karakter) helyettesítő karaktereket. A \_LOB például jelentheti azt, hogy BLOB és CLOB, míg a %név jelentheti akár azt is, hogy fájlnev. A CONTAINS karaktorsorozat művelettel csak azok a séma attribútumok kereshetők, amelyek meg lettek jelölve kereshetőként, ilyen például az IKF\_CONTENT.
- A LIKE keresési karaktorsorozata is tartalmazhat helyettesítő karaktereket, csakúgy, mint az SQL esetén.
- A jelenleg támogatott attribútumneveket "Az Információbányászat alapelvei" oldalszám: 51 szakasz sorolja fel.

Lekérdezési példák:

- Egyszerű szöveges keresések:  
 ("IKF\_CONTENT" CONTAINS "'Dél-Afrika'") AND NOT  
 ("IKF\_CONTENT" CONTAINS "'Cape'")
  - Egyszerű kategória keresések:  
 ("DKIKFCATEGORY" >= "madarak/gyümölcssevők")
  - Kombinált szöveges és kategória lekérdezés:  
 ("IKF\_CONTENT" CONTAINS "'Dél-Afrika'") AND  
 ("DKIKFCATEGORY" >= "madarak/ragadozók és dögevők/sólyom")
  - Attribútum lekérdezés:  
 ("IKF\_SUMMARY" LIKE "trópusi kolibrik")
- vagy
- ("IKF\_FEATURES" LIKE "Goethe") AND ("IKF\_TITLE" = "Faust")

## Programozási illesztők

Az Információbányászat funkciói alkalmazásfejlesztési céllal a következő formákban érhetők el:

- Java szolgáltatási API
- InformációbányászatJavaBeans

A **Java szolgáltatási API** a teljes Információbányászat funkcionalitást biztosítja, kivéve a katalógusok karbantartását, ami Enterprise Information Portal szolgáltatás lévén az Információstruktúrálási eszköz része. Az API Java RMI alapján ügyfél/kiszolgáló kommunikációt biztosít.

A Java szolgáltatási alkalmazásprogram illesztőt (API) használó alkalmazások a következőket tehetik meg:

- Dokumentum nyelvének meghatározása
- Szöveges dokumentumok összegzésének elkészítése
- Kategóriák hozzárendelése a dokumentumokhoz



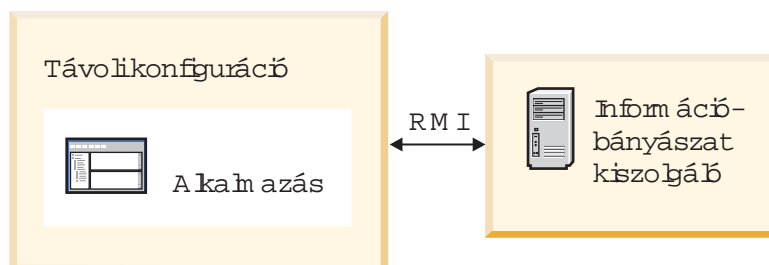
- Információk, például nevek, szókapcsolatok vagy kifejezések kinyerése a szöveges dokumentumokból
- Hasonló dokumentumok csoportosítása
- Katalógusban található dokumentumok metaadatainak tárolása és kikeresése
- Szöveges keresések végrehajtása bizonyos kategóriákba tartozó dokumentumokon és attribútumokon, például egy összegzésen

A Java szolgáltatási API képes a helyi futásra közvetlen metódushívások felhasználásával, illetve futhat távoli módban is a Java Távoli metódushívás (RMI) alkalmazásával. A távoli futtatás lehetővé teszi, hogy az egyik kiszolgálót webes alkalmazásokat futtató alkalmazáskiszolgálónak, egy másikat pedig a szövegelemzéseket, -indexelést és keresést végző Információbányászat kiszolgálónak állítsa be. A kiszolgáló feladatok mechanizmusa biztosítja annak lehetőségét, hogy az Információbányászat kiszolgáló (távoli számítógép) teljes feladatokat kapjon, és minden ezekkel kapcsolatos feldolgozás azon a számítógépen fusson le.

A web bejáró használatakor a megfelelő hozzáférési módszereket JavaBeans komponensek segítségével kell megvalósítani. A web bejáró elérése nem használható a Java szolgáltatási API szintjén.

Az Információbányászat Java szolgáltatási alkalmazásprogram illesztőjét (API) az *Application Programming Guide for Windows* című kiadvány írja le részletesen.

Az Információbányászat távoli konfigurációját a 8. ábra: oldalszám: 60 szemlélteti.



8. ábra: Információbányászat távoli konfiguráció

A JavaBeans meghatározásoknak megfelelő **Információbányászat Bean-ek** magasszintű Java alkalmazásprogram illesztőt biztosítanak a gyors alkalmazásfejlesztéshez. A bean-ek nem támogatják a kiszolgáló feladatok feldolgozását, így a bean-ek felhasználásával végzett alkalmazásfejlesztésnek egy számítógépen kell történnie.

A bean-ek mindegyike használható eseményvezérelt módban, néhányuk azonban biztosít közvetlenül hívható metódusokat is. A meglévő Enterprise Information Portal bean-ekkel való integrációt az Enterprise Information Portal bean-ek által használt eredmény esemény támogatása biztosítja. Ez azt jelenti, hogy az egyesített keresés és web bejáró eredmény eseményei kompatibilisek az Információbányászat szolgáltatásokkal, így az Információbányászat eredményeit az EIP is fel tudja dolgozni. Az Információbányászat bean-ek részletes leírását az *Application Programming Guide for Windows* című kiadvány tartalmazza.



---

## Első lépések

Az Enterprise Information Portal Információbányászat Első lépések olyan leíró útmutató, amely az IBM Információbányászat technológiájának egy adott üzleti környezetre alkalmazását valódi példán keresztül mutatja be a tudásszakértőknek, adminisztrátoroknak és alkalmazásfejlesztőknek. Az útmutató szerkezete a következő:

- Rövid bevezetés
- Az Első lépések megtétele előtt elvégzendő feladatok
- Minta adatok szervezése
- Adatok elérése
- Példa ügyfél használata
- Minta adatok eltávolítása
- További hivatkozások

Az Információbányászat Első lépések használatához nyissa meg a <CMBROOT>\ikf\firststeps\first\_steps.html fájlt.

Az Információbányászat Első lépések futtatásához Windows alatt válassza a Start menü **Enterprise Information Portal for Multiplatforms 8.1 → Információbányászat → Első lépések** menüpontját.

Az Enterprise Information Portal Információbányászat Első lépések a telepítés ellenőrzéseként is szolgálnak.

---

## Osztályozás összeállítása

Az Információstrukturalási eszköz egy web alapú alkalmazás, amely lehetőséget ad egy könyvtárnak nevezett katalóguskészlet létrehozásához és karbantartásához. A katalógusok tárolják a metaadatokat, és mindegyik katalógus egy osztályozáshoz tartozik, amely meghatározza a katalógusban található információk szervezését. Az osztályozás a kategóriáknak olyan hierarchikus felépítése, amely a dokumentumokat a téma szerinti tartalmuknak megfelelően osztályozza. Egy osztályozás felső szintje tartalmazhatja például az *üzlet*, *kultúra* és *sport* felső szintű kategóriákat, és a *sport* kategória továbbosztható *csapatsportok* és *egyéni sportok* alsóbb szintű kategóriákra. A *csapatsportok* egy továbbosztása megadhatja például a *foci*, *vízilabda* és *kosárlabda* részkategóriákat.

A kategóriák betanítási dokumentumainak meghatározásával, és a kategória betanításával az Információstrukturalási eszköz létrehoz egy olyan kategorizálási modellt, amelynek segítségével a kategorizálási eszköz kategóriákat rendelhet a dokumentumokhoz.

## Az Információstrukturalási eszköz telepítése

Az Információstrukturalási eszközt kiszolgáló kisalkalmazás tárolóba, például WebSphere alkalmazáskiszolgáló kiszolgáló kisalkalmazás alrendszerre kell telepíteni webes alkalmazásként.

Megjegyezzük azonban, hogy két Információstrukturalási eszköz webes alkalmazás (például IST1 és IST2) nem használhatja ugyanazt az Információbányászat példányt.

Az Információstrukturalási eszköz telepítése előtt gondoskodjék a helyes WAS változat telepítéséről és beállításáról. A részleteket az *Enterprise Information Portal tervezése és telepítése* című kiadványban találja.

Az Információstrukturalási eszköz telepítéséhez szükséges felhasználói jogosultságok:

- Windows: rendszergazda jogosultság
- AIX platformon: root jogosultságok
- Sun platformon: root jogosultságok

## Kezdeti lépések

Az alkalmazás webes felületet biztosít az osztályozások meghatározásához, karbantartásához betanításához. Felülete két keretből áll. A bal oldali keret neve *katalógus* nézet. Itt hozhatók létre és karbantarthatók az osztályozások. A jobb oldali keret, a *jegyzetömb* jeleníti meg az alkalmazás használatára vonatkozó információkat. A párbeszédablak lapjain taníthatók be és értékelhetők ki a katalógus dokumentumai. Az alkalmazás használatához először létre kell hozni a katalógusokat és kategóriákat a bal oldali katalógus keretben.

## Hozzáférési jogok

Az Információstrukturálási eszközben, a felhasználói nevet és jelszót az Enterprise Information Portal kezeli. Csak olyan felhasználó és jelszó adható meg, amelyet az Enterprise Information Portal ismer.

Az Információstrukturálási eszköz indításakor megjelenik egy biztonsági üzenet, amely arra figyelmeztet, hogy a betanítási dokumentumok feltöltését végző Java kisalkalmazás olvasási hozzáférést igényel a fájlrendszerhez. Ha megtagadja a jogosultságot, akkor nem lesz képes a betanítási dokumentumok feltöltésére, és így módon az osztályozások betanítására.

Az Információstrukturálási eszköz többfelhasználós környezetben is tud futni. Ahhoz, hogy több felhasználó is képes legyen megjeleníteni ugyanazt az osztályozást, az Információstrukturálási eszköz egy zárolási mechanizmust biztosít a katalógus és a benne található kategóriák hozzáféréseinek felügyelete érdekében.

A felhasználók kifejezetten megadhatják a zárolást a katalógus használatának megkezdése előtt annak kijelölésével és zárolásával, vagy megkezdhetik a katalógussal végzett munkát és a dokumentumok betanítását, amely esetben a rendszer automatikusan zárolja a katalógust a hozzáférési ütközések megelőzése érdekében. A többi felhasználó ilyenkor megtekintheti ugyan a katalógust, de a zárolás feloldásáig semmilyen módosítást nem hajthatnak végre.

Megjegyezzük, hogy ha az Információstrukturálási eszközt futtató alkalmazáskiszolgáló leáll, akkor minden zárolás törlődik.

## Osztályozás meghatározása

A katalógus az osztályozások rögzített formája, amely fa struktúrában elhelyezkedő kategóriákból áll.

Új katalógus meghatározásának és a megfelelő kategóriák kiválasztásának lépései:

1. A meghatározni kívánt kategóriák meghatározása, és az új osztályozás létrehozása.

A katalógus nézetben válassza ki a **Könyvtár** elemet, majd kattintson a jobb oldali egérgombbal. Megjelenik egy menü. Válassza az **Új katalógus** menüpontot, melynek hatására létrejön egy katalógus ikon. Az ikon átnevezéséhez írja be a katalógus nevét, majd kattintson a **Beírás** gombra. Létrejön egy azonos nevű mappa is. Ez a gyökér kategória. A jegyzetömb nézet tartalma megváltozik.

Új katalógusnak a könyvtárhoz adására egy másik módszer az Információstrukturálási eszközön kívül, például a fájlrendszerben létrehozott meglévő osztályozások importálása. További információk: "Betanítási dokumentumok feltöltése" oldalszám: 64.

Az osztályozások tevékeny használatához a katalógust zárolnia kell. A katalógus létrehozásakor ez automatikusan bekövetkezik. Meglévő katalógus zárolásához válassza ki a katalógust, majd kattintson a jobb oldali egérgombbal. Megjelenik egy menü.

Válassza a **Katalógus zárolása** menüpontot, melynek hatására a katalógus állapotikonja megváltozik. A katalógus állapotát az ikonok változása jelzi:



Az osztályozás fa össze van húzva, és egyik felhasználó sem zárolja.



Az osztályozás fa ki van bontva, és egyik felhasználó sem zárolja.



Az osztályozási fa össze van húzva, és az aktuális felhasználó zárolja.



Az osztályozási fa ki van bontva, és az aktuális felhasználó zárolja.

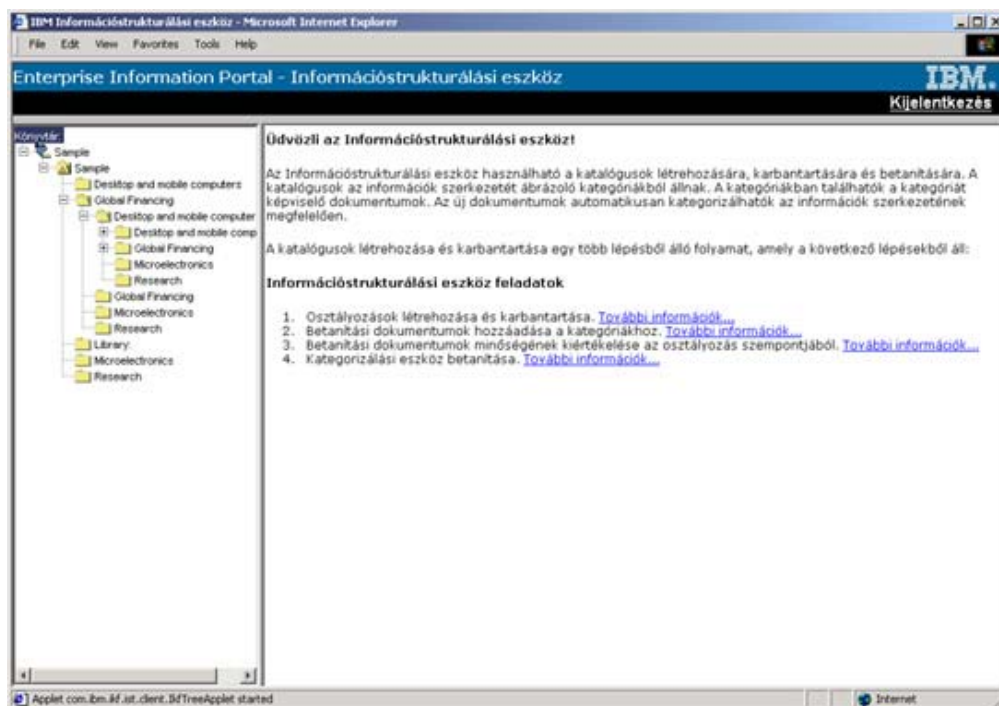


Az osztályozási fa össze van húzva, és egy másik felhasználó zárolja.



Az osztályozási fa ki van bontva, és egy másik felhasználó zárolja.

A 9. ábra: két katalógust mutat be példaként.



9. ábra: Katalógus példa

2. A katalógus létrehozása és zárolása után lehetőség van a kategóriák hozzáadására, átnevezésére vagy törlésére. Új katalógus létrehozásakor a katalógussal megegyező néven

létrejön egy gyökér kategória is. Ez a kategória átnevezhető. Ehhez válassza ki a kategóriát, majd kattintson a jobb oldali egérgombbal. Válassza az előugró menü **Átnevezés** menüpontját, írja be az új nevet, majd nyomja meg az **Enter** billentyűt.

Új kategóriák hozzáadásához jelölje ki a katalógust, amelyhez új részkategóriát kíván hozzáadni, majd kattintson a jobb egérgombbal. Válassza az előugró menü **Új kategória** menüpontját, írja be az új nevet, majd nyomja meg az **Enter** billentyűt. A fa szerkezet azonos szintjén lévő kategóriák neveinek egyedinek kell lenniük.

A gyökér kategória nem törölhető. Ennek törlésére csak a katalógus törlésekor kerül sor. A kategóriák törlése valamennyi alkategória, betanítási dokumentum és rekord törlésével jár.

3. A kijelölt katalógusra vonatkozó további jellemző információk a **Tulajdonságok** lapon jelennek meg. Az első alkalommal a leírás mező üres. Katalógus leírás hozzáadásához vagy módosításához kattintson a **Leírás szerkesztése** gombra. Megjelenik egy ablak, amelyben megadható a leírás.

## Betanítási dokumentumok kiválasztása

A kategorizálási modellek minősége erősen függ az egyes kategóriákhoz rendelt betanítási dokumentumok minőségétől.

A betanítási dokumentumoknak az Enterprise Information Portal által támogatott formátumok valamelyikében kell lenniük. A támogatott formátumok listáját a 9. fejezet, "Dokumentumformátumok", oldalszám: 117 szakasz sorolja fel.

A megfelelő betanítási dokumentumok kiválasztása rendkívül fontos. A dokumentumoknak meg kell felelniük a következő feltételeknek:

- Képviseli a kategóriát.
- Jelentékeny mennyiségű leíró szöveget tartalmaz, és nincs benne túl sok jelölőszó vagy szólista.
- Azonos stílusban íródott, tehát nincs prózai stílusban írt szöveg a jelentési stílusban írt dokumentumok között.
- Mindegyik hasonló hosszúságú, és lehetőleg nem túl hosszú. A betanítási dokumentumok hosszának közelítőleg meg kell egyeznie a kategorizálni kívánt dokumentumok hosszával.

Kategóriánként megközelítőleg 40 betanítási dokumentum ajánlott. Ha egy adott kategória általánosabb, akkor több dokumentumra van szükség. A kategóriákat figyelmesen kell kiválasztani, és sokatmondóknak kell lenniük. Az emberi áttekintésre is pontatlan és határozatlan hatást keltő kategóriák és betanítási dokumentumok minden bizonnyal az automatikus feldolgozás során is problémákat fognak felvetni.

## Betanítási dokumentumok feltöltése

Betanítási dokumentumok hozzáadásához válassza ki a megfelelő kategóriát, melynek hatására megjelenik a **Betanítási dokumentumok listája** ablak. Ha egy dokumentumot hozzá kíván adni a listához, akkor kattintson a **Dokumentum hozzáadása...** gombra.

Megjelenik a **Betanítási dokumentum hozzáadása** ablak. Az ablakot nem kell bezárnia a fájlok feltöltésének befejezése után, hanem felhasználható más kategóriák vagy katalógusok további dokumentumainak feltöltésére is.

Betanítási dokumentumok hozzáadásához kattintson a **Tallózás...** gombra, majd a **Megnyitás** ablakban válassza ki a megfelelő fájlokat vagy egy könyvtárat.

Feltöltésre egy vagy több azonos könyvtárban található fájl, illetve egy teljes könyvtár is megadható. Ha a kijelölt könyvtár üres, akkor hibaüzenet jelenik meg.

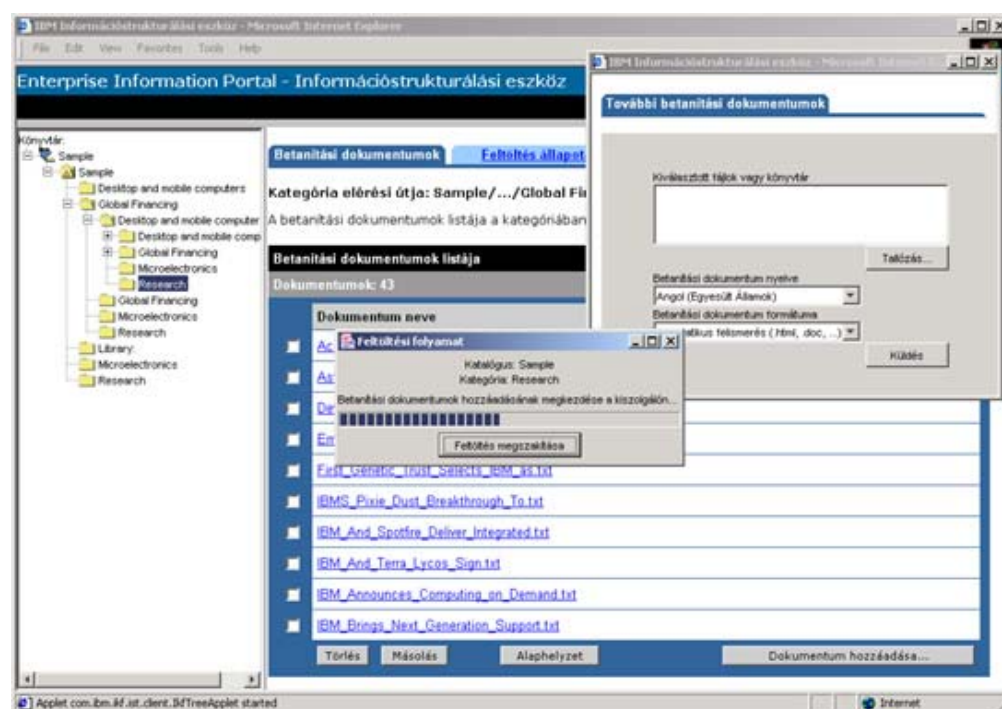
Ez lehetővé teszi az Információstrukturalási eszközön kívül, például a fájlrendszerben létrehozott meglévő osztályozások importálását és kezelését.

Ha a fájlrendszerben kijelölt könyvtár, például Fejlesztés tartalmaz egy Tervezés nevű alkönyvtárat, és az osztályozási fa Fejlesztés nevű kategóriájának is van Tervezés részkategóriája, akkor az alkönyvtárban található fájlok a részkategóriába kerülnek feltöltésre. Ha a részkategória nem létezik, akkor létrejön, és a fájlok az újonnan létrehozott részkategóriába kerülnek.

Válassza ki a kijelölt fájlok nyelvét és formátumát. A sima szövegfájlok kivételével mindig az "automatikus felismerés" beállítást adja meg.

A fájlok hozzáadásához a betanítási dokumentumok listájához kattintson az **Elküldés** gombra.

A betanítási dokumentumokat a 10. ábra: mutatja be.



10. ábra: Betanítási dokumentumok hozzáadása

Megjelenik a **Feltöltési folyamat** ablak. Az ablakban a feltöltési folyamat megszakítható. A feltöltési folyamat megszakításakor további fájlok hozzáadására nem kerül sor. A már hozzáadott betanítási dokumentumok eltávolításra kerülnek. A feltöltés során létrejött részkategóriák viszont nem kerülnek törlésre a katalógusból.

Ha a fájlok nem tölthetők fel, akkor automatikusan megjelenik egy **Feltöltés állapota** ablak a sikertelen feltöltés okára vonatkozó információkkal. Például:

- Egy azonos nevű fájl már létezik.
- A fájl üres.
- A fájlt nem lehet feltölteni a kiszolgálóra.

A Feltöltés állapota ablak **Betanítási dokumentumok** lapjának kiválasztásával térjen vissza a Betanítási dokumentumok listához. Itt jelennek meg a kategóriába sikerrel feltöltött betanítási dokumentumok. A lista összes dokumentumának megjelenítéséhez használja az **Előző** és **Következő** gombokat.

Ha egy adott dokumentumot több kategóriához is hozzá kíván adni betanítási dokumentumként, akkor a fájlt csak az első kategóriába kell feltölteni, onnan átmásolható a többibe. A fájlt nem kell ismét feltölteni. A másoláshoz válassza ki az átmásolni kívánt dokumentumokat a Betanítási dokumentumok listája ablakban, majd kattintson a **Másolás** gombra. A megjelenő ablakban kattintson a **Tallózás...** gombra a dokumentumok másolási céljaként megjelölni kívánt kategóriák kiválasztásához, majd kattintson az **Elküldés** gombra.

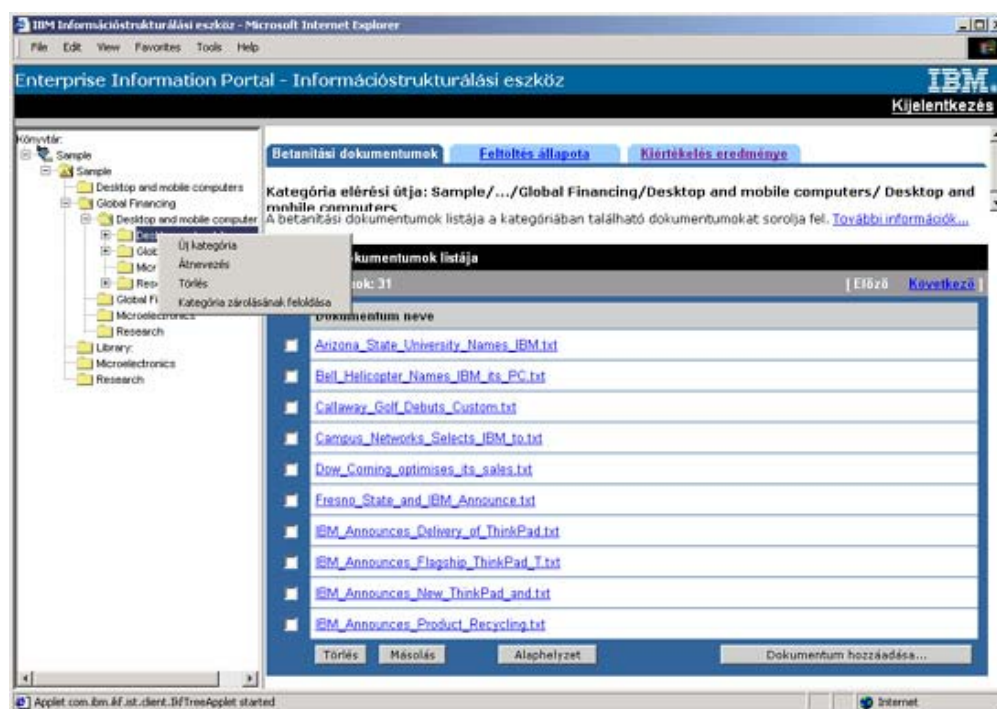
A fájlok feltöltése során az alábbi műveletek nem megengedettek:

- Katalógus zárolásának feloldása
- Kijelentkezés az Információstrukturálási eszközből
- Katalógus betanítás vagy kiértékelés indítása
- Katalógus átnevezése
- Katalógus feltöltési állapotinformációinak törlése

A következő műveletek viszont elvégezhetők:

- Másik fájl feltöltési folyamat indítása akár ugyanabba, akár egy másik kategóriába
- Másik katalógus kezelése

A betanítási dokumentumok listáját a 11. ábra mutatja be.



11. ábra: Betanítási dokumentumok listája

Ha böngészőt a fájl feltöltés során bezárja, akkor a kategóriába azok a dokumentumok kerülnek be betanítási dokumentumként, amelyek már átvitelre kerültek a kiszolgálóra. Az át nem vitt fájlok nem kerülnek hozzáadásra.



Az összes betanítási dokumentum feltöltésének befejezésekor válassza ki a katalógust a kiértékelés megkezdéséhez.

## Kategorizálási modell kiértékelése

Az osztályozások meghatározása, és az egyes kategóriák betanítási dokumentumainak hozzáadása után az osztályozást ki kell értékelni. Az osztályozás kiértékelése segít meghatározni a betanítási dokumentumok minőségét az előre meghatározott osztályozás tekintetében. Ez egy ismétlődő folyamat, amely a következő lépésekből áll:

1. Kiértékelés indítása
2. Kiértékelés eredményeinek felbecsülése
3. Osztályozás vagy a betanítási dokumentumok módosítása
4. Kiértékelés újrafuttatása

Az egyes kiértékelési szakaszokban a kiértékelési folyamat a következőket teszi:

- Felosztja a betanítási dokumentumokat egy betanítási készletre (a dokumentumok megközelítőleg 80%-a) és egy tesztkészletre (a dokumentumok megközelítőleg 20%-a). Elvégzi a katalógus betanítását a betanítási készlet alapján, majd végrehajtja a kategorizálást a tesztkészleten.
- Ellenőrzi, hogy a dokumentumok elég nagy megbízhatósági érték mellett a megfelelő kategóriákba esnek-e. A tartomány 0 és 1 között lehet, ahol az 1 az ideális dokumentumot jelöli. A megbízhatósági érték beállítható. Az alapértelmezés 0,5.

A folyamatnak 3-5 ismétlés adható meg. Az alapértelmezés a három ismétlés, amely már értékelhető eredményt biztosít az osztályozás erősségeiről és gyengeségeiről. Öt ismétlés megadásakor minden dokumentum szerepel betanítási és tesztdokumentumként is.

A kiértékelési folyamat megkezdéséhez kattintson a **Kiértékelés indítása** gombra.

Az egyes kiértékelések során az alábbi értékek kerülnek kiszámításra minden kategóriánál:

- *Helyes* dokumentumok. A C kategória dokumentumai közül a kiértékelés során a C kategóriához hozzárendelt dokumentumok száma.
- *Kimaradó* dokumentumok. A C kategória dokumentumai közül a kiértékelés során nem a C kategóriához hozzárendelt dokumentumok száma.
- *Bejövő* dokumentumok. A C kategóriához hozzárendelt, de más kategóriából származó dokumentumok száma.
- *Hozzárendelés nélküli* dokumentumok. A C kategória dokumentumai közül a kiértékelés során egyik kategóriához sem rendelhető dokumentumok száma. Ide tartozhatnak azok a dokumentumok, amelyek bár hozzárendelésre kerültek egy kategóriához, de nem érték el a beállított megbízhatósági értéket.

A kiértékelési folyamat katalógus szintű előrehaladását egy áttekintő ablak mutatja, amelyben a következő információk jelennek meg:

- A kiértékelés állapota, amely jelzi, hogy a kiértékelés fut még, vagy már leállt.
- A legutóbbi kiértékelés dátuma és időpontja.
- A befejeződött kiértékelési ismétlések száma.
- Ismétlések teljes száma.
- Átlagos összesített pontosság, amely a helyes dokumentumok százalékos arányát mutatja, vagyis az eredetileg is a kategóriához rendelt, és a bejövő dokumentumokat.
- Az átlagos összesített visszahívás, amely a helyes kategóriában található dokumentumok százalékos arányát mutatja.
- A helyes dokumentumok, vagyis a megfelelően hozzárendelt dokumentumok száma.

- A rosszul elhelyezett dokumentumok, vagyis a kimaradó és bejövő dokumentumok száma.
- Hozzárendelés nélküli dokumentumok

A pontosság és a visszahívás szorosan kapcsolódik a hozzárendelt megbízhatósági értékhez. Ha a megbízhatósági érték alacsony, akkor a pontosság leesik, és megnő a visszahívás, illetve ez visszafelé is igaz. A magas pontosság érték azt jelenti, hogy sok betanítási dokumentum került helyesen hozzárendelésre; a magas visszahívási érték arra utal, hogy a betanítási dokumentumok legnagyobb része hozzárendelésre került egy kategóriához, más szóval nincs, vagy nagyon kevés hozzárendelés nélküli dokumentum van.

## Kiértékelés eredményei

A kiértékelés részletes eredményeinek megjelenítéséhez kattintson a **Kiértékelés eredményei** lapra. A katalógus és a lap kiválasztásakor a teljes katalógus eredményei megjelennek; kategória kiválasztásakor a kategóriára vonatkozó eredmények jelennek meg.

A katalógus szintű kiértékelési eredményeket a 12. ábra szemlélteti.

Kategória	Pontosság	Visszahívás	Bejövő	Kimenő	Helyes	Nincs hozzárendelve
Sample/Research	100,0	100,0	0,0	0,0	100,0	0,0
Sample/Microelectronics	100,0	97,0	0,0	2,0	97,0	0,0
Sample/Desktop and mobile computers	87,0	81,0	10,0	8,0	72,0	8,0
Sample/Global Financing	85,0	85,0	12,0	12,0	75,0	0,0

12. ábra: Kiértékelés katalógus szintű eredményei

A kiértékelést érdemes a katalógus eredményeivel kezdeni. A vörössel (kritikus) és kézzel (kevésbé kritikus) megjelenő eredmények helytelenségre utalnak, vagy a kategóriával, vagy a betanítási dokumentumokkal kapcsolatban. Mivel a kategóriák minőségét a betanítási dokumentumok minősége határozza meg, ajánlott tanulmányozni a dokumentumok mozgását a bejövő és kimaradó besorolások között.

A kategórián illetve a betanítási dokumentumokon végrehajtandó módosítások erősen függenek attól, hogy a problémák a pontosságra és a visszahívásra is, vagy inkább csak a pontosságra vonatkoznak. Minél magasabb a pontosság érték, annál jellemzőbb a kategória az osztályozás többi kategóriájához képest. Másrésztől minél magasabb a visszahívás érték, annál kevesebb hozzárendelés nélküli betanítási dokumentum van.

A kiértékelés eredményei két szinten jeleníthetők meg:

### 1. Katalógus szint:



A katalógus minden kategóriájánál megjelennek a következők:

- A pontossági és visszahívási százalékok
- A bejövő dokumentumok százalékos aránya
- A kimaradó dokumentumok százalékos aránya
- A helyes dokumentumok százalékos aránya
- A hozzárendelés nélküli dokumentumok százalékos aránya

## 2. Kategória szint:

A kimaradó és bejövő dokumentumok esetén a következők jelennek meg:

- A betanítási dokumentumok, illetve a származási vagy célkategóriák.

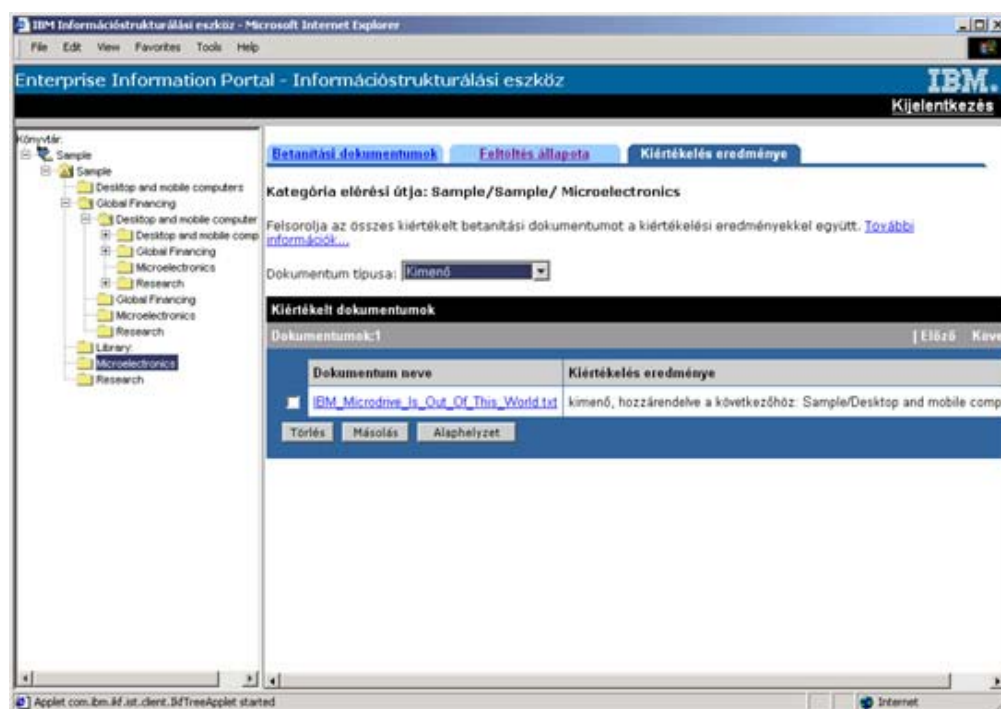
Helyes dokumentumok esetén:

- A betanítási dokumentum

## Kiértékelés eredményeinek értelmezése

Bár a következő szakasz tanácsokat ad a kiértékelés eredményeinek értelmezéséhez, sose feledje el, hogy az osztályozás egészében működik, és az egyik részében végrehajtott módosítások pont az ellenkező hatást válthatják ki az osztályozás más részein.

A kategória szintű kiértékelési eredményeket a 13. ábra: szemlélteti.



13. ábra: Kiértékelés kategória szintű eredményei

A kiértékelés katalógus szintű eredményeivel kezdődően válassza ki a viszonylag alacsony pontossági és visszahívási értékekkel rendelkező, illetve a vörössel (kritikus) vagy kézzel (kevésbé kritikus) megjelölt dokumentumokat tartalmazó kategóriákat. Valamennyi kategóriánál ellenőrizze a következőket:

- Több **bejövő** dokumentum esetén:
  - A kategóriába több más kategória dokumentumai is bekerülnek:
  - A kategória nem elég jellemző. Másolja át a dokumentumokat ebbe a kategóriába.

- Tegye a kategóriát jellemzőbbé megfelelőbb betanítási dokumentumok választásával, vagy ossza fel a kategóriát részkategóriákra.
- A kategóriába egy vagy két másik kategória dokumentumai kerülnek be:
  - Másolja át a dokumentumokat ebbe a kategóriába.
  - Ellenőrizze a származási kategóriát. Érdemes megtartani a kategóriát, vagy össze lehet őket vonni? Az összevonáshoz másolja át az összes dokumentumot a megtartani kívánt kategóriába, a másikat pedig törölje le.
- Több **kimaradó** dokumentum esetén:
  - A kategóriából több más kategóriába is kerülnek dokumentumok:
    - A kategória nem elég jellemző. Válasszon megfelelőbb betanítási dokumentumokat, vagy fontolja meg a kategória törlését.
  - A kategóriából egy vagy két másik kategóriába kerülnek dokumentumok:
    - Másolja át ezeket a dokumentumokat a másik kategóriákba.
    - Ellenőrizze a problémás kategóriát. Érdemes megtartani a kategóriát, vagy össze lehet őket vonni?
- Több **hozzárendelés nélküli** dokumentum esetén:
  - Hasonlítsa össze a hozzárendelés nélküli dokumentumokat a helyesekkel a következő szempontok szerint:
    - Méret
 

Ha a hozzárendelés nélküli dokumentumok rövidebbek, akkor például fontolja meg az összefűzésüket egy hosszabb dokumentum megalkotása érdekében.
    - Stílus
 

A stílusok eltérése esetén törölje a dokumentumot.
    - Témakör
 

Ha a témakör kissé eltérő, de még mindig kapcsolódik a helyes dokumentumok témájához, akkor fontolja meg további ilyen témájú betanítási dokumentumok keresését, amelyet adjon hozzá ugyanehhez vagy egy újonnan létrehozott kategóriához.

Ha a visszahívási és pontosság értékek megfelelők, de ugyanazok a dokumentumok több egymás utáni kiértékelés során is kimaradnak, akkor törölje ezeket a dokumentumokat.

A kiértékelési folyamatot csak akkor kell újraindítani, ha az osztályozás megváltozott, például letörölt egy kategóriát, összevont kategóriákat, vagy nagy mennyiségű dokumentumot helyezett át kategóriák között. A kategóriák kisebb módosításai, például új betanítási dokumentumok hozzáadása vagy dokumentumok áthelyezése több kategórián is elvégezhető, mielőtt a kiértékelési folyamatot újra kellene indítani. Ne feledje, hogy nincs visszavonás funkció, tehát a módosítások nem vonhatók vissza.

A kiértékelés akkor tekinthető befejezettnek, ha nincsenek vörössel vagy kézzel megjelölt kategóriák, illetve ha 90% feletti pontosság és visszahívás értéket sikerült elérnie.

## Katalógus betanítása

Az osztályozás megfelelő eredményeket felmutató kiértékelése után el kell végezni a betanítást az összes betanítási dokumentumon a kategorizálási modellnek nevezett metaadatok létrehozásához, amelynek alapján az Információbányászat kategorizálási szolgáltatása az új dokumentumok kategorizálását elvégzi.

A betanítási fázis megkezdéséhez válassza ki a betanítási dokumentumokat tartalmazó katalógust, kattintson a **Betanítás** lapra, majd a **Betanítás indítása** gombra.

A jobb oldali keret tetején megjelenik a kijelölt katalógus neve. Emellett látható a katalógus betanításának állapota is. A betanítás állapota a következők valamelyike lehet:

- A dokumentumok nem kategorizálhatók. Végezze el a katalógus betanítását a kategorizáló frissítéséhez. (Ez az alapértelmezett az újonnan létrehozott katalógusoknál is. Ebben az esetben a **Legutóbbi betanítás** dátuma üres.) A katalógus új, és még nem került sor a betanítására, vagy a legutóbbi betanítás eredményei érvénytelenek, mivel a kategóriát átnevezték vagy törölték a katalógus legutóbbi betanítása óta. A kategorizálási szolgáltatás használata új dokumentumok kategorizálásának megkísérlésekor hibákat fog eredményezni.
- A katalógusban található betanítási dokumentumok megváltoztak. A dokumentumok a legutóbbi betanítás eredményei szerint kerülnek kategorizálásra. Végezze el a katalógus betanítását a kategorizáló frissítéséhez.
- A katalógus betanítása fut.
- A katalógus naprakész, nincs szükség betanításra.

A betanítási folyamat leállításához kattintson a **Betanítás leállítása** gombra. Megjegyezzük, hogy a betanítás leállításakor a katalógus betanítási folyamatát ismét le kell futtatni.

A betanítás során új betanítási dokumentumok nem tölthetők fel a katalógusba.

---

## Teljesítményhangolás

Ahogy a katalógusban jönnek létre a rekordok, a szövegesen kereshető attribútumok (például IKF\_CONTENT) szövegindexei igen nagyra nőnek, ami egy idő után a keresési teljesítmény érezhető csökkenéséhez vezet. A tárterület optimalizálása és a teljesítmény növelése érdekében a szövegindexeket rendszeres időközönként újra kell szervezni, különösen nagyobb mértékű frissítések után.

Az szövegindexek újraszervezésére a legalkalmasabb időszak a számítógép alacsony kihasználtsága, például éjszaka. Az index újraszervezés futtatásához futtassa a következő parancsfájlokat:

- Windows platformon: ...\\ikf\\lkfReorg.cmd
- AIX platformon: .../ikf/bin/lkfReorg.sh
- Solaris platformon: .../ikf/bin/lkfReorg.

A paraméterek: lkfReorg <felhasználói azonosító><jelszó><adatbázis>

---

## Az IBM Web Crawler használata

Ez a szakasz írja le az IBM Web Crawler szolgáltatás beállítását. A szolgáltatást az EIP telepítőprogramjának Szolgáltatások összetevőjével lehet telepíteni.

Az EIP 8.2 változata egy web bejárót, egy Lotus Notes bejárót, a bejárt fájlok adatainak kinyerésére egy összegzőt, HTML alapú dokumentációt, konfigurációs példákat és segédprogramokat tartalmaz. Az IBM Web Crawler (más néven GCS) a Java SDK 1.3 vagy újabb változatát igényli.

Az IBM Web Crawler Java alapú tartalom bejáró és bányászó. Tartalomra elindítva lekérdezi és feldolgozza a tartalmat.

Az IBM Web Crawler a tartalom bejárást végezheti intraneten, extraneten, az Interneten, Lotus Notes adatbázisokban natív módban vagy Domino kiszolgálón keresztül, és helyi fájlrendszereken. Az IBM Web Crawler könnyen kiegészíthető új protokollokkal. A tartalom típusa tetszőleges lehet, például HTML, Notes mellékletek és multimédia.

Az IBM Web Crawler több tartalomtípusból a metaadatokat és szövegeket is ki tudja nyerni. A HTML tartalomból kiemelhető például:

- URL
- Cím
- Szövegtörzs
- Legutóbbi módosítás időpontja
- Metajelölők, például szerző, kulcsszavak, leírás, stb.

Egy adott tartalomtípushoz többféle előre beállított bejáróból is választhat. A tartalom és/vagy a kinyert metaadatok a helyi lemezen kerülnek mentésre. Az IBM Web Crawler a Network Solutions Outside In technológiájának segítségével több, mint 200 féle tartalomtípusból tudja kinyerni a szöveget, ideális partneri szerepet vállalva így a keresési alkalmazások mellett. Az IBM Web Crawler emellett könnyen kiegészíthető új bejárókkal is.

Az IBM Web Crawler Windows NT 4.0 és Windows 2000 operációs rendszereken áll rendelkezésre. Az IBM Web Crawler fél óra alatt telepíthető, beállítható és használatba vehető. A tartalmat másodpercenként 10 fájl sebességgel szerzi meg és elemzi egy 500 MHz-es processzorra épült számítógépen. A tesztelés során 1.000.000 objektummal is elboldogult (ez 200.000 feljegyzésnek felel meg). Több felhasználót, és felhasználónként több bejárás/bányászati konfigurációt is támogat, emellett lehetővé teszi a felhasználó által előnyben részesített nyelvi beállítás használatát.

## Az IBM Web Crawler képességei

A telepítőprogram két fájlt telepít:

`x:/<telepítési_könyvtár>/run`

Az IBM Web Crawler webes változatának parancsfájlljai és példakonfigurációi.

`x:/<telepítési_könyvtár>/notes-run`

A web bejáró Notes változatának parancsfájlljai és példakonfigurációi.

`x:/<telepítési_könyvtár>/lib`

Az IBM Web Crawler .JAR, .ZIP és szűrő fájlljai.

## Az IBM Web Crawler webes változatának beállítása és futtatása

Ez a szakasz írja le, hogyan állítható be és használható az IBM Web Crawler webes tartalomhoz. A IBM Web Crawler webes változata HTTP, FTP, news és fájl kiszolgálókat tud elérni, és HTML dokumentumok mellett más objektumok összegzéseit is elő tudja állítani. Az összegzés fájlokból dokumentumonként vagy objektumonként egy jön létre, ezekbe kerülnek a metaadatok és a teljes szöveg.

### Alapkonfiguráció

Ez a szakasz adja meg az IBM Web Crawler XML formátumú konfigurációs fájljának szerkesztésére vonatkozó tudnivalókat. A kezdeti lépéseket két példa konfigurációs fájl segíti:

- A `config-db2.xml` fájl írja le az IBM Web Crawler használatát a DB2 UDB adatbáziskezelővel.
  - A `config-sample.xml` fájl írja le az IBM Web Crawler használatát a DB2 UDB nélkül.
1. Nyisson meg egy parancssort.
  2. Váltson be az IBM Web Crawler telepítési könyvtárának `run` alkönyvtárába. Ha például az IBM Web Crawler szolgáltatást Windows platformon telepítette, akkor írja be, hogy `cd x:/<cmbrroot>/gcs/run`. Ha az IBM Web Crawler AIX alatt történt, akkor használja a `cd /usr/lpp/cmb/gcs` parancsot

**Tipp:** Nagyon fontos, hogy a fájlban tartsa meg az eredeti változatát. A fájl hibái megakadályozhatják az IBM Web Crawler megfelelő működését. A szerkesztést figyelemmel végezze.

3. Ha az IBM Web Crawlert DB2 UDB adatbázissal kívánja használni (jobb méretezhető, de lassabb megoldás), akkor módosítsa a `config-db2.xml` fájlt. Ehhez írja be például az `edit config-db2.xml` parancsot a parancssorba.
4. Ha az IBM Web Crawlert nem DB2 UDB adatbázissal használja (kevésbé méretezhető, de gyorsabb megoldás), akkor kiindulásként a `config-sample.xml` fájlt használja. Ehhez írja be például az `edit config-sample.xml` parancsot a parancssorba.

Ha  $n$  darab URL bejárását szeretné futtatni adatbázis nélkül, akkor körülbelül  $n/1000$  megabyte RAM-ra lesz szüksége azon a számítógépen, amely a bejárando URL metaadatait tartalmazza. 500.000 URL bejárásához például 512 MB RAM szükséges. A memória megfelelő kihasználásához növelje a `crawlweb.bat` fájlban található JVMXmx értéket.

## Web bejáró beállítása az információbányászathoz

A web bejáró és az információbányászat együttes felhasználásához, vagyis ahhoz, hogy képes legyen információbányászati funkciók alkalmazására a bejárt dokumentumokon, a fenti minta konfigurációs fájlokban (`config-sample.xml` és `config-db2.xml`) megadottaktól eltérő alábbi egyedi konfigurációs beállítások szükségesek:

```
<globals ...
    max-urls="5000"
    temp-filepool-class="FullPathFilePool"
    summaries-dir="webspaces/ikf/disks/1/"
    summaries-filepool-class="DistributedEIPFilePool"
    ... >
...
</globals>
<group-list>
  <group ...>
    ...
    <summarizer-config>
      ...
      <resource-handler content-type="*"
                          summarizable="EipHtmlSummarizable"
                          summary-maker="EipHtmlRawSummaryMaker" />
      ...
    </summarizer-config>
    ...
  </group>
</group-list>
```

A beállítások megtalálhatók az Információbányászat összetevő minta konfigurációs fájljában (`im-crawler-config-sample.xml`) is. A fájl a következő könyvtárban található:

Windows platformon:

```
<CMBROOT>\samples\java\beans\infomining\webcrawler\
```

UNIX (AIX és Solaris) platformon:

```
<CMBROOT>/samples/java/beans/infomining/webcrawler/
```

## Az IBM Web Crawler DB2 működésnek beállítása

A DB2 működés beállításához létre kell hozni egy adatbázist. Ehhez DB2 adminisztrátori jogosultságra van szükség. Elképzelhető, hogy át kell váltania a DB2 adminisztrátor fiók használatára. Az adatbázist elnevezheti a DB2 szabályainak megfelelően, de ha az adatbázisnak nem a `gcs` nevet adja, akkor frissítenie kell a `dbname` beállítást a web bejáró konfigurációs fájljában.

Ha rendelkezik adatbázis adminisztrációs jogosultsággal, akkor az adatbázis létrehozásához futtathatja a következő parancsot egy DB2 parancssorban:

```
db -createdb <felhasználó><jelszó>[adatbázisnév]
```

Ha nem adja meg az adatbázis nevét, akkor a rendszer a **gcs** nevet fogja használni. Az adatbázis létrehozása után hozza létre az IBM Web Crawler táblákat a következő paranccsal:

```
db -createtables <felhasználó><jelszó>[adatbázisnév]
```

Az IBM Web Crawler adatbázist és táblákat létre kell hozni az IBM Web Crawler használatára előtt.

Az új adatbázis használatához a konfigurációs fájl **urlpool-config** szakaszában a következő beállításokat kell megadni: **dbname**:

- Az adatbázis neve (a fenti létrehozás alapján), például: **gcs**.
- Felhasználói név: a felhasználó neve, például: **db2admin**.
- Jelszó: a felhasználó jelszava, például: **db2admin**.

Adja meg az adatbázis, a felhasználói név és a jelszó tulajdonságok megfelelő értékeit. Ne módosítsa a gyorsítótár méretét és az illesztőprogramot. A fájl további részeiben adja meg a bejárás hatókörét a rendszeren.

## Bejárás hatókörének beállítása

Ezekre a konfigurációs fájl beállításokra szükség van a bejárás hatókör megadásához, függetlenül attól, hogy használ-e DB2-t vagy sem.

Ellenőrizze a következő beállításokat a **crawler-config** részben, és állítsa be az értékeket a környezetnek megfelelően.

### seed list

Legalább egy kezdeti abszolút URL. Az URL címnek elérhetőnek kell lennie. Az elérhetőséget ellenőrizze egy böngészővel, például **http://www.<webhely>.com/**.

### content-type-pattern-list

Az oldalon talált URL címeket csak akkor járja be, ha a fájl kiterjesztés megfelel ezeknek a mintáknak, például: **htm\***.

### include-pattern-list

Az oldalon talált URL címeket csak akkor járja be, ha megfelelnek ezeknek a mintáknak, például: **<webhely>.com**.

Emellett az alábbi bejegyzéseket is beállíthatja:

### recursion-depth

A kiindulási ponttól hivatkozásokban számított maximális bejárando távolság. A korlátlan mélységet a **-1** jelöli.

### exclude-pattern-list

Az oldalon talált URL címeket csak akkor járja be, ha nem felelnek meg ezeknek a mintáknak, például: **\*cgi-bin\***.

### system properties

Ha a tűzfalon keresztül szeretné végezni a bejárást egy Socks támogatással nem rendelkező gépről, akkor a fájl **socksProxy** értékeit is meg kell adni.

## Az IBM Web Crawler indítása

Ha módosította az **.xml** konfigurációs fájlt, akkor mentse el.

Az IBM Web Crawler indításához használja a **crawlweb** parancsfájlt és egy konfigurációs fájlt. Nyisson egy parancssort, majd írja be a következő parancsot:

- Windows esetén: **crawlweb.bat <konfigurációs\_fájl>**
- AIX esetén: **crawlweb.sh <konfigurációs\_fájl>**



DB2 UDB adatbázist használó bejárás indításához írja be a `crawlweb config-db2.xml` parancsot, majd nyomja meg az Entert. Ha a futtatás DB2 UDB használata nélkül történik, akkor írja be a `crawlweb config-sample.xml` parancsot, majd nyomja meg az Entert.

**Tipp:** A bejárás/összegzési folyamatot tervezze úgy, hogy rendszeres időközönként jelentés készüljön. A célok bejárásakor az összegzések a `summaries-dir` paraméterben beállított helyre kerülnek. Az alapértelmezett összesítők az eredeti objektumot és a metaadat előtagot .HTML fájlként helyezik el egy fában. A bejárás során vagy után további információkat a naplófájlból szerezhet.

### További beállítások

Ezen a ponton érdemes megismerkedni a konfigurációs beállításokkal. Konfigurációs példaként tekintse meg a 7. fejezet, “IBM Web Crawler mintafájlok”, oldalszám: 103 szakaszban található `config-sample2.xml` fájlt. A példa az alábbi konfigurációt mutatja be:

- Bejáró és összegző szálak
- Grafikus figyelő
- Naplózási beállítások
- SOCKS
- Lotus Domino bejárás
- Többféle tartalomtípus
- További kizárások
- Az InsoSummarizable használata az objektumok, például .pdf fájlok összegzésének megszerzéséhez

A konfigurációs fájlban használható paraméterek formális definícióját a `config.dtd` fájlban találja. **Tanács:** Ne módosítsa ezt a fájlt. Készítsen egy másolatot a fájlról, és a másolatot nevezze át.

## Az IBM Web Crawler konfigurációs fájl

A konfigurációs fájl egy XML fájl, ez írja le az IBM Web Crawler számára az összegyűjtendő webes erőforrásokat és az összegzésük módját. Ez a rész a `config.xml` fájlban megadható elemeket és attribútumokat írja le. Az IBM Web Crawler Notes változatának használatáról az “IBM Web Crawler Notes változata” oldalszám: 86 szakaszban olvashat.

Az IBM Web Crawler ellenőrzi, hogy a konfigurációs fájl tartalma megfelel-e a `gcs-config.dtd` fájl meghatározásainak. Jelentős hibák esetén, például bejárandó URL hiányában az IBM Web Crawler kilép, és hibaüzenetet ír ki. Kisebb problémánál (ismeretlen attribútumok vagy értékek) a program figyelmeztetést naplóz a naplófájlban, és folytatja a feldolgozást. **Tanács:** A szerkesztés előtt mentse el a konfigurációs fájlt. A fájl hibái megakadályozhatják az IBM Web Crawler megfelelő működését.

A példa konfigurációs fájlok az IBM Web Crawler telepítés részei.

### <gcs-config>

A `gcs-config` fájl két szakaszból áll: **globals** és **group-list**. A `gcs-config` fájlra a 7. fejezet, “IBM Web Crawler mintafájlok”, oldalszám: 103 szakasz mutat be egy példát.

**globals** A `globals` elem az IBM Web Crawler beállításait írja le, például a fájlrendszerrel, teljesítménnyel és hálózati információkkal kapcsolatban.

### group-list

A `group-list` elem a csoportok bejárását és összegzését állítja be, ahol a csoportok erőforráskészletek, például üzleti vagy hálózati tartományok lehetnek.

## <globals>

A globals elemek képviselik az IBM Web Crawler globális beállításait. A beállítások globális attribútumokként és utód elemekként vannak megadva.

A globális attribútumokat az alábbi lista sorolja fel. A globális leszármazott elemeket a “<logger-config>” oldalszám: 77 szakasz tartalmazza.

### max-urls

A bejárható URL-ek maximális száma. Pozitív egész számnak kell lennie. Alapértelmezett értéke: 100000.

### summaries-dir

A könyvtár, ahová az erőforrás összegzések kiírásra kerülnek. Alapértelmezett értéke a summaries/ könyvtár.

### summaries-filepool-class

Az erőforrás összegzésekhez használt fájlátroló típusa. Meghatározza az összegző fájlok elnevezését, és a használt könyvtár szerkezetet (ha van). Az alapértelmezett beállítás a FullPathFilePool, amely létrehoz egy könyvtárat a hosztnak, majd az URL címekben szereplő alkönyvtár szerkezetet és fájlneveket használja.

### num-crawlers

Bejárási szálak száma. Pozitív egész számnak kell lennie. Alapértelmezett értéke: 20.

### num-summarizers

Összegző szálak száma. Pozitív egész számnak kell lennie. Alapértelmezett értéke: 5. A num-crawlers és num-summarizers beállításakor kövesse az alábbi lépéseket:

1. A bejárók számát a rendszer sebességéhez képest adja meg úgy, hogy a processzorfrekvenciát hússzal osztja. Egy 600 MHz-es processzorral rendelkező rendszeren például ez 30.
2. Az összegzők számát állítsa az 1. lépésben megadott érték negyedére, például: 8.
3. Végezzen egy próbafuttatást úgy, hogy közben figyeli a Windows Feladatkezelő Teljesítmény lapját. Ha a processzor kihasználtsága bármikor egy másodpercnél hosszabb időre 100%, akkor térjen vissza az első lépéshez, és állítsa be az eredeti értéknek mondjuk a háromnegyedét.

Ha a próbafuttatás során a szövegfigyelő Összegző: Teendő számai jóval elmaradnak a megadott összegzők számától, akkor csökkentse az összegzők számát (a kevesebb jobb), és növelje a bejárók számát (a több jobb) a 3. lépés figyelembe tartásával. A legjobb teljesítmény elérése érdekében használja a lehető leggyorsabb hálózatot, az összegzéseket, az adatbázist, az ideiglenes tároló területet és a naplót pedig különböző lemezekre helyezze el, amennyiben erre lehetőség van.

### text-monitor

Ha be van kapcsolva, akkor a szövegfigyelő minden öt másodpercben a szabványos kimenetre nyomtatja az IBM Web Crawler állapotát. Ha egy tizedes érték van megadva, akkor a szövegfigyelő beállítja a szövegkimenetek közötti frissítési időt (másodpercekben). Az alapértelmezett beállítás az off.

### graph-monitor

Ha be van kapcsolva, akkor a grafikonfigyelő utasítja az IBM Web Crawler szolgáltatást, hogy az állapotát egy grafikus felhasználói felületen jelenítse meg. Ha egy tizedes érték van megadva, akkor a grafikonfigyelő beállítja a figyelő grafikus felhasználói felületek közötti frissítési időt (másodpercben). Az alapértelmezett beállítás az off.

**log-file** Megadja a használandó elsődleges naplófájlt. Az alapértelmezés log/log.txt

**Tipp:** A logger-config elemben további naplózási információkat határozhat meg.



**log-priority**

Beállítja az alapértelmezett naplózási prioritást. Adja meg az info, warn vagy error beállítást. Az alapértelmezett beállítás a warn.

**Tipp:** A logger-config elemben további naplózási információkat határozhat meg.

**temp-dir**

Az ideiglenes fájlok tárolására használt könyvtár neve. **Tipp:** Az IBM Web Crawler a könyvtár összes fájlját törölheti. Ne módosítsa az alapértelmezett `x:/temp/gcs` beállítást.

**temp-filepool-class**

Az ideiglenes fájlokhoz használandó fájl tároló típusa. **Tanács:** Ne módosítsa az alapértelmezett TempFilePool beállítást.

**content-dir**

Az IBM Web Crawler által a tartalom fájlok kiírására használt könyvtár. A content-dir általában megegyezik a temp-dir beállítással.

**content-filepool-class**

A tartalom fájlokhoz használandó fájl tároló típusa. Ez általában megegyezik a temp-filepool-class beállítással.

**how-often-to-gc**

Bejárható URL-ek száma a tárterület felszabadítási kérések között. **Tanács:** Adjon meg egy 50-et vagy egy 50-nél nagyobb egész számot. Az alapértelmezett beállítás a 100.

**max-resource-pool-size**

Az összegzésre várakozó erőforrások maximális sormérete. **Tanács:** Adjon meg 10-et vagy egy 10-nél nagyobb egész számot. Az alapértelmezett beállítás 10 várakozó erőforrást engedélyez összegzőnként.

**connect-timeout**

Az a várakozási idő ezredmásodpercekben, amely után a hálózati kapcsolat túllépi az időkorlátot. Az alapértelmezett beállítás a 4000. A megadható értékek: 1000-60000.

**read-timeout**

Az a várakozási idő ezredmásodpercekben, amely után a hálózati olvasás túllépi az időkorlátot. Az alapértelmezett beállítás a 6000. A megadható értékek: 1000-60000.

**cookies**

Ez a beállítás adja meg a rendszernek, hogy ellenőrizze-e a cookie-kat a HTTP fejlécekben, és tárolja-e azokat egy adatbázisban. Az alapértelmezett beállítás az off. Ha engedélyezni szeretné a cookie-kat, akkor adja meg az on beállítást.

**locale** Az összegzésekhez és a naplózáshoz használt nyelv. Az alapértelmezett érték `en_US` (amerikai angol).

A globális leszármazott elemek közé a logger-configs, az urlpool-config és a system-properties tartozik.

**<logger-config>**

A logger-config fájl további beállításokat biztosít a naplózásra kerülő információkkal, az információk formázásával és a naplófájl helyével kapcsolatban. A default log-file és a log-priority globális attribútumként vannak megadva. A naplózásról további információkat "Az IBM Web Crawler naplózása" oldalszám: 83 szakaszban talál.

**category**

A beállítani kívánt naplózó kategóriája, például `gcs.crawler`. Ha nincs megadva, akkor a beállítás az alapértelmezett naplózóra vonatkozik. Ne feledje, hogy egy adott kategória beállításai az összes leszármazott kategóriára is vonatkoznak.

**priority**

Az a minimális prioritás, amellyel egy üzenetnek rendelkeznie kell a naplózáshoz. Ha nincs megadva, akkor a naplózó a prioritást a szülő kategóriából veszi (tehát végső soron a log-priority globális alapértelmezéséből).

**log-file**

A naplófájl helyét adja meg. Ha az érték '+' jellel kezdődik, akkor naplófájl a más (szülő) naplófájlok mellett kerülnek felhasználásra. Ha nincs megadva, akkor a rendszer a szülő naplófájlt használja (tehát végső soron a log-file globális alapértelmezett beállítását).

**Tipp:** Ne adja meg ugyanazt a log-file beállítást több naplózónak, mert felülírják egymást.

**log-layout**

A naplófájlba kerülő üzenetek elrendezését adja meg.

**<urlpool-config>**

Az urlpool-config fájl határozza meg az IBM Web Crawlernek az URL címeket tároló összetevőjét. Az URL tárolóval kapcsolatban több beállítási lehetőség is van. A tároló lehet a memóriában, DB2 adatbázisban, vagy használhat kis mennyiségű memóriát igénylő változatot, amely az egyes URL-ek vonatkozásában nem tartalmaz túl sok információt. Ha nem ad meg urlpool-config elemet, akkor az URL-ek a memóriában tárolódnak. Az urlpool-config elemnek rendelkezhetnek urlpool-param leszármazott elemekkel különféle értékek, például adatbázis információk megadásához.

**urlcontainer-class**

A használandó URL tároló típusa. Adja meg az alábbiak valamelyikét:

- DB2URLContainer - bejárás a DB2 UDB felhasználásával.
- MemoryURLContainer - bejárás a DB2 UDB nélkül (alapértelmezett).
- BigMemoryURLContainer - bejárás a DB2 UDB nélkül, többletmemória használata (a hivatkozó URL címek és más információk tárolásához).

**urlcollection-class**

A használandó URL gyűjtemény típusa. Adja meg az alábbiak valamelyikét:

- DB2URLCollection - bejárás a DB2 UDB felhasználásával.
- MemoryURLCollection - bejárás a DB2 UDB nélkül (alapértelmezett).
- BigMemoryURLCollection - bejárás a DB2 UDB nélkül, többletmemória használata (a hivatkozó URL címek és más információk tárolásához).

**<urlpool-param>**

Ezzel adhatók át paraméterek az urlcollection-class elemnek. Egy példát a 7. fejezet, "IBM Web Crawler mintafájlok", oldalszám: 103 szakaszban bemutatott DB2 UDB adatbázist használó konfigurációs fájl tartalmaz az adatbázis kapcsolati információk megadásával.

**name** A paraméter neve.

**value** A paraméter értéke.

**Tipp:** A paramétereket különös odafigyeléssel használja, mivel ezeken nincs hibaellenőrzés.

**<system-properties>**

A system-properties fájl tartalmazza a rendszertulajdonságok beállításainak listáját.

**<property>**

Példaként tekintse át a speciális konfigurációs példa SOCKS átjárót használó konfigurációját.

**name** A paraméter neve.

**value** A paraméter értéke.

Alternatívaként az IBM Web Crawler beállítható úgy is, hogy a külső kiszolgálókhoz PROXY átjárón keresztül csatlakozik:

```
<system-properties>
  <property name="proxySet" value="true"/>
  <property name="proxyHost" value="proxy.hosztnev"/>
  <proxy port name="proxyPort" value="80"/>
</system-properties>
```

**Tipp:** A paramétereket különös odafigyeléssel használja, mivel ezeken nincs hibaellenőrzés.

### <group-list>

A group-list legalább egy csoport elemet tartalmazó lista.

### <group>

A group elem erőforrások olyan csoportját képviseli, amelyek bejárására és összegzésére hasonló módon kerül sor. Minden csoportnak egyedi név attribútummal kell rendelkeznie, emellett mindegyiknek lennie kell legalább egy crawler-config leszármazott elemének, amely a bejárando elemeket határozza meg. A group attribútum rendelkezhet leszármazott summarizer-config elemmel, ha nem kívánja használni az alapértelmezett összesítőket. **Tipp:** Az átfedésben lévő csoportok (ugyanaz az URL több csoportban is benne van) nem várt eredményekre vezethetnek. A több csoporthoz tartozó URL címeket a rendszer csak az első olyan csoporthoz társítja, amelyben az URL megtalálható.

**name:** a csoport egyedi neve (kötelező).

### <crawler-config>

Az alábbi szabályokkal adhatja meg a bejárás hatókörét. A bejáró az alapérték lista minden URL címét visszakeresi, elemzi az URL címek tartalmát, és hozzáadja a megfelelő URL címeket a bejárando címek listához:

- legalább egy szabály a content-type-pattern-list elemében
- legalább egy szabály az include-pattern-list elemében
- az exclude-pattern-list szabályok egyike sem

A crawler-config elemnek meg kell adni egy **recursion-depth** attribútumot is. A recursion-depth határozza meg, hogy a bejáró a kiindulási helytől hány hivatkozásnyira távolodhat el. Az alapértelmezett érték a -1, ami végtelen mélységet ad meg.

### <seed-list>

Ez egy URL alapérték lista, esetleg hitelesítési információkkal.

### <seed>

A seed érték egy induló URL alapértéket jelent a bejáró számára, URL attribútumokkal, és esetleg hitelesítési információkkal. Minden alapértéknek abszolút URL címnek kell lennie, például `http://<web.hely>.com/`. Kerülje az átirányított, nem elérhető vagy nem szöveges oldalakra mutató alapértékeket. Hasznos egy olyan oldalra mutatni, amelyet előzőleg szerkesztett, és tartalmazza az alapértékeket. Az ilyen oldalt könnyű frissíteni, átnézni és tesztelni egy böngészővel.

**URL** Egy alap URL a bejárás megkezdéséhez.

### <authentication>

Az RFC2617 szerinti *Alap hitelesítéssel* védett alapérték URL címeknek elküldött nem kötelező hitelesítés.

#### username

A hitelesítéshez használt felhasználói név.

**password**

A hitelesítéshez használt jelszó.

Például:

```
<seed url="http://web.hely.com/"><authentication username="én"
password="titok"/></seed>
```

**<content-type-pattern list>**

A bejárando tartalomtípusokat tartalmazó minták listája fájlkiterjesztés szerint. A lista url-name-pattern beállításainak megfelelő URL fájlkiterjesztések (.html, .gif, .doc stb.) átmennek a teszten. A kiterjesztéssel nem rendelkező URL-ek alapértelmezésben átmennek a teszten. Ha a content-type-pattern-list elem nincs megadva vagy üres, akkor a rendszer csak a fájlkiterjesztés nélküli URL-eket fogja elfogadni.

**<include-pattern lista>**

Ez a bejárando URL címeket kiszolgáló vagy tartománynév szerint tartalmazó minták listája. A lista url-obj-pattern, url-regex-pattern, url-name-pattern vagy url-predicate-pattern mintájának megfelelő URL-ek átmennek a teszten. Ha az include-pattern-list nincs megadva vagy üres, akkor a rendszer minden URL-t elfogad.

**<exclude-pattern lista>**

A bejárásból kizárt URL-ek mintáit tartalmazó lista. A lista url-obj-pattern, url-regex-pattern, url-name-pattern vagy url-predicate-pattern mintájának megfelelő URL-eket a rendszer nem járja be. Ha az exclude-pattern-list nincs megadva vagy üres, akkor a rendszer semmilyen URL-t nem fog visszautasítani.

**<url-obj-pattern>**

Ez a minta az URL különböző részeinek (protokoll, hoszt, stb.) illesztésére szolgál. Használható az exclude-pattern-list és az include-pattern-list listákban is. A minta minden egyes részében lehet egy '\*' helyettesítő karakter a minta elején vagy végén, ami bármit jelenthet. A minta közepén viszont nem lehetnek helyettesítő karakterek. A megfelelés megkülönbözteti a kis- és nagybetűket. A kihagyott URL rész minták automatikusan megfelelnek.

Az alábbi példa bemutatja, hogyan bontja szét a Java és az IBM Web Crawler a <http://www.ibm.com/products/index.html?query#ref> URL címet:

- Protokoll: http
- Hoszt: www.ibm.com
- Port: -1 (nincs megadva)
- Fájl: /products/index.html?query
- Elérési út: /products/index.html
- Könyvtár: /products/
- Fájlnev: index.html
- Kiterjesztés: .html
- Lekérdezés: query
- Hivatkozás: ref

Az alábbi lista részletes információkat tartalmaz az url-obj-pattern minden eleméről:

**protokoll**

Ennek a helyettesítő karakter mintának kell megfelelnie az URL protokolljának, például: http.

**host**

Ennek a helyettesítő karakter mintának kell megfelelnie az URL hosztjának, például: \*.ibm.com

- port** Ennek a helyettesítő karakter mintának kell megfelelnie az URL portjának, például: 80.
- file** Ennek a helyettesítő karakter mintának kell megfelelnie az URL fájlnek, például: \*.htm\*. Az URL fájl része a hoszt utáni első osztásjellel kezdődik, és tartalmazhat lekérdezést is, hivatkozást azonban nem. A `http://www.ibm.com/products/index.html?query#ref` URL *fájl* része a `/products/index.html?query`.
- path** Ennek a helyettesítő karakter mintának kell megfelelnie az URL elérési útnak, például: \*.html. Az URL elérési út része a hoszt utáni első osztásjellel kezdődik, és nem tartalmaz sem lekérdezést, sem hivatkozást. A példában a `http://www.ibm.com/products/index.html?query#ref` URL *elérési út* része a `/products/index.html`.
- dir** Ennek a helyettesítő karakter mintának kell megfelelnie az URL könyvtárnak, például `/products/`. A könyvtár az URL elérési út részének az első és utolsó osztásjel közé eső szakasza. A példában a `http://www.ibm.com/products/index.html?query#ref` URL *könyvtár* része a `/products/`. Nem tartalmaz lekérdezést vagy hivatkozást. Megjegyezzük, hogy az utolsó osztásjelet elhagyó szabálytalan URL címek, például `http://www.ibm.com/products` nem megfelelően fognak illeszkedni a könyvtár szakaszra. A rossz URL példában az URL könyvtár a `/`.
- filename** Ennek a helyettesítő karakter mintának kell megfelelnie az URL fájlnevének, például `index.html`. A fájlnev az URL utolsó osztásjel utáni szakasza. A példában a `http://www.ibm.com/products/index.html?query#ref` URL *fájlnev* része az `index.html`. Nem tartalmaz lekérdezést vagy hivatkozást.
- extension** Ennek a helyettesítő karakter mintának kell megfelelnie az URL fájlkiterjesztésének, például `htm*`. Ha lehetséges, akkor inkább használja a `content-type-pattern-list` beállítást.
- query** Ennek a helyettesítő karakter mintának kell megfelelnie az URL lekérdezésnek.
- ref** Ennek a helyettesítő karakter mintának kell megfelelnie az URL hivatkozásnak (a HTTP nem használja). Például: az `<url-obj-pattern host="*.ibm.com"/>` minden IBM weboldalnak megfelel.

### <url-regex-pattern>

Az `url-regex-pattern` egy olyan minta, amely reguláris kifejezéssel felel meg egy URL-nek. Használható az `exclude-pattern-list` és az `include-pattern-list` listában is. A `com.ibm.regex` csomagot használja (regex4j), és tartalmazza a Perl 5 reguláris kifejezés funkcionalitásának nagy részét. Két reguláris kifejezést lehet megadni, az egyiknek az URL-nek *meg kell felelnie*, a másiknak az URL *nem felelhet meg*. Emellett más paraméterek is megadhatók, például az i a kis- és nagybetűk azonosnak vételéhez. Részletes információkat a `Regex4j` reguláris kifejezések részben talál.

**match** Perl 5 stílusú reguláris kifejezés, amelynek az URL-nek meg kell felelni.

#### no-match

Perl 5 stílusú reguláris kifejezés, amelynek az URL nem felelhet meg.

#### options

Elhagyható módosítók, például az `i`, ami a kis- és nagybetűk azonosságát adja meg.

Például: az `<url-regex-pattern match="^http://www\.ibm\.com/.*\.html?$"/>` a fő IBM weboldal összes HTML oldalát jelenti.

### **<url-name-pattern>**

Ez egy példa minta helyettesítő karakterekkel, amely egy teljes URL-nek vagy egy URL fájlkiterjesztésnek felel meg. A `content-type-pattern-list`, az `include-pattern-list` és az `exclude-pattern-list` listában használható. Lehet benne egy `'*` karakter a minta karakterlánc elején vagy végén, ami bármit jelenthet. A név közepén viszont nem lehetnek helyettesítő karakterek. A megfelelés megkülönbözteti a kis- és nagybetűket.

Például: az `<url-name-pattern name="*.ibm.com/*"/>` az IBM weblap összes fájlját jelenti, de az `<url-name-pattern name="*.ibm.com/*.html"/>` érvénytelen, mert helyettesítő karakter van a karakterlánc közepén.

**name** Helyettesítő karakteres minta, amelynek az URL karakterláncnak meg kell felelnie, nem kötelező `'*` helyettesítő karakterrel a karakterlánc elején és/vagy végén.

### **<url-predicate-pattern>**

Ez a minta betölti az egyező URL egyik Java `UnaryPredicate` osztályát. Az `exclude-pattern-list` és az `include-pattern-list` listában is használható. Az osztálynak rendelkeznie kell egy `public boolean execute(URL url)` metódussal, amely `true` értéket ad vissza, ha az URL megfelel a predikátumnak.

**class** Teljesen meghatározott `UnaryPredicate` osztálynév.

### **<summarizer-config>**

Az összegző konfigurációja, amely tartalmazza az utód erőforráskezelőket. Jelenleg csak egy `summarizer-config` használható csoportonként.

### **<resource-handler>**

A tartalomtípus (például `text/html`) vagy a fájlkiterjesztés (például `htm`) alapján meghatározza, hogy milyen típusú összegzések jönnek létre az erőforrásokhoz (például weboldalakhoz vagy hírcsoport cikkekhez). Ha egy erőforrás készen áll az összegzésre, akkor az IBM Web Crawler ellenőrzi az erőforráskezelőket, és az első olyan erőforráskezelőt használja, amely megfelel a tartalomtípusnak vagy a fájlkiterjesztésnek. Ha nincs egyezés, akkor a rendszer a `Copy2RdfSummarizable` és a `Copy2RdfSummaryMaker` kezelőket használja alapértelmezésben. Ezt a beállítást felülbírálhatja, ha olyan erőforráskezelőt ad hozzá a lista végére, amelynek nincs `content-type` vagy `file-extension` jellemzője).

A `resource-handler` rendelkezhet `summarizer-param` leszármazottal, amely speciális paramétereket adhat át a saját `SummaryMaker` osztályának.

#### **content-type**

Helyettesítő karaktereket tartalmazó minta, amelynek a tartalomtípusnak meg kell felelni. Például: `*htm*`.

#### **file-extension**

Helyettesítő karaktereket tartalmazó minta, amelynek a fájlkiterjesztésnek meg kell felelni. Például: `htm*`.

#### **summarizable**

Az erőforrás `Summarizable` osztályának neve. Például: `HtmlRawSummarizable`.

#### **summary-maker**

Az erőforrás `SummaryMaker` osztályának neve. Például: `HtmlRawSummaryMaker`.

A `content-type` és `file-extension` mintákban használhat helyettesítő karaktereket. A mintában lehet egy `*` karakter a minta karakterlánc elején vagy végén, ami bármit jelenthet. A közepén viszont nem lehetnek helyettesítő karakterek. A megfelelés megkülönbözteti a kis- és nagybetűket.

Az `resource-handler` akkor felel meg, ha a `content-type` és `file-extension` minta is megfelel. A meg nem határozott minta mindig megfelel, így a

```
<resource-handler content-type="*htm*"
summarizable="*HtmlRawSummarizable" summary-maker=
"HtmlRawSummaryMaker"/>
```

feltételeknek minden olyan fájl megfelel, amelynek tartalomtípusa `text/html`, függetlenül a fájlkiterjesztéstől.

A `summarizable` és `summary-maker` elemeknél nem kell az osztályok teljes elérési útvonalát megadni, ha az osztályok benne vannak a `com.ibm.IBM Web Crawler.summarizer.resource` csomagban.

### <summarizer-param>

A `SummaryMaker` osztálynak átadott speciális karakterek. A használata az osztályra jellemző.

**name** A paraméter neve.

**value** A paraméter értéke.

**Tipp:** A paramétereket különös odafigyeléssel használja, mivel ezeken nincs hibaellenőrzés.

## Az IBM Web Crawler naplózása

Ez egy bevezető az IBM Web Crawler naplózási szolgáltatásaiba.

Az IBM Web Crawler hatékony felügyeletet biztosít a naplózott adatok, a naplózás helye és a naplók formázása felett. Például: a bejárt oldalak válaszkódjait beírhatja egy fájlba, az IBM Web Crawler állapotát (hány URL-t járt be, hány szót használt stb.) egy másik fájlba, az összegzett URL-eket egy harmadik fájlba, az IBM Web Crawler figyelmeztetéseket egy negyedik fájlba, és a hálózati segédprogram csomag naplózási üzeneteit egy ötödik fájlba, hibakeresési céllal.

Egy példa naplóelemzési fájlt az “IBM Web Crawler napló elemzési fájl példa” oldalszám: 105 szakasz mutat be.

### Naplók használata

A naplózás hasznos a hálózat, a web, a bejárás és az összegzés megfigyeléséhez, a többi alkalmazás összetevővel való kommunikációhoz és az IBM Web Crawler hibakereséshez.

A bejárás és bányászat megfigyelése számos érdekes jelenséget hozhat előtérbe, például a rosszul konfigurált kiszolgálókat, a hiányzó oldalakat és az objektumok számát tartalomtípusonként. A `loganalysis.pl` Perl parancsfájl egy példa napló összegző értékelést biztosít. Elképzelhető, hogy az alkalmazásoknak információkra van szüksége az IBM Web Crawler szolgáltatástól, például a tartalom eltávolításának időpontjára vonatkozóan.

### Naplózók beállítása

Az IBM Web Crawler konfigurációs fájlban egynél több naplózó beállításait is megadhatja. A `globals` elem `log-priority` és `log-file` attribútuma határozza meg az alapértelmezett naplózási stratégiát.

Ha ki szeretné terjeszteni a naplózási stratégiát, akkor hozzon létre `logger-config` utasításokat a `globals` elem leszármazottaiként. Minden elem az IBM Web Crawler naplóüzeneteinek egy részhalmazát választja ki, átirányítja egy adott fájlba, és a megadott formátumban kiírja. A naplózott üzenetek részhalmazát a `priority` és `category` attribútumokkal választhatja ki. A választható `priority` értékek: `trace`, `debug`, `info` és `warn` (a kis- és nagybetűk eltérőnek számítanak).



- A priority érték beállítása megadja a naplózó által naplózott információkat, miközben a nyomkövetés a lehető legtöbb információt rögzíti.
- A trace és debug karbantartási szintek. Az üzenetek csak angolul jelennek meg.
- Az info és warn felhasználói szintek, ezek nemzeti nyelven jelennek meg.
- Az info szint nagyon sok üzenetet hoz létre. Az üzenetek mennyiségét a warn prioritás megadásával csökkentheti.

## Naplózási konfigurációs példák

**Forrás/cél hivatkozások naplózása dátum/idő/szál információk nélkül a log/fromto.txt fájlba**

```
<logger-config category="gcs.url.fromto" priority="info"
log-layout="%m\n" log-file="log/fromto.txt"/>
```

**Összesített objektumok naplózása a log/resources.txt fájlba**

```
logger-config category="gcs.summaries.list.resource"
priority="info" log-file="log/resources.txt"/
```

**Kihagyott URL címek naplózása a kihagyás okával**

```
<logger-config category="gcs.url.skipped"
priority="info" log-file="log/urls_skipped.txt"/>
```

**Speciálisan feldolgozott HTTP válaszkódok naplózása**

```
<logger-config category="gcs.http.302"
priority="info" log-file="log/urls_redirected.txt"/>
<logger-config category="gcs.http.404"
priority="info" log-file="log/urls_not_found.txt"/>
```

**Minden összegzési üzenet naplózása prioritással együtt**

```
<logger-config category="gcs.summarizer"
priority="TRACE" log-file="summarizer_trace.txt"
log-layout="%d: %t: %c: %p: %m\n"/>
```

## Hibaelhárítás

Ha hibákba ütközik, akkor először az alábbiakat ellenőrizze:

**Elérhetők a kiindulási oldalak?**

A lapoknak létezniük kell (érdeemes elkerülni az átirányított oldalakat), és elérhetőnek kell lenniük a rendszerről.

**A kiindulási oldalak szabványos HTML oldalak?**

A keretek, a Flash, a javascriptek és az egyéb elemek nem megfelelő alapérték választások. Válasszon szabványos HTML oldalakat.

**Ha DB2 UDB használatával végzi a bejárást, akkor sor került már a bejárásra?**

A DB2 UDB nyomon követi a már bejárt oldalakt. Ha minden oldal feldolgozásra került, akkor nem történik semmi. Új bejárás indításához használja a db-emptytables parancsot.

**Ha DB2 UDB használatával végzi a bejárást, akkor a konfigurációs fájl adatbázishozzáférési bejegyzési helyesek?**

Ha az adatbázishoz csatlakozás nem sikerül, akkor a bejárás megszakad.

**Figyelmesen ellenőrizte a konfigurációs fájl módosításait?**

A hibák megakadályozhatják az IBM Web Crawler működését. Átnézte az esetleges túl megszorító max-urls vagy recursion-depth értékeket?

**Még mindig fennáll a probléma?**

Nyissa meg a konfigurációs fájlt, és módosítsa a naplózási prioritást "debug" értékre. Indítsa el ismét a web bejárót, majd a leállás után vizsgálja meg a naplófájlt.



## Összegzők kiválasztása

Az összegző célja, hogy egy erőforrás (például egy weboldal) vagy egy hoszt (például egy webkiszolgáló) alapján egy olyan fájlt hozzon létre, amely könnyen használható formátumban tartalmazza a hasznos információkat.

Az IBM Web Crawler különböző tartalomtípusokat kezelő összegzőket tartalmaz, amelyek különböző adatokat nyerne ki a forrásokból, és különböző formátumban állítják elő a kimeneti fájlt. Ez a szakasz írja le a rendelkezésre álló összegzők szolgáltatásait és követelményeit. Ha egyik összegző sem rendelkezik olyan funkcióval, amelyre szüksége van, akkor írhat saját összegzőt is.

Az IBM Web Crawler összegzőinek két típusa van. Az erőforrás összegző egyetlen erőforrás, például egy weboldal, míg a hoszt összegző egy hoszt, például egy webkiszolgáló összegzését végzi. Egyelőre csak az erőforrás összegzőket lehet konfigurálni.

### Erőforrás összegző kiválasztása

Az erőforrás összegző kiválasztásakor vagy írásakor vegye figyelembe az alábbiakat:

- Mi a bemeneti formátum? HTML oldalak, PDF, WordPro dokumentum, XML fájl.
- Milyen metaadatokat kell kinyerni? HTTP fejléc, cím, feljegyzett hivatkozások, szövegtörzs.
- Mi a kimeneti formátum? XML, HTML, RDF.

Az IBM Web Crawler config fájljának resource-handler elemével megadhatja, hogy milyen összegzőt szeretne használni az egyes erőforrástípusokhoz. Először adja meg az összegző tartalomtípusát és/vagy fájlkiterjesztését. Utána adja meg a munkát elvégző Java Summarizable és SummaryMaker osztályokat. A summarizable osztály az összegzendő erőforrásokat jelöli, a summary-maker osztály pedig létrehozásra kerülő összegzés típusát.

### Alapértelmezett összegző (Copy + RDF Summarizer)

A copy + RDF összegzőt a rendszer azokhoz az objektumokhoz használja, amelyek tartalomtípusához nincs kifejezetten megadva más összegző. Az összegző bármilyen erőforrástípussal használható, és két fájlt ír. Az első fájl az eredeti erőforrás pontos másolata, a második fájl pedig az eredeti URL-t, a tárolt fájl fájlnevét és a HTTP fejléc információit tartalmazó RDF összegzés. Pontosan beállítható a DefaultSummarizable és Copy2RdfSummaryMaker elemekkel.

### HTML oldalak összegzője (nyers HTML összegző)

A HTML erőforrásoknál a nyers HTML összegző egyszerűen lemásolja a fájlt az URL és HTTP fejléc információkkal egy megjegyzésben az oldal tetején. A HtmlSummarizable és HtmlRawSummaryMaker elemekkel állítható be.

```
<resource-handler content-type="*htm*"
    summarizable="HtmlSummarizable"
    summary-maker="HtmlRawSummaryMaker" />
```

### HTML oldalak összegzője (EIP HTML összegző)

A HTML erőforrásoknál a nyers HTML összegző egyszerűen lemásolja a fájlt az URL és HTTP fejléc információkkal egy megjegyzésben az oldal tetején. Ez az EIPHtmlSummarizable és az EIPHtmlRawSummaryMaker paraméterekkel állítható be.

```
<resource-handler content-type="*htm*"
    summarizable="EIPHtmlSummarizable"
    summary-maker="EIPHtmlRawSummaryMaker" />
```

### HTML oldalak összegzője (nem író HTML összegző)

Ez az összegző bejárja a HTML oldalakat és követi a hivatkozásokat, de nem ír összegző fájlt. Ez akkor lehet hasznos, ha például egy oldal minden PDF fájlját be szeretné jární (az INSO -

XML összegzővel), de nem szeretné eltárolni a HTML fájlokat. A beállítása a HtmlSummarizable és a NoWriteSummaryMaker paraméterekkel lehetséges.

```
<resource-handler content-type="*htm*"
                  summarizable="InsoSummarizable"
                  summary-maker="InsoSummaryMaker" />
```

## Más tartalomtípusok összegzői (INSO - XML összegző)

Ez az összegző XML összegzést készít több, mint 200 féle erőforrástípusból, egyebek között Microsoft Word dokumentumokból, PDF fájlokból, PowerPoint bemutatókból. Az összegző rendelkezik némi metainformációval, a szövegtörzset pedig a Network Solutions INSO szűrők nyerik ki (INSO licenc szükséges hozzá). Az InsoSummarizable és az InsoSummaryMaker paraméterekkel állítható be.

```
<resource-handler content-type="pdf"
                  summarizable="InsoSummarizable"
                  summary-maker="InsoSummaryMaker" />
```

## Egyéb összegzők

Ha további erőforrástípusokat kíván összegezni, más típusú adatokat kíván kinyerni, vagy eltérő kimeneti formátumokat kíván használni, akkor keresse fel igényével az IBM-et, vagy készítsen egyéni összegzőt.

## IBM Web Crawler Notes változata

Ez a szakasz írja le az IBM Web Crawler Notes változatának beállítását és használatát. Az IBM Web Crawler Notes változata Lotus Notes adatbázisokat ér el, és Notes dokumentumokból és csatolt fájlokból hoz létre összegzéseket. Az összegzéseket dokumentumonként vagy mellékletenként egy XML formátumú fájl tartalmazza, amelyben megtalálható az objektum és a teljes szöveg.

## Előfeltételek

Az IBM Web Crawler Notes változatának futtatása előtt teljesülniük kell a következő előfeltételeknek:

- Lotus Notes 5.0.5 vagy újabb változat
- Önkicsomagoló ZIP fájl mellékletek kezeléséhez PKZIP 2.50.

## Teszt bejárás végrehajtása

Válassza a **Start** → **Programok** → **Parancssor** menüpontot. Az új ablakban váltson át abba a könyvtárba, ahová az IBM Web Crawler Notes változatát terméket telepítette, majd váltson a notes-run alkönyvtárba. Például:

```
cd c:\<telepítési_könyvtár>\gcs\notes-run
```

A Notes bejárásokat két fájl vezérli:

- A források listája, amelyben megadhatja a bejárando Notes adatbázisokat. Ez tartalmazza a Notes kiszolgálóneveket, IP címeket, .NSF fájlneveket és néhány további beállítást. Egy forráslista tartalmazhatja például 34 Notes adatbázis nevét. A bejárásra kerülő adatbázisok az alábbiakban részletezett beállításokkal adhatók meg. A források listája egy XML fájl vagy egy Notes adatbázis (NSF fájl) lehet.
- A konfigurációs fájl határozza meg a forrás listát, a bejárando forrásokat, a feldolgozásra kerülő melléklet típusokat, a kimeneti formátumokat, és így tovább. A konfigurációs fájl mindig egy XML fájl.

Az IBM Web Crawler megfelelő telepítésének ellenőrzéséhez járja be a tesztadatbázist. Egy szövegszerkesztőben nyissa meg a testSources.xml fájlt, és győződjön meg róla, hogy a test.nsf adatbázis elérési útja és fájlneve megfelelő. A fájl az IBM Web Crawler telepítéskor a notes-run alkönyvtárba kerül. Készítsen biztonsági másolatot az eredeti fájlokról. **Tanács:** A szerkesztést figyelmesen végezze: A fájl hibája az IBM Web Crawler működésképtelenségét okozhatja. Mentse a változásokat.

A termékkel szállított `test.nsf` bejárásával ellenőrizze a telepítést. Írja be a `crawlNotes` `crawlTestXml` parancsot.

A `crawlNotes.bat` fájl a `crawlTestXml` konfigurációs fájlal indítja el az IBM Web Crawlert. A rendszer a konfigurációs fájl nevéhez automatikusan hozzáfűzi az `.xml` kiterjesztést. Az IBM Web Crawlernek jelentenie kell a bejárást, és el kell végeznie két melléklettel rendelkező dokumentum bejárását és összegzését.

Az IBM Web Crawler befejezése után ellenőrizze a summary könyvtárban található összegzéseket, illetve a konfigurációs fájlban megadott könyvtárban található bejárési naplófájlokat.

## Egyéni Notes bejárás beállítása

A sikeres tesztbejárás után meg lehet kezdeni a tényleges bejárásokat.

1. Hozzon létre egy saját adatbázis forráslistát. Adja hozzá a bejárando Notes adatbázisokat a forráslistához.

Ha a bejárando Notes adatbázisokat egy XML fájlban szeretné megadni, akkor kezdje a `testSources.xml` fájl szerkesztésével. Ha a bejárando Notes adatbázisokat Notes adatbázisban kívánja megadni, akkor a Notes segítségével nyissa meg, és módosítsa a `testSources.nsf` Notes adatbázist. A forrásfájlokban beállítható paramétereket a Forráslisták szerkesztése szakasz magyarázza el.

2. Adja meg a bejáró konfigurációját. Ehhez egy XML formátumú konfigurációs fájlt kell szerkeszteni.
  - Ha a források szerepelnek az XML fájlban, akkor kezdje a `crawlTestXml.xml` fájl módosításával, és állítsa be a `sourcesInXmlFile` attribútumot úgy, hogy a források fájlra mutasson.
  - Ha a források Notes adatbázisban vannak, kezdje a `crawlTestNsf.xml` fájlal, és állítsa be a `sourcesInNotesDB` paramétert úgy, hogy a források adatbázisra mutasson. A konfigurációs fájlokban megadható paraméterek magyarázatát "Az IBM Web Crawler konfigurációs fájl" oldalszám: 75 szakasz adja meg.

Ha a forráslista és a konfiguráció kész, akkor hívja meg az IBM Web Crawler szolgáltatást: `crawlNotes saját_konfiguráció`.

Ha a forráslista Notes adatbázisban van, akkor az IBM Web Crawler szolgáltatást a következőképpen indítsa el: `crawlNotes crawlTestNsf`.

Ha az IBM Web Crawler Notes bejárása befejeződött, akkor ellenőrizze a summary könyvtárban található összegzéseket, illetve a konfigurációs fájlban megadott könyvtárban található bejárési naplófájlokat.

## Forráslista paraméterei

A forráslista a bejárható Notes adatbázisok leírását tartalmazza. A `file.xml` formátumában megadott forráslisták egy `notesDataSources` elemet, és egy vagy több `oneDBinfo` elemet tartalmaznak. Az egyes `oneDBinfo` elemek a következőket tartalmazzák:

**id** Az adatbázis azonosítószáma. Erre a konfigurációs fájl tartomány paramétere hivatkozik.

### **serverName**

Az adatbázist tartalmazó kiszolgáló neve. Helyi adatbázis használatához adjon meg üres karaktersorozatot: "".

### **pathAndFileName**

Az adatbázis teljes elérési útja és fájlneve a kiszolgálón. Az elérési útnak és fájlnevnak `.nsf` kiterjesztéssel kell végződnie.

**viewName**

A bejárando adatbázis Notes nézetének neve.

**ipAddress**

Elhagyható. Ez adja meg a kiszolgáló IP címét. Megadása esetén nincs DNS kikeresés. Itt adhatja meg az IP címet abban az esetben is, ha a DNS nem tudja feloldani a célkiszolgáló nevét. Windows rendszereken az IP címet az nslookup kiszolgáló\_neve paranccsal határozhatja meg.

**dateLastCrawled**

Elhagyható. Az adatbázis legutóbbi bejárásának dátuma. Az értéket a bejáró automatikusan beállítja, kivéve, ha a konfigurációs fájlban az update Date Last Crawled paramétert kapcsolja.

**tries** Elhagyható. Itt adható meg, hogy sikertelen bejárás (időtúllépés) esetén hány kísérlet történjen az adatbázis ismételt bejárására.

**fieldSubstitutions**

A leképezések határozzák meg a Notes adatbázisok mezőneveinek cseréjét a kimeneti XML dokumentumokban. Legalább egy helyettesíthető elemet tartalmaz, ezek mindegyike két attribútummal rendelkezik:

- Original: Az a mezőnév, amelyet helyettesíteni fog a kimeneti XML dokumentumban, ha a mező létezik.
- Replace: Az új mező neve, amely az eredeti mező helyett fog megjelenni a kimeneti XML dokumentumban.

A Notes adatbázisokban tárolt forráslistákat a Notes ügyféllel lehet megvizsgálni és frissíteni. Indítsa el a Notes programot, majd válassza a **Fájl → Adatbázis → Megnyitás** menüpontot. Kattintson a **Tallózás** gombra a testSources.nsf adatbázis megkereséséhez az x:\<telepítési\_könyvtár>/gcs/notes-run könyvtárban. A termék részét képező tesztadatbázis is megvizsgálható és frissíthető a Notes ügyféllel. Indítsa el a Notes programot, majd válassza a **Fájl → Adatbázis → Megnyitás** menüpontot. Kattintson a **Tallózás** gombra a test.nsf adatbázis megkereséséhez az x:\<telepítési\_könyvtár>/gcs/notes-run könyvtárban.

**Konfigurációs fájl paraméterek**

Ez a rész a konfigurációs fájlban beállítható paramétereket írja le. Az alapértelmezett értékkel rendelkező paramétereket ki lehet hagyni.

A forráslistát egy sourcesInXml fájl vagy egy sourcesInNotesDB elem azonosítja. A forrás egy XML fájl vagy egy Notes adatbázis, amely tartalmazza a bejárando adatbázis vagy adatbázisok információit.

A runInfo elemek a futás irányításához tartalmazznak paramétereket. Ez azt jelenti, hogy ezek a paraméterek egy adott bejáró használata során az összes bejárt Notes adatbázisra vonatkoznak.

**rangeSpecify**

A bejárando adatbázisok azonosítói. Az azonosítók a forráslista azonosító mezőiben megadott számok. Megadható egyedi azonosítók vesszővel elválasztott listájaként, vagy kötőjellel jelölt tartományként, például: 1-4, 15, 25-31.

**SummaryDirectory**

A kimeneti összegzések főkönyvtárát adja meg. A rendszer az összegzéseket ennek a könyvtárnak az alkönyvtáraiba írja.

**MaxThreads**

A párhuzamos bejárési szálak maximális számát adja meg. Az egyes Notes adatbázisokat egyetlen szál járja be. Több adatbázis párhuzamosan kerül bejárásra.

**doIncrementalCrawl**

Az alapértelmezett beállítás a No. A Yes beállítás esetén csak a 'summarizeThisDateAndLater' óta készült/megváltozott Notes dokumentumok kerülnek feldolgozásra. Ha a 'summarizeThisDateAndLater' nincs megadva, akkor a bejáró egyes adatbázisok saját DateLastCrawled mezőjét használja az adatbázisok forráslistájában. Ha a doIncrementalCrawl=no beállítás van megadva, akkor a rendszer minden dokumentumot feldolgoz, függetlenül a dátumtól.

**summarizeThisDateAndLater**

A mező formátuma: HH/n/n/éééé óó:pp a IZ. Például: 01/01/2000 01:11 PM CET . Ha nincs megadott dátum és időpont, akkor az összes dokumentum összegzésre kerül vagy a forráslistában feljegyzett legutóbbi bejárás óta (doIncrementalCrawl = Yes), vagy időponttól függetlenül (doIncrementalCrawl = No).

**detachAttachments**

Az alapértelmezett beállítás a Yes. Ha a Yes beállítást adja meg, akkor a rendszer leválasztja és összegzi a csatolt fájlokat. A feldolgozandó csatolt fájl típusokat a konfigurációs fájl tartalmazza. A No beállítás esetén a mellékletek figyelmen kívül maradnak.

**attachmentFilenameFormat**

Az alapértelmezés az l (hosszú). A másik lehetséges beállítás az s (rövid). A hosszú fájlnevek a típust, a kiszolgálót, az adatbázis nevét és a Notes azonosítót adják meg. A rövid fájlnevekben a típus és a Notes azonosító található meg.

**processAttachmentsAfterwards**

Az alapértelmezett beállítás a No. A Yes beállítás használata esetén a melléklet fájlok nem kerülnek összegzésre a Notes adatbázis összegzés során. A bejáró minden csatolt fájlról egy bejegyzést ír a notesCrawl-attachments.bat fájlba, megadva a fájlok összegzésére szolgáló parancsokat. Ezután írhat egy olyan parancsfájlt, amely összegzi és törli a mellékleteket. A mellékletek későbbi feldolgozása általában jelentős lemezes tárterületet igényel.

**saveAttachmentFiles**

Az alapértelmezett beállítás a No. A Yes beállítás használatakor az eredeti melléklet fájlok nem kerülnek törlésre a feldolgozás végén. A beállítás csak akkor érvényes, ha a processAttachmentsAfterwards értéke No. Ha a processAttachmentsAfterwards beállítása Yes, akkor a mentett mellékletek tárolásához lemezterületre van szükség.

**MaximumNumberOfDetachingErrors**

Az alapértelmezett érték 10. A mellékletek kezelésekor bekövetkező hibák maximális száma, például lemezterület elfogyás a mellékletek mentésekor. A maximális számú hiba bekövetkezése után a bejárás megszakad.

**saveURLsToFile**

Az alapértelmezett beállítás a No. Ha meg van adva, akkor a Notes dokumentumban talált URL címek az alábbi nevű fájlba kerülnek:  
adatbázisnév (elérési út és .NSF nélkül) + ".html".

**updateDateLastCrawled**

Az alapértelmezett beállítás a Yes. A No beállítás használatakor a dateLastCrawled bejegyzés nem kerül frissítésre a forráslistában.

**tempDirectory**

Az alapértelmezett beállítás a c:\temp. A rendszer ezt a könyvtárat használja az ideiglenes fájlok írásához.

### logSummaryDirectory

Az alapértelmezett beállítás a log. Ebbe a könyvtárban kerülnek mentésre a naplófájlok.

### loggerPriority

Az alapértelmezett beállítás az info. Ez a beállítás határozza meg a naplózó prioritását. A lehetséges értékek: error (legmagasabb), warn és info (legalacsonyabb). Ha például a naplózó prioritása warn, akkor csak a warn és error prioritással rendelkező üzenetek kerülnek be a naplóba.

Az attachments elem a feldolgozandó csatolt fájl kiterjesztését meghatározó elemeket tartalmaz (például: .prz).

## Az IBM Web Crawler kizárása a kiszolgálókról

Biztonsági és teljesítmény szempontok miatt az EIP adminisztrátornak néhány kiszolgálót vagy oldalt ki kell zárnia a bejárasi folyamatból. Szükség lehet arra is, hogy korlátozni lehessen a bejárók tevékenységeit az üzemeltetett kiszolgálókon és oldalakon.

Az IBM Web Crawler *hozzáférési házirend* fájljal utasítható bizonyos kiszolgálók és oldalak elkerülésére. A fájl a *Robotok kizárásának szabványa* (lásd:

<http://info.webcrawler.com/mak/projects/robots/norobots.html>) szerint állítható össze.

- Az IBM Web Crawler a kiszolgáló minden bejárása során először lekéri a <http://kiszolgáló/robots.txt> hozzáférési házirend fájlt.
- A fájl az alábbi formájú sorokból áll:

*mező*:<elhagyható\_szőköz>*érték*<elhagyható\_szőköz>

Ha a *mező* User-Agent, az *érték* pedig IBM-WebCrawler vagy \*, akkor az alábbi Disallow sorok (a következő User-Agent sorig) adják meg az elkerülendő részleges címeket. Ez teljes vagy részleges elérési út lehet. Az értéknek megfelelő egyik cím sem kerül visszaadásra.

Például:

Disallow: /help

A /help.html és a /help/index.html URL-eket is letiltja.

Disallow: /help/

Letiltja a /help/index.html URL-t, de engedélyezi a /help.html oldalt.

Az üres értékek az összes cím visszaadását engedélyezik.

- A sorok között üres sorok is lehetnek.
- Megjegyzések a # karakterrel szűrhetők be. A sor további részeit a rendszer megjegyzésnek tekinti.

Néhány példa:

- Az alábbi /robots.txt megadja, hogy a kiszolgálót az összes robotnak el kell kerülnie.  
# mindenki kitiltva

User-agent: \*  
Disallow: /

- Az alábbi /robots.txt megadja, hogy csak az IBM Web Crawler végezhet bejárást, erre viszont nem vonatkoznak korlátozások.

```
# csak az IBM engedélyezett
```

```
User-agent: *  
Disallow: /  
User-agent: IBM-WebCrawler  
Disallow: # semmi sem tiltott
```

- Az alábbi /robots.txt megadja, hogy minden robotnak el kell kerülnie a **temp**, **development** és **testing** htmldocs alkönyvtárakat, továbbá az IBM-WebCrawler nevű robotnál kivételt képez a development és a testing könyvtár (engedélyezett), valamint az xyz és wxyz robotoknak a teljes kiszolgálót el kell kerülniük.

```
# életszerűbb példa
```

```
User-agent: *  
Disallow: /htmldocs/temp  
Disallow: /htmldocs/development  
Disallow: /htmldocs/testing
```

```
User-agent: xyz  
User-agent: wxyz  
Disallow: /
```

```
User-agent: IBM-WebCrawler  
Disallow: /htmldocs/temp
```





---

## 6. fejezet Bevezetés a munkafolyamatokba

Az EIP munkafolyamatok segítségével felügyelhetővé válik az üzleti folyamatok előrehaladása és teljesítménye. Amikor a felhasználók az egyesített keresések eredményeit használják, igen gyakran kell döntéseket hozniuk a végrehajtandó tevékenységekről. Az EIP munkafolyamat funkciók ezt a folyamatot azzal segítik, hogy segítenek előzetesen meghatározni a felhasználók által elvégzendő feladat jellegét.

A munkafolyamatokat automatizálhatja. Ehhez létre kell hozni azokat a profilokat és szabályokat, amelyek alapján a munkafolyamat összetevői együttműködnek. A felhasználói hozzáférések és jogosultságok a jogosultságkészletek és hozzáférés felügyeleti listák segítségével felügyelhetők.

---

### Munkafolyamatok magyarázata

A legtöbb üzleti művelet egymástól függő folyamatként írható le. A munka egyik dolgozótól a másikig, illetve az egyik osztálytól a másikhoz kerül. Az egyszerűbb folyamatok lehet, hogy csak néhány lépésből állnak, a bonyolultabbak viszont különféle osztályok több dolgozóját is érinthetik.

A munkafolyamat szolgáltatás lehetővé teszi a feladatok továbbítását a folyamatokban, és döntések meghozatalát a folyamat során. Az XYZ Biztosító például sok keresetet kap levélben. Az ellenőrzési folyamat során a biztosítási kárbecslőknek különféle dokumentumokat, például fotókat, értékeléseket és szakértői jelentéseket kell összegyűjteniük. A végső elfogadás előtt az alkalmazottak minden nap órákat töltenek az információk megnyitásával, rendezésével, kartotékozásával és áttekintésével, valamint az odaillő dokumentumok begyűjtésével.

Az információk fogadása és ellenőrzése során ez a feladatcsoport az egyik alkalmazottól a másikhoz kerül. A keresettel kapcsolatos teendők befejeződéséig lehet, hogy több osztály alkalmazottaihoz is eljut az anyag.

---

### Munkafolyamatok alkalmazása

A fenti példában szereplő XYZ Biztosítóhoz hasonlóan a dokumentumokat alkalmazó cégek általában az alábbi feladatok valamilyen részét hajtják végre:

- Dokumentumok iktatása a későbbi visszakereséshez.
- Dokumentumok, űrlapok, jelentések és információk összegyűjtése különféle forrásokból, majd a dokumentumok eljuttatása megfelelő helyre a feldolgozáshoz.
- Beérkező levelek és a feldolgozás alatt álló dokumentumok illesztése.

A *munkafolyamat* a munka folyamatát jelenti. Ez írja le az egy vagy több dokumentumból vagy tartalomcsoportból álló csoportokon végrehajtható tevékenységeket, valamint a dokumentumcsoport által a folyamat során bejárt útvonalat. A munkafolyamat a végrehajtásnak megfelelően tükrözi a feladatokat, pontosan tisztázott hatáskörökkel és határokkal. Meghatározza a feladatok és tevékenységek sorrendjét, valamint az ezek között fennálló kapcsolatokat és viszonyokat. A munkafolyamat adja meg azokat a feltételeket, amelyek meghatározzák a munka során hozott döntéseket. A munkafolyamatok létrehozásával kapcsolatban nézze meg a *Munkaállomás alkalmazás programozói kézikönyv* című kiadványt. A munkafolyamat szolgáltatással használható ügyfelekre vonatkozóan nézze meg az *e-Ügyfél telepítése, beállítása és kezelése* című kiadványt.

## Felhasználói azonosítók és csoportok szinkronizálása

Ez a szakasz írja le a felhasználói azonosítók szinkronizálását az EIP, a Content Manager és az MQSeries Workflow között.

Ha olyan Content Manager vagy Enterprise Information Portal rendszer felhasználóit kezeli, amelynek az MQSeries Workflow is része, akkor ez az MQSeries Workflow felhasználók kezelését is jelenti. Ennek megfelelően ha létrehoz, módosít vagy töröl egy felhasználói azonosítót vagy csoportot a Content Manager vagy az Enterprise Information Portal rendszerben, akkor ezt el kell végezni az MQSeries Workflow kiszolgálón is.

Mivel a Content Manager és az Enterprise Information Portal megosztja a felhasználói azonosítókat és csoportokat, ha olyankor hoz létre egy felhasználói azonosítót vagy csoportot a rendszeradminisztrációs ügyfélben a munkafolyamatok engedélyezése mellett, amikor az MQSeries Workflow kiszolgáló nem fut, akkor hibaüzenet jelzi, hogy a felhasználót vagy csoportot nem lehetett létrehozni az MQSeries Workflow kiszolgálón. Bár a felhasználói azonosító létezik a Content Manager vagy az Enterprise Information Portal rendszeren, az MQSeries Workflow kiszolgálón nem létezik.

A tartalomkezelési rendszer és az MQSeries Workflow felhasználói azonosítónak és csoportjainak szinkronizálásához le kell futtatni a Workflow felhasználó szinkronizálási segédprogramot. Ha az MQSeries Workflow kiszolgáló telepítését saját kezűleg végezte, akkor győződjön meg róla, hogy az MQSeries Workflow kiszolgáló elindult. Nyissa meg a Szolgáltatásokat, és ellenőrizze, hogy az állapotánál az MQSeries Workflow *X.X* - FMC bejegyzés látható (ahol az *X.X* a telepített kiszolgáló verziószáma). Ha az MQSeries Workflow nem fut, akkor indítsa el a szolgáltatást, vagy ha a telepítést csendes módban végezte, akkor lépjen be a WFInstall könyvtárba, és futtassa a CMBWFStart.bat fájlt. A szinkronizálási segédprogram futtatásához tegye a következőket:

1. Váltson be az Enterprise Information Portal telepítési könyvtárba. Az alapértelmezett könyvtár a C:\CMBROOT.
2. Írja be az EIPUser2WF.bat parancsot.

**Fontos:** Távoli munkafolyamat kiszolgáló esetén az EIPUser2WF.bat indítása előtt el kell indítani az RMI kiszolgáló parancsfájlját (cmbsvregist81.bat vagy cmbsvregist81.sh) is. Emellett az EIPUser2WF.bat két INI fájl (cmbsvcs.ini és cmbsvclient.ini) információi segítségével határozza meg a helyi vagy távoli kiszolgáló helyét. Győződjön meg róla, hogy a cmbsvcs.ini fájlban a LOCAL érték szerepel helyi kiszolgáló, illetve REMOTE távoli kiszolgáló esetén. A cmbsvclient.ini fájl tartalmazza a kiszolgáló helyét.

3. Adja meg az EIP adatbázis nevét, a felhasználói azonosítót, a jelszót és a sémát.

A szükséges információk beírása után a szinkronizálási segédprogram átmásolja a felhasználókat a Content Manager vagy Enterprise Information Portal kiszolgálóról az MQSeries Workflow kiszolgálóra. A segédprogram befejeződése után nem kap arra utaló hibaüzenetet, hogy egy felhasználó vagy csoport ismeretlen az MQSeries Workflow kiszolgálón.

Ha töröl egy felhasználói azonosítót vagy csoportot a Content Manager vagy Enterprise Information Portal kiszolgálóról, akkor a felhasználói azonosítót az MQSeries Workflow kiszolgálóról is törölni kell. Ha egy felhasználói azonosító vagy csoport nem létezik az MQSeries Workflow kiszolgálón, akkor azt nem törölheti a Content Manager vagy Enterprise Information Portal kiszolgálóról. Tegyük fel például, hogy létrehoz egy felhasználói azonosítót a Content Manager vagy Enterprise Information Portal rendszerben az EIP munkafolyamatok letiltása mellett. Ezután engedélyezi a munkafolyamat kiszolgáló beállítását, majd megpróbálja törölni a felhasználói azonosítót vagy csoportot. Mivel a felhasználói azonosító vagy csoport nem létezik az MQSeries Workflow kiszolgálón, hibaüzenet jelzi,

hogy a felhasználói azonosító nem létezik az MQSeries Workflow kiszolgálón. A hiba kijavításához le kell futtatni az EIPUser2WF segédprogramot a felhasználói azonosítók és csoportok szinkronizálásához, majd törölje a felhasználói azonosítót vagy csoportot.

## EIP kiszolgáló újratelepítése a munkafolyamatok engedélyezése esetén

Ha engedélyezte a munkafolyamatokat az EIP rendszerben, és újra kívánja telepíteni az EIP kiszolgálót, akkor az MQSeries Workflow kiszolgálóról el kell távolítani az EIP összes munkafolyamatokkal kapcsolatos adatát.

Az EIP kiszolgáló eltávolítása előtt az alábbi lépések végrehajtásával távolíthatja el az EIP munkafolyamatok adatait az MQSeries Workflow kiszolgálóról. **Fontos:** A lépéseket a megadott sorrendben kell végrehajtani.

1. Szüntesse meg az összes folyamatpéldányt az e-Ügyfélből vagy a munkafolyamat API-k segítségével. A folyamatpéldányok megszüntetése az összes feladat törlésével jár.
2. Törölje a feladatlistákat és folyamatsablonokat az EIP rendszeradminisztrációs ügyfélből vagy a rendszeradminisztrációs API-k segítségével.
3. Törölje az összes felhasználót és csoportot, még amikor a munkafolyamatok engedélyezettek, mivel ez biztosítja, hogy a felhasználói azonosítók és csoportok törlése az MQSeries Workflow kiszolgálón is megtörténjen.
4. Konfigurálja újra az EIP MQSeries Workflow kiszolgálóját.
  - a. Helyezze be az EIP telepítő CD-t. Egy parancssorban lépjen be a WflInstall könyvtárba.
  - b. A parancssorban írja be a következő parancsot:

```
fmcibie -iCMBWFAdmin.fdl -uadmin -ppassword -o -f
```

**Fontos:** Ha az EIP adatbázist még az előtt távolítja el, hogy törölte volna az EIP és MQSeries Workflow felhasználói azonosítókat és csoportokat, akkor a törölt EIP adatbázisban korábban szerepelt felhasználói azonosítók és csoportok ismételt létrehozására tett kísérlet meghiúsul. A megjelenő hibaüzenet: DGL2616A: A(z) XXX felhasználó nem adható hozzá. DGL2485A: A munkafolyamat felhasználó már létezik. A probléma elhárításához tegye a következőket:

1. Tiltsa le az EIP munkafolyamat szolgáltatást az EIP rendszeradminisztrációs ügyfélben.
2. Jelentkezzen ki az EIP rendszeradminisztrációs ügyfélből, majd jelentkezzen be ismét. A munkafolyamatok letiltása és a kilépés után ismét létrehozhatja azokat a felhasználói azonosítókat és jelszavakat, amelyek még megtalálhatók az MQSeries Workflow kiszolgálón.
3. Az MQSeries Workflow kiszolgálón már létező felhasználói azonosítók és csoportok létrehozása után engedélyezze az EIP munkafolyamat szolgáltatásokat.

## Felhasználói azonosítók és csoportok frissítése az MQSeries Workflow és az EIP adatbázis között

A felhasználói azonosítók és csoportok szinkronizálásához tegye a következőket az összes olyan felhasználónál és csoportnál, amely létezik az MQSeries Workflow kiszolgálón, de nem létezik az EIP adatbázisban:

1. Hozzon létre egy fájlt, és írja bele a következő szöveget:
  - CODEPAGE 1250
  - FM\_RELEASE V3R3 2
  - DELETE PERSON 'User1'
  - DELETE PERSON 'User2'

- DELETE PERSON 'User3'

ahol a User1, User2 és User3 a törölni kívánt felhasználói azonosítók. Tetszőleges számú felhasználó szerepelhet a fájlban.

2. Mentse a fájlt, és futtassa a következő parancsot egy parancssorban:

```
fmcibie -u admin -ppassword -i DeletePersons.fdl -f -o
```

A DeletePersons.fdl az előző lépésben létrehozott fájl neve.

Tegye a következőket az összes olyan felhasználónál, amely létezik az MQSeries Workflow kiszolgálón, de nem létezik az EIP adatbázisban:

1. Hozzon létre egy szövegfájlt, és írja bele a következőket:

- CODEPAGE 1250
- FM\_RELEASE V3R3 2
- DELETE ROLE 'Group1'
- DELETE ROLE 'Group2'
- DELETE ROLE 'Group3'

ahol a Group1, Group2 és Group3 a törölni kívánt csoportok. Tetszőleges számú csoport szerepelhet a fájlban.

2. Mentse a fájlt, és írja be a következő parancsot egy parancssorba:

```
fmcibie -u admin -p password -i DeleteGroups.fdl -f -o
```

A DeleteGroups.fdl az előző lépésben létrehozott fájl neve.

Amikor töröl egy felhasználói azonosítót az EIP rendszerből, és a munkafolyamat szolgáltatás engedélyezett, akkor a következő hibaüzenet jelenhet meg:

A felhasználó törlése meghiúsult [RC=12]

Ha egy felhasználó törlési művelet meghiúsul, akkor a rendszer előállít egy temp.log fájlt, amelyben megtalálja a hiba részleteit. A temp.log fájl helye az x:\CMBROOT könyvtár.

**Tipp:** Ha azt a felhasználói azonosítót próbálja meg törölni, amellyel bejelentkezett az MQSeries Workflow ügyfélbe, akkor a törlési művelet meghiúsul.

---

## Munkafolyamat tervezése

A munkafolyamat meghatározása előtt elemezni kell a vállalatban végzett feladatokat, a végrehajtás helyét és módját, valamint a feladatokat végrehajtó személyeket. Ezt a tervezési lépést az adminisztrátor vagy üzleti elemző végzi el.

Mi a végtermék? A végtermék lehet a teljes vállalat munkájának eredménye, a vállalat egy osztályán végzett munka eredménye, vagy különféle osztályokon tevékenykedő dolgozók munkájának eredménye. A már említett biztosítótársaságnál a végtermék az a levél, amit a kötvénytulajdonos kap a kereset elfogadásáról vagy elutasításáról.

Elemezni kell a végtermék előállításához feldolgozott információkat, meg kell határozni a végrehajtandó tevékenységeket, ezek végrehajtási helyét, és az információáramlás módját a munkafolyamaton keresztül.

## Feldolgozott információk

Gondolja át a vállalat felhasználói által feldolgozásra kerülő információkat. Milyen jellegű bemenet felel meg a végterméknek? Mik azok az adott dokumentumok, amelyeket fel kell dolgozni?

A *feladatok* nem dokumentumok. A *feladatok* dokumentumokra vonatkozó hivatkozásokat tartalmaznak, emellett további információkkal is rendelkezhetnek a dokumentumokról, például a dokumentum állapotával vagy létrehozási idejével. *Feladat* a tartalomkiszolgálók tetszőleges tartalma (dokumentum vagy objektum) lehet. Az XYZ biztosítótársaság például először a kereset űrlapot kapja meg, és a későbbiek során futnak be a kiegészítő dokumentumok, például a fotók, becslések vagy szakértői jelentések.

## Információkezelés módja

Ki tudja a legjobban kezelni a folyamat egyes lépéseit? Egy adminisztrációs asszisztens például ellenőrizheti, hogy egy kereseti űrlap teljes-e, majd iktathatja az igénylést, amíg egy adott dokumentum be nem érkezik a kötvénytulajdonostól. A dokumentum megérkezésekor a biztosítási kárbecslő lehet felelős a dokumentum és a kereset párosításáért, és a dokumentum elfogadásáért.

A *feladatlisták* úgy tekinthetők, mint a dolgozók számára létrehozott teendőlisták. A kereseti űrlapok egy *feladatlistába* csoportosíthatók, amelyeket a megadott biztosítási kárbecslők érhetnek el. A feladatlista a feladatok megfelelően szűrt megjelenése. Az alkalmazottak a feladatlistának csak azokat az elemeit látják, amelyekre jogosultak.

A folyamaton belül feladatlisták meghatározhatók a kereset munkafolyamatának minden egyes részéhez, például a fényképek, becslések és szakértői jelentések összegyűjtéséhez. A feladatlista tartalmazhat más folyamatokból származó feladatokat is. Egy biztosítási kárbecslő feladatlistája például tartalmazhatja egy kereset felbecsléseit, egy másik kereset fotóit és egy harmadik szakértői jelentéseit. A kárbecslő által végrehajtott tevékenység a feladatlista minden feladatánál eltérő lehet. A felbecslés áttekintése után például elfogadhatja az első keresetet. Mielőtt a második kereset fotóival foglalkozni lehetne, lehet hogy további információkat kér ezzel kapcsolatban. A harmadik kereset szakértői jelentését pedig továbbíthatja egy másik alkalmazottnak.

## Végrehajtandó tevékenységek

A feladat tartalmával kapcsolatos teendőket át kell gondolni. Egy biztosítási kárbecslő például elfogadhat egy keresetet, vagy formai hiányosságok miatt visszautasíthatja. A felhasználók által a feladatokon végrehajtható tevékenységeket a *tevékenységlisták* határozzák meg.

Attól függően például, hogy a kereset megfelel-e a kezdeti követelményeknek, a biztosítási kárbecslők egyik lehetőségként továbbengedhetik a kereseti űrlapot a szokásos folyamaton, másik lehetőségként pedig visszautasíthatják azt.

## Hogyan haladnak végig az információk a folyamaton?

Gondolja át, hogyan haladnak végig az információk és tevékenységek a folyamaton. Például mikor kerül sor a kereseti űrlap kezdeti áttekintésére? Milyen további dokumentumok szükségesek ahhoz, hogy a kérelem továbbhaladhasson a folyamat következő lépéséhez? Milyen feltételek határozzák meg, hogy egy kereset elfogadásra vagy visszautasításra kerül? Ez az információáramlás a munkafolyamat alapja.

A munkafolyamat a feladatot a folyamaton végigvezető útvonalakból áll. Honnan származnak a bemeneti információk? A munkafolyamatnak el kell kezdődnie valahol. Az XYZ Biztosítótársaság esetén a munkafolyamatot a kötvénytulajdonos által küldött kereseti űrlap indítja el.

Ha minden dokumentum beérkezett, akkor a feladat folytathatja útját a folyamat végéig — például a kereset elfogadásáig.

## A részek összekapcsolódása

Miután elemezte a feldolgozásra kerülő információkat, megállapította a végrehajtandó tevékenységeket, és meghatározta információáramlás módját, készen áll a munkafolyamat diagramjának létrehozására, amely tulajdonképpen a munkafolyamat grafikus ábrázolása. A folyamatábra létrehozásához az EIP grafikus munkafolyamat tervezője használható.

A folyamatábra bemutatja a munka útvonalát a folyamat különféle tevékenységei között, feljegyezve a tevékenységgel járó feladatokat is. Leírja a munkafolyamat áramlását, főbb elemeit és a lényeges pontjait.

A folyamatábra minden szimbóluma egy olyan pontot jelképez, amelyben egy feladat kerül elvégzésre. A biztosítási kereseteket át kell tekinteni, be kell gyűjteni hozzá a kiegészítő dokumentumokat, és az adott feltételek alapján el kell utasítani, vagy el kell fogadni a keresetet. A folyamat tervezőben alkalmazott folyamat szimbólumokról bővebben lásd: "Folyamatok létrehozása" oldalszám: 100.

## Enterprise Information Portal munkafolyamat összetevők használata

Ez a szakasz írja le a munkafolyamat szolgáltatás összetevőit. Minden összetevő az adminisztrációs ügyfélből érhető el. **Tipp:** Az EIP 8. változatában lévő munkafolyamat szolgáltatások több változáson is keresztülmentek, megváltozott például a 7.1 változat tárolója a Content Manager v8 architektúra kiterjesztései miatt.


## Munkafolyamat tervező használata

A folyamat tervező segítségével grafikus formában határozhatja meg és állíthatja össze a munkacsoportok, osztályok vagy a teljes vállalat munkafolyamatait. **Korlátozás:** Az EIP áttérési folyamat a 7.1 adatbázisokból a felhasználókat veszi át. Az EIP 8.2 változata nem biztosít automatikus áttérést a munkafolyamatokkal kapcsolatos adatokhoz. A 7.1 változatban használt munkafolyamat grafikonokat újra kell rajzolni az EIP 8.2 folyamat tervezőjével, és ismét ki kell alakítani a 7.1 változatban használt folyamatokat.

A munkafolyamat tervező használata előtt létre kell hozni a jogosultságkészleteket, hozzáférés felügyeleti listákat, felhasználókat, felhasználói csoportokat, tevékenységeket, tevékenységlistákat és feladatlistákat. Amikor az adminisztrációs ügyfélben kialakítja a munkafolyamatot, akkor a teljes folyamatra vonatkozóan beállítható egy alapértelmezett tevékenységlista. A folyamat minden egyes csomópontjában egyénileg is beállíthatók a különféle tevékenységlisták. Az említett feladatokról a "Tevékenységlisták meghatározása" oldalszám: 100 és a "Feladatlisták meghatározása" oldalszám: 99 szakasz, illetve az online súgó tartalmaz további információkat.



Bár a munkafolyamat tervező felhasználható munkafolyamatok létrehozására, nem használható a munkafolyamatok futtatására. Az ügyfél használatakor a felhasználók a feladatlistákat és a feladatokat látják illetve kezelik. Saját Enterprise Information Portal munkafolyamat ügyfelek programozásával kapcsolatban olvassa el a *Munkaállomás alkalmazás programozói kézikönyv* és az online API kézikönyv kiadványokat. A minden munkafolyamatban előforduló három ikont a 7. táblázat: írja le. Emellett más munkafolyamat ikonok is vannak, például részmunkafolyamatok, események, adatgyűjtések és felhasználói kilépési pontok. Az eszköztár ikonjait részletesen az EIP online súgója írja le.

7. táblázat: Alapvető munkafolyamat ikonok

Ikon	Leírás
	A kezdő csomóponttal kezdődik a munkafolyamat. Minden munkafolyamatnak rendelkeznie kell egy, de csakis egy kezdő csomóponttal.



7. táblázat: Alapvető munkafolyamat ikonok (Folytatás)

Ikon	Leírás
	A befejező csomópont zárja le a munkafolyamatot. Minden új munkafolyamat ábra tartalmaz egy befejező csomópontot. A folyamatok létrehozásakor automatikusan létrejön egy befejező csomópont is. A befejező csomópont a rajzolási felület tetszőleges pontjára helyezhető. Minden munkafolyamatnak rendelkeznie kell egy, de csakis egy befejező csomóponttal.
	A munkavégzési csomópont társít egy feladatlistát és tevékenységlistát a munkafolyamat egy adott pontjával. A munkavégzési csomópont a munkafolyamatnak olyan pontját jelzi, ahol munkavégzés történik. Minden csomópontnak, még a kezdő és befejező csomópontnak is szüksége van egy tevékenységkészletre és egy személyre, aki ezeket a tevékenységeket elvégzi.

## Munkafolyamat szolgáltatások használata

Az Enterprise Information Portal által biztosított munkafolyamat szolgáltatások a munkafolyamatokra vonatkozó információkat kezelik. A munkafolyamat tervezőben létrehozott munkafolyamat és tevékenységlista meghatározások az Enterprise Information Portal adminisztrációs adatbázisban és az IBM MQSeries Workflow adatbázisban találhatók.

Miután a rendszeradminisztrátor létrehozott egy feladatlistát, a hozzá tartozó információk állandó tárolásra kerülnek az adminisztrációs adatbázisban. A rendszeradminisztrátor a feladatlisták frissítését, törlését és hozzáadását az EIP segítségével végzi. Amikor a rendszeradminisztrátor kiiktat egy folyamatot, akkor a folyamat zárolásra kerül az Enterprise Information Portal adatbázisban, és a rendszer kiiktatottként jelöli meg, így akadályozva meg, hogy bárki más módosíthassa, amíg az adott felhasználó be nem fejezte vele a munkát.

## Feladatlisták meghatározása

A feladatlista tulajdonképpen az elvégezhető feladatok egy szűrése. A feladatlista egy adott felhasználóhoz vagy felhasználói csoporthoz tartozó feladatok megfelelően szűrt listája. Amikor a felhasználók bejelentkeznek az Enterprise Information Portal rendszerbe, akkor megjelenik a nekik kiosztott feladatok listája. A feladatlistákat az Enterprise Information Portal adminisztrációs ügyfélben lehet meghatározni.

A feladatlisták meghatározásába beletartoznak azok a szabályok, amelyek a feladatok megjelenítésére, állapotára és biztonságára vonatkoznak. Az egyes feladatlistákra vonatkozó szabályokat a feladatlista létrehozásakor adja meg. A feladatlista hozzáféréseinek felügyeletére hozzáférés felügyeleti lista hozható létre. A feladatlisták meghatározásáról további információkat az online súgó tartalmaz. A feladatlisták meghatározásába az alábbiak tartoznak:

### Hozzáférés felügyeleti lista

A hozzáférés felügyeleti listák legalább egy egyedi felhasználói azonosító vagy felhasználói csoport azonosítójából, illetve a hozzájuk tartozó jogosultságok listájából állnak. A jogosultságkészletek határozzák meg a felhasználó jogosultságait a feladaton végzett bizonyos tevékenységekre vonatkozóan. A hozzáférés felügyeleti listák segítségével korlátozható a felhasználók hozzáférése a feladatlista elemeihez.

### Feladatlisták szűrése és rendezése

Azok a feltételek, amelyek alapján a felhasználói feladatlisták szűrése és rendezése történik.

### Feladatlista elemeinek maximális száma

A feladatlisták elemeinek maximális száma.

---

## Tevékenységlisták meghatározása

A tevékenységlisták azon tevékenységek összefoglaló felsorolásai, amelyeket a felhasználók végrehajthatnak a munkafolyamat feladatain.

A tevékenységek és tevékenységlisták meghatározásáról további részleteket az adminisztrációs ügyfél online súgójában talál.

---

## Folyamatok létrehozása

A tevékenységek, tevékenységlisták és feladatlisták meghatározása után a folyamat tervezővel létrehozható a folyamat modellje. A tevékenységek és tevékenységlisták meghatározásáról további részleteket az adminisztrációs ügyfél online súgójában talál. A Folyamat tervező vizuális jelzésekkel segíti a folyamatok létrehozását.

---

## Munkafolyamat tervező engedélyezése

Ebben a lépésben kerül sor a munkafolyamat szolgáltatások elindítására az adminisztrációs adatbázisban. **Korlátozás:** A munkafolyamat szolgáltatásokhoz kijelölt adatbázisnak ugyanazon a kiszolgálón kell lennie, mint az MQSeries terméknek, továbbá az MQSeries szolgáltatásokat el kell indítani.

EIP munkafolyamatok engedélyezése és egy munkafolyamat meghatározás létrehozása:

1. Jelentkezzen be az adminisztrációs ügyfélbe.
2. Ha több adminisztrációs adatbázissal rendelkezik, akkor kattintson annak az ikonjára, amelyben engedélyezni kívánja a munkafolyamat szolgáltatásokat.
3. Válassza az Eszközök → Szolgáltatások menüpontot. Kattintson a Munkafolyamatok engedélyezése beállításra.
4. Jelentkezzen ki az ügyfélből, majd jelentkezzen be ismét. Ha több adatbázissal rendelkezik, akkor kattintson annak az ikonjára, amelyben engedélyezte a munkafolyamat szolgáltatásokat. Megjelenik a **Munkafolyamatok** mappa ikonja.
5. Az Enterprise Information Portal adminisztráció főablakának bal oldalán kattintson duplán a **Munkafolyamatok** mappára.
6. Kattintson a jobb egérgombbal a **Folyamatmeghatározások** elemre, majd válassza az előugró menü **Új** menüpontját egy új folyamat létrehozásához.

**Követelmény:** Folyamatok meghatározása előtt létre kell hozni legalább egy hozzáférés felügyeleti listát, egy tevékenységet és egy tevékenységlistát.

---

## MQSeries Workflow kiszolgáló indítása

Az MQSeries Workflow kiszolgáló indításához írja be a `cmbwfstart` parancsot egy parancssorba. Megjelenik az MQSeries Workflow kiszolgáló két ablaka. A kiszolgáló futásának folytatásához hagyja nyitva az ablakokat.

Ha a munkafolyamat szolgáltatást az Enterprise Information Portal kezdeti telepítése után telepíti, akkor az Enterprise Information Portal rendszert be kell állítani a munkafolyamat szolgáltatás használatára. A konfigurációt akkor is módosítani kell, amikor a munkafolyamat szolgáltatást nem arra a munkaállomásra telepíti, amelyre az adminisztrációs ügyfél is telepítve van.

1. Az Adminisztráció ablakban kattintson az **Eszközök** elemre.



2. Kattintson a menü **Szolgáltatások** pontjára.
3. Válassza ki a **Munkafolyamat** jelölőnégyzetet.
4. A konfiguráció befejezése után lépjen ki az Enterprise Information Portal adminisztrációs ügyfélből, majd jelentkezzen be ismét a munkafolyamat szolgáltatás inicializálásához. Miután bejelentkezett az Enterprise Information Portal adminisztrációs ügyfélbe, a bal oldali panelen megjelenik a **Munkafolyamat meghatározások** ikon.

**Tipp:** Az adminisztrátorok csak akkor látják a **Munkafolyamat meghatározások** ikont, ha jogosultak a munkafolyamat szolgáltatás felügyeletére. Ha korlátozni kívánja a munkafolyamat szolgáltatás elérését, akkor nézze meg az egyes tartalomkiszolgálók megfelelő rendszeradminisztrációs kézikönyvét. Az online súgó további információkat is tartalmaz az adminisztrátorok feljogosításáról a munkafolyamat szolgáltatás felügyeletére.

Az ügyfél elkészíthető egy egyéni alkalmazásból az EIP kapcsolati eszközkészlet és példák alapján, illetve használható az EIP minta ügyfele is.



---

## 7. fejezet IBM Web Crawler mintafájlok

Ez a szakasz két konfigurációs példát mutat be. A config-sample2.xml mintafájl példákat mutat be a <gcs-config> konfigurációs paraméterekre. A naplóelemzési példa egy jelentést mutat be egy befejeződött bejárás információiról.

---

### config-sample2.xml példa

Az alábbi példa a gcs-config fájl használatát mutatja be.

```
<!DOCTYPE gcs-config SYSTEM "config.dtd">
<gcs-config>
  <!-- Globális beállítások -->
  <globals max-urls="1000000"
    num-crawlers="30"
    num-summarizers="8"
    summaries-dir="summaries"
    log-file="log/LOG.txt"
    temp-dir="temp"
    log-priority="warn"
    text-monitor="60"
    graph-monitor="2"
    connect-timeout="120"
    read-timeout="100">

    <!-- Naplózási beállítások -->
    <logger-config category="gcs.summaries.list.resource"
priority="info" log-file="log/resources.txt"/>
    <logger-config category="gcs.summaries.list.host" priority="info"
log-file="log/hosts.txt"/>
    <logger-config category="gcs.url.skipped" priority="info"
log-file="log/skipped_urls.txt"/>
    <logger-config category="gcs.url.fromto" priority="info"
log-layout="%m\n" log-file="log/fromto.txt"/>
    <logger-config category="gcs.http" priority="info"
log-file="log/http.txt"/>
    <logger-config category="gcs.http.connect" priority="info"
log-file="log/connecterrs.txt"/>

    <!-- Adatbázis meghatározása -->
    <urlpool-config urlcontainer-class="DB2URLContainer"
urlcollection- class="DB2URLCollection">
      <urlpool-param name="dbname" value="gcs"/>
      <urlpool-param name="user" value="xxxxxx"/>
      <urlpool-param name="password" value="xxxxxx"/>
      <urlpool-param name="cachesize" value="1000"/>
      <urlpool-param name="driver"
value="COM.ibm.db2.jdbc.app.DB2Driver"/>
    </urlpool-config -->

    <!-- SOCKS proxy beállítása -->
    <system-properties>
      <property name="socksProxySet" value="true"/>
      <property name="socksProxyHost" value="socks2.server.ibm.com"/>
      <property name="socksProxyPort" value="1080"/>
    </system-properties -->

  </globals>

  <group-list>
    <group name="ibm">
```

```

<crawler-config recursion-depth="-1">
  <seed-list>
    <!-- A bejárás kezdeti URL címei: -->
    <seed url="http://gcs.stl.ibm.com/gcs/testurl.html"/>
    <seed url="http://gcs.stl.ibm.com/gcs/stl.html"/>
    <seed url="http://gcs.stl.ibm.com/gcs/ibm.html"/>
  </seed-list>

  <content-type-pattern-list>
    <!-- Az alábbi mintáknak nem megfelelő URL fájlkiterjesztések
    nem kerülnek bejárásra: -->
    <url-name-pattern name="htm"/>
    <url-name-pattern name="pdf"/>
    <url-name-pattern name="gif"/>
    <url-name-pattern name="zip"/>
    <url-name-pattern name="txt"/>
  </content-type-pattern-list>

  <include-pattern-list>
    <!-- Az alábbi mintáknak nem megfelelő URL címek nem
    kerülnek bejárásra -->
    <url-obj-pattern host="*.ibm.com"/>
    <!-- sbo - url-obj-pattern query="*OpenDocument*" / -->
    <!-- sbo - url-obj-pattern query="*OpenView*" / -->
    <!-- url-obj-pattern query="*OpenDocument =>
OpenDocument&ExpandAll*" / -->
    <!-- url-obj-pattern query="*OpenView =>
OpenView&ExpandAll&Count=999999*" / -->
  </include-pattern-list>

  <exclude-pattern-list>
    <!-- Az alábbi mintáknak megfelelő URL címek nem kerülnek bejárásra -->
    <!-- Intranet általános részeinek kihagyása -->
    <url-obj-pattern file="*news*" />
    <url-obj-pattern file="*search*" />
    <url-obj-pattern file="*/afs/*" />
    <url-obj-pattern file="*/.../*" />
    <url-obj-pattern file="*bluepages*" />
    <!-- Személyes honlapok kihagyása -->
    <url-obj-pattern file="*/~*" />
    <!-- SOCKS kihagyása: Ezt egy URL sem adhatja meg közvetlenül -->
    <url-regex-pattern match=".*:1080/*" />
    <!-- Átjárók kihagyása: Egyszerű halandók számára ajánlott -->
    <url-regex-pattern match=".*[?\\=\\;\\%&quot;&amp;].*" />
    <!-- Ellenkező esetekben az itt beállított átjárók bejárása -->
    <!-- Domino kihagyása -->
    <url-obj-pattern file="*.nsf*" />
    <!-- Ellenkező esetben a Domino bejárása:
    csak az OpenDocument engedélyezése -->
    <url-obj-pattern query="*OpenServer*" />
    <url-obj-pattern query="*OpenDatabase*" />
    <url-obj-pattern query="*OpenElement*" />
    <url-obj-pattern query="*OpenView*" />
    <url-obj-pattern query="*OpenAbout*" />
    <url-obj-pattern query="*OpenHelp*" />
    <url-obj-pattern query="*OpenIcon*" />
    <url-obj-pattern query="*OpenForm*" />
    <url-obj-pattern query="*OpenNavigator*" />
    <url-obj-pattern query="*OpenAgent*" />
    <url-obj-pattern query="*CreateDocument*" />
    <url-obj-pattern query="*DeleteDocument*" />
    <url-obj-pattern query="*EditDocument*" />
    <url-obj-pattern query="*SaveDocument*" />
    <url-obj-pattern query="*SearchSite*" />
    <url-obj-pattern query="*SearchView*" />
    <url-obj-pattern query="*&login*" />
    <url-obj-pattern query="*Command*" />

```

```

        <!-- Domino bejárás: OpenDocument permutációk elkerülése -->
        <url-obj-pattern    query="*ExpandSection*" />
        <url-obj-pattern    query="*Navigate*" />
        <url-obj-pattern    query="*Start*" />
    <!-- -->

    </exclude-pattern-list>
</crawler-config>

<summarizer-config>
<!-- A Copy2Rdf az alapértelmezett összesítő.
      Használat az alábbi típusoknál: -->

    <resource-handler content-type="*htm*"
                      summarizable="EipHtmlSummarizable"
                      summary-maker="EipHtmlRawSummaryMaker" />
    <resource-handler content-type="*pdf"
                      summarizable="InsoSummarizable"
                      summary-maker="InsoSummaryMaker" />

</summarizer-config>
v    </group>
    </group-list>
</gcs-config>

```

---

## IBM Web Crawler napló elemzési fájl példa

```
D:\gcs\run\log>perl loganalysis.pl log.txt
```

Elapsed time in Log.txt for 7710 lines was 1.84 minutes.

```

GCS was configured for 20 crawlers
999 total crawls attempted
137 - total crawl failures:
      21 GCSHttpConnection.ABANDONING
      12 GCSHttpConnection.CONNECT_ERROR
      16 GCSHttpConnection.UNKNOWN_HOST
      4  HTTP 403
      29 HTTP 404
      2  HTTP 500
      8  HTTP 599
      1  Read timed out
      39 Robots not allowed
      4  over max redirects
      1  unknown protocol

```

```

-----
862 = successfully crawled
0   - unchanged since earlier crawl

```

```

-----
862 = new or changed
468 crawled per minute

```

```

GCS was configured for 5 summarizers
855 total summaries attempted
0 - total summary failures:
-----
855 = successfully summarized
144 gcs.summaries.list.host
855 gcs.summaries.list.resource
465 summarized per minute

```

```

GCS successfully crawled 134 servers to obtain 862 URL:
afqa0854.mop.ibm.com: 15
als1f1.yamato.ibm.com: 1
apache.btv.ibm.com: 1

```

apc.endicott.ibm.com: 2  
 as400service.ibm.com: 1  
 atlas.bocaraton.ibm.com: 1  
 autopproxy.ibm.com: 1  
 cer.si.ibm.com: 1  
 commerce.www.ibm.com: 1  
 crmweb.boulder.ibm.com: 3  
 d02ntcl01.ibm.com: 1  
 dacs.endicott.ibm.com: 1  
 duke.toraix.can.ibm.com: 1  
 ebcweb.austin.ibm.com: 1  
 ecspubs.ibmus2.ibm.com: 5  
 edaw3.fishkill.ibm.com: 1  
 endwww.endicott.ibm.com: 1  
 gcs.stl.ibm.com: 1  
 gustwick.austin.ibm.com: 1  
 ibmfnsys.somers.hqregion.ibm.com: 1  
 ibmpnyil.somers.hqregion.ibm.com: 2  
 ifw-www.mul.ie.ibm.com: 1  
 iplswww.nas.ibm.com: 2  
 itirc.ibm.com: 1  
 logosite.services.ibm.com: 1  
 lt.lahulpe.ibm.com: 17  
 messaging.ibm.com: 1  
 mrsmrn04.leeds.uk.ibm.com: 1  
 online.lahulpe.ibm.com: 1  
 page.sg.ibm.com: 1  
 procure.sbyl.ibm.com: 1  
 reso.somers.hqregion.ibm.com: 1  
 ristal.leipzig.de.ibm.com: 1  
 rrhhar.argentina.ibm.com: 1  
 seashore.stl.ibm.com: 1  
 secureway.raleigh.ibm.com: 15  
 service.software.ibm.com: 1  
 software.ibmus2.ibm.com: 1  
 techcenter.austin.ibm.com: 1  
 tr2.fishkill.ibm.com:8080: 1  
 ucd.torolab.ibm.com: 1  
 usmweb.boulder.ibm.com: 1  
 w3-1.ibm.com: 32  
 w3-2.ibm.com: 3  
 w3-3.ibm.com: 108  
 w3-5.ibm.com: 4  
 w3.a-nz.au.ibm.com: 1  
 w3.academy.ibm.com: 1  
 w3.almaden.ibm.com: 2  
 w3.alphaworks.ibm.com: 1  
 w3.ap.ibm.com: 1  
 w3.asca.ibm.com: 7  
 w3.austin.ibm.com: 3  
 w3.boulder.ibm.com: 1  
 w3.br.ibm.com: 1  
 w3.btv.ibm.com: 1  
 w3.can.ibm.com: 40  
 w3.chq.ibm.com: 4  
 w3.coc.ibm.com: 1  
 w3.corporatetechnology.ibm.com: 1  
 w3.cupertino.ibm.com: 1  
 w3.dds.dfw.ibm.com: 17  
 w3.demopkg.ibm.com: 4  
 w3.design.ibm.com: 1  
 w3.developer.ibm.com: 3  
 w3.education.ibm.com: 1  
 w3.emea.ibm.com: 14  
 w3.enterlib.ibm.com: 7  
 w3.finsys.ibm.com: 1  
 w3.gcg.ibm.com: 1

w3.globalfinancing.de.ibm.com: 1  
w3.hakozaki.ibm.com: 1  
w3.houston.ibm.com: 1  
w3.hursley.ibm.com: 5  
w3.iabc.ibm.com: 1  
w3.ibm.com: 180  
w3.ibmfax.ibm.com: 1  
w3.ibmlla.ibm.com: 14  
w3.isicc.de.ibm.com: 1  
w3.itso.ibm.com: 1  
w3.japan.ibm.com: 1  
w3.knowledge.raleigh.ibm.com: 1  
w3.linux.ibm.com: 1  
w3.marketing.ibm.com: 1  
w3.micro.ibm.com: 2  
w3.mtlisc.can.ibm.com: 1  
w3.munich.ibm.com: 1  
w3.ode.raleigh.ibm.com: 1  
w3.paylink.au.ibm.com: 1  
w3.pisc.uk.ibm.com: 1  
w3.pl.ibm.com: 1  
w3.printers.ibm.com: 1  
w3.pssc.mop.ibm.com: 1  
w3.pssed.au.ibm.com: 1  
w3.raleigh.ibm.com: 3  
w3.rchland.ibm.com: 1  
w3.research.ibm.com: 3  
w3.reserve.ibm.com: 1  
w3.rs6000.ibm.com: 1  
w3.security.ibm.com: 1  
w3.software.ibm.com: 6  
w3.ssd.ibm.com: 1  
w3.stl.ibm.com: 1  
w3.techline.ibm.com: 1  
w3.techsupp.yamato.ibm.com: 1  
w3.torolab.ibm.com: 2  
w3.usergroup.ibm.com: 1  
w3.vendor.pok.ibm.com: 1  
w3.viewblue.ibm.com: 1  
w3.watson.ibm.com: 2  
w3.wdg.uk.ibm.com: 1  
w3.ytal.yasu.ibm.com: 1  
w3.zurich.ibm.com: 1  
w3chq.disbursements.ibm.com: 1  
w3is.lagaude.ibm.com: 1  
w3md.btv.ibm.com: 1  
w3ssd.mainz.de.ibm.com: 1  
w3vm.demopkg.ibm.com: 1  
widweb.raleigh.ibm.com: 1  
wtscpok.itso.ibm.com: 1  
wwas.raleigh.ibm.com: 1  
www-1.ibm.com: 63  
www-3.ibm.com: 4  
www-4.ibm.com: 86  
www.almaden.ibm.com: 1  
www.as400.ibm.com: 1  
www.chips.ibm.com: 1  
www.ibm.com: 52  
www.ieg.ibm.com: 1  
www.patents.ibm.com: 1  
www.pc.ibm.com: 2  
www.rs6000.ibm.com: 23  
www.software.ibm.com: 9  
www.storage.ibm.com: 1  
www.watson.ibm.com: 1

GCS timed out 1 times:

w3-3.ibm.com: 1

GCS ignored 42 URL prohibited by robots.txt:

- reso.somers.hqregion.ibm.com: 1
- w3.education.ibm.com: 1
- w3.rchland.ibm.com: 34
- w3.zurich.ibm.com: 1
- www.ibm.com: 5

GCS skipped 3846 URL (requires gcs.url logging)

- 59 specified an unsupported protocol:
  - protocol not supported gopher: 6
  - protocol not supported mailto: 53
- 1206 had content-types (lower or UPPER case, > 10) that were not included
  - .2: 12
  - .faq: 13
  - .1: 14
  - .asp: 16
  - .cgi: 21
  - .shtml: 90
  - .pl: 92
  - .gif: 157
  - .nsf: 160
  - .jpg: 214
  - .css: 240
- 516 URL were on servers and/or paths that were not included
- 2065 were excluded for these reasons:
  - URL longer than 254: 1
  - excluded by rule 1: 1210
  - excluded by rule 2: 854



---

## 8. fejezet A szöveges keresés és a QBIC használata

A szakasz első része a szöveges keresés és a Lekérdezés képtartalom alapján (QBIC) beállítását és használatát mutatja be. A két szolgáltatás csak akkor érhető el, ha a Content Manager 7.1 kapcsolatot telepíti az EIP telepítésekor. A szakasz második része a példa alkalmazásokhoz használt példa szövegek és kép adatok betöltéséről tartalmaz információkat.

---

### Dokumentumok keresése a szöveg keresőmotorral

A Content Manager 7.1 kiszolgálókba integrálható a szöveg keresési funkció, így lehetségessé válik a Content Manager környezetben tárolt dokumentumok automatikus indexelése, keresése és visszaadása. A felhasználók szavakra vagy kifejezésekre keresve találhatják meg a dokumentumokat. A szöveg keresési kiszolgáló az egybyte-os és duplabyte-os karakterkészleteket is támogatja, és AIX vagy Windows rendszereken fut.

A szöveg keresés tartalmazza a strukturált XML, HTML és címkézett ASCII dokumentumok támogatását, amely lehetővé teszi a kifejezések keresését a dokumentumok meghatározott szakaszaiban. Az adatok egymásba ágyazott szakaszokban is kereshetők. Az XML teljes szöveggörnyezete kereshető, megadható például olyan keresés, amely az IBM kifejezést keresi a címek között, valamint olyan is, amely az IBM kifejezést egy adott szakasz címében keresi. DTD elérési út meghatározásakor a szöveg keresés dinamikusan használja az egyes dokumentumok megfelelő DTD fájljait, feltéve, hogy a DTD a dokumentum metaadataként van tárolva.

A szöveg keresést használó EIP rendszer tervezéséről és telepítéséről az *Enterprise Information Portal tervezése és telepítése* kiadványban talál információkat.

### Szöveg keresési kiszolgáló engedélyezése

Ha szöveg keresési kiszolgálót kíván használni, akkor engedélyeznie kell az adminisztrációját, mielőtt elindítaná az IBM Content Manager for Multiplatforms adminisztrációs ügyfelet. Az adminisztráció engedélyezése:

1. Indítsa el az IBM Content Manager for Multiplatforms könyvtárkiszolgálót.  
Hagyja, hogy a könyvtárkiszolgáló befejezze az indexosztályok felépítését.
2. A következő paranccsal indítsa el a szöveg keresési kiszolgálót azon a munkaállomáson, amelyre telepítette:  
`imlss -start dlinst`

ahol a *dlinst* a szöveg keresési kiszolgálónak a telepítés során, vagy az *imlcfgsv* parancs használatával megadott példányneve.

---

### Képek keresése a képek tartalma alapján (QBIC)

Ez a szakasz a Lekérdezés képtartalom alapján (QBIC) szolgáltatást, valamint a QBIC szolgáltatás beállítását mutatja be. A QBIC szolgáltatás csak akkor használható, ha telepítette a Content Manager 7.1 kapcsolatot. A QBIC szolgáltatás Windows és AIX operációs rendszerekkel kompatibilitás.

### Bevezetés a kép keresésbe

A kép keresési kiszolgáló az IBM QBIC (lekérdezés képtartalom alapján) technológiáját alkalmazza az objektumok vizuális tulajdonságainak, például szín vagy kitöltés alapján végrehajtott kereséséhez. A kép keresési kiszolgáló elemzi a képeket, és egy adatbázisban

tárolja az információkat. A felhasználók ezzel olyan kép kereséseket hajthatnak végre, amelyek a szóbeli körülírás helyett a kép vizuális tulajdonságain alapulnak, például színek, kitöltési minták, és ezek elhelyezkedésének megadásával. A tartalomalapú lekérdezések és a szöveges vagy kulcsszó alapján végrehajtott keresések kombinálásával hatékony lekérdezési módszerek alakíthatók ki a kép és multimédia adatokhoz.

Minden kép keresési kiszolgáló rendelkezik egy adatkönyvtárral, amely a kép keresési katalógusokat tartalmazó kép keresési adatbázisokat tárolja. A kép keresési katalógusok tartalmazzák a képgyűjtemények vizuális jellemzőire vonatkozó adatokat. A tényleges képjelölőket az IBM Content Manager for Multiplatforms rendszer objektumkiszolgálói tárolják. A kép keresési kiszolgáló AIX és Windows rendszereken fut.

A kép keresés telepítéséről az *Enterprise Information Portal tervezése és telepítése* kiadványban talál információkat.

## Kép keresés beállítása

Az alábbi utasításokat a Kép keresés telepítés után kell alkalmazni. A kép keresést a telepítő a Content Manager 7.1 kapcsolat kiválasztása esetén automatikusan telepíti. A kép keresés beállítása a következő feladatokat jelenti:

1. Környezet beállítása
2. Kép keresési kiszolgáló beállítása
3. Kép keresési ügyfél beállítása
4. Mintaadatok betöltése

**AIX rendszereken a telepítő varázsló használatakor** nem kell lefuttatni a konfigurációs és beállítási parancsfájlokat, illetve nem kell kiadni a kiszolgáló konfigurációs parancsait. Ezeket a feladatokat a varázsló végzi el.

**Windows alatt** viszont el kell végezni ezen feladatokat is.

### Környezet beállítása

A szakasz környezeti beállításokkal kapcsolatos feladatait a kiszolgáló és ügyfél számítógépeken is végre kell hajtani. A kép keresési kiszolgálónak az alábbi környezeti változókra van szüksége:

#### QBICTOP

Ennek segítségével történik a fájlnevek feloldása a kép keresés beállítása során.

#### QbicImagePath

Ennek segítségével történik a kiszolgáló képfájlok fájlneveinek feloldása.

#### QbicMaskPath

Ennek segítségével történik a kiszolgáló maszkfájlok fájlneveinek feloldása.

#### QbicSketchPath

Ennek segítségével történik a kiszolgáló vázlatfájlok fájlneveinek feloldása.

#### QbicTextPath

Ennek segítségével történik a kiszolgáló szövegfájlok fájlneveinek feloldása.

A kép keresési ügyfélnek csak a QBICTOP környezeti változóra van szüksége.

**AIX példa:** AIX rendszeren futtassa a beállítási parancsfájl előállító konfigurációs parancsfájl, majd futtassa a beállítási parancsfájl a környezet beállításához.

1. Futtassa a következő konfigurációs parancsfájl:  
`/usr/lpp/cmb/bin/frnconfg.iss QBICTOP`

ahol a QBICTOP a vezérlőfájlok (\*.ini) könyvtárának elérési útja. A QBICTOP változót állítsa be a /felhasználó/cmb/qbic értékre, ahol a /felhasználó a kép keresés adminisztrátori felhasználójának saját könyvtára. A kép keresési felhasználói azonosítónak olvasás/írás hozzáféréssel kell rendelkeznie a könyvtárhoz.

A parancsfájl előállítja az frnsetup.iss beállítási parancsfájlt.

2. A kép keresési felhasználói azonosító saját könyvtárából futtassa a következő parancsot:  
./frnsetup.iss

Ez a parancsfájl tölti fel a kép keresési kiszolgálók és ügyfelek környezeti változóit.

**Windows példa:** A környezeti változók beállítása:

1. Kattintson a **Start** → **Beállítások** → **Vezérlőpult** elemekre.
2. Kattintson duplán a **Rendszer** ikonra.
3. Kattintson a **Környezet** lapra.
4. Állítsa be a 8. táblázat: helyen felsorolt környezeti változókat a megfelelő mezők kitöltésével és a **Beállítás** gomb megnyomásával.

**Követelmény:** A kép keresési ügyfélnek csak a QBICTOP változóra van szüksége. Ügyfél környezetekben csak a QBICTOP változót állítsa be.

8. táblázat: Kép keresési környezeti változók

Változó	Érték
QBICTOP	d:\cmbroot\iss
QbicImagePath	d:\cmbroot\iss
QbicMaskPath	d:\cmbroot\iss
QbicSketchPath	d:\cmbroot\iss
QbicTextPath	d:\cmbroot\iss

Ahol a d: a kép keresés telepítési meghajtója.

## Kép keresési kiszolgáló beállítása

A kép keresési kiszolgálót a futtatás előtt be kell állítani. A kiszolgáló beállítása a kezdeti konfigurálásból és a kapcsolat ellenőrzéséből áll.

A kiszolgáló beállítása:

1. Indítsa el a parancsértelmezőt a qbicadm parancssal.
2. Írja be a **config server** parancsot. Például:  
config server LIBSRVRN FRNADMIN JELSZÓ 9999

ahol a LIBSRVRN a könyvtárkiszolgáló neve, az FRNADMIN a Content Manager felhasználói azonosító, a JELSZÓ a Content Manager jelszó, és a 9999 a kép keresési kiszolgáló portszáma.

További információkat a “Kapcsolat ellenőrzése” oldalszám: 112 részben talál.

## Kép keresési ügyfél beállítása

A kép keresési ügyfél, valamint a kép keresés rendszeradminisztrációs programjának indítása előtt be kell állítani az ügyfelet. A Content Manager rendszeradminisztráció egy álnév hozzárendelését igényli. Ellenőrizze a konfigurációt a kapcsolat ellenőrzésével.

**Álnév hozzárendelése:** Mielőtt használhatná a kép keresés kép keresési ügyfélként működő Content Manager rendszeradminisztrációs programját, hozzá kell rendelnie legalább egy kiszolgáló álnevet.

Álnév hozzárendelése:

1. Indítsa el a parancsértelmezőt a `qbicadm` paranccsal.
2. Írja be az **add alias** parancsot. Például:  
`add alias QBICSRV HOSZT 9999`

ahol a QBICSRV az álnév, a HOSZT a kép keresési kiszolgáló hosztneve, a 9999 pedig a kép keresési kiszolgáló portszáma.

#### Kapcsolat ellenőrzése:

##### Fontos:

1. A kép keresési kiszolgálóra csatlakozáshoz futnia kell a könyvtárkiszolgálónak is.
2. A kép keresés rendszeradminisztrációjához szükség van egy meglévő Content Manager felhasználói azonosítóra. Ahhoz, hogy sikeresen csatlakozzon a könyvtárkiszolgálóhoz, a kép keresési felhasználói azonosítónak és a könyvtárkiszolgáló felhasználói azonosítónak meg kell egyeznie. A felhasználói azonosító alapértelmezett értéke az `frnadmin`. Ha módosítja az értéket, akkor győződjön meg róla, hogy az azonosítók megegyeznek.

A kapcsolat ellenőrzése:

1. A kép keresési kiszolgáló beállítása, és az álnév létrehozása után indítsa el a kiszolgálót a parancssorból kiadott `commsrv` paranccsal.
2. A parancsértelmező indításához használja a `qbicadm` parancsot.
3. A parancsértelmezőben adja ki a `connect` parancsot.  
`connect QBICSRV FRNADMIN JELSZÓ`

ahol a QBICSRV az álnév, az FRNADMIN a Content Manager felhasználói azonosító, a JELSZÓ pedig a Content Manager jelszó.

A sikeres csatlakozás után megjelenik a Könyvtárkiszolgáló a(z) LIBSRVRN üzenet.

4. A kapcsolatbontáshoz írja be a `disconnect` parancsot.
5. A parancsértelmezőből a `quit` paranccsal léphet ki.

---

## Mintaadatok betöltése és indexelése

Ez a rész írja le a példa alkalmazásokban felhasználható mintaszövegek és képadatok betöltését és indexelését. Ez a rész csak akkor alkalmazható, ha telepítette a Content Manager 7.1 kapcsolatot, és kiválasztotta a szöveg keresés lehetőséget.

Az Enterprise Information Portal CD-n többféle mintabetöltési program is található. Ebben a szakaszban azt írjuk le, hogyan tölthetők be a szöveges és képi adatok a LoadSampleTSQBICDL mintabetöltési program segítségével. A szöveges és képi adatok külön is betölthetők, így ellenőrizhető mindkét szolgáltatás megfelelő működése.

## Mielőtt betöltené a mintaadatokat

A betöltési program futtatása előtt el kell végezni az alábbiakat:

1. Jelentkezzen be az EIP adminisztrációs ügyfélbe. Kattintson a Start → Programok → **Enterprise Information Portal** for Multiplatforms 8.2 → Adminisztráció menüpontra.
2. Válasszon ki egy adatbázist, majd jelentkezzen be a megfelelő felhasználói azonosítóval és jelszóval. Ha az alapértelmezett adatbázist (`icmnlsdb`) választja, akkor akkor adja meg az **icmadmin** felhasználói azonosítót és a **password** jelszót. Ha más adatbázist használ, akkor adja meg a megfelelő felhasználói azonosítót.

3. Hozzon létre egy könyvtárkiszolgáló konfigurációt a Content Manager rendszeradminisztrációs programjával. A feladatról a rendszeradminisztrációs program súgója nyújt segítséget.
4. Módosítsa a könyvtárkiszolgáló konfigurációjának **Hozzáférési** tulajdonságait a következők végrehajtásával:
  - a. Kattintson a jobb egérgombbal az új konfigurációra, majd válassza az előugró menü **Tulajdonságok** menüpontját a Tulajdonságok párbeszédablak megjelenítéséhez.
  - b. Kattintson a **Hozzáférés** lapra.
  - c. Jelölje ki a **Korlátlan számú munkamenet tetszőleges munkaállomásról** választógombot.

## Szöveg keresési index létrehozása

Az adatok betöltése előtt létre kell hozni egy üres szöveg keresési indexet, amelyet a szöveges példák indexeléséhez fog használni. **Tipp:** Szöveg keresési index csak Content Manager 6.1 és 7.1 kiszolgálókon hozható létre.

Üres szöveg keresési index létrehozása:

1. A következő paranccsal indítsa el a szöveg keresési kiszolgálót azon a munkaállomáson, amelyre telepítette:

```
imlss -start dlinst
```

ahol a **dlinst** a szöveg keresési kiszolgálónak a telepítés során, vagy az **imlcfgsv** parancs használatával megadott példányneve.

2. Indítsa el a Content Manager rendszeradminisztrációs programját, és jelentkezzen be.
3. A bal felső ablakrész található listájából válassza ki a **Szöveg keresés** elemet.
4. Kattintson duplán a bal oldali panel **Keresési kiszolgálók** elemére.
5. Kattintson duplán a **TM** elemre. A TM a szöveg keresési kiszolgáló keresési kiszolgáló álneve.
6. Kattintson a bal oldali panel **Indexek** mappájára. Ha megjelenik az **RC\_EMPTY\_LIST** üzenet, akkor válassza ki a menüsor **Kijelölt** → **Új** elemeit egy index létrehozásához.
7. Az **Új** index ablakban állítsa be az új indexet. Az egyes mezők részletes leírásának megjelenítéséhez kattintson a **Súgó** gombra.

Például:

**Windows rendszereken:**

Név      TMINDEX

Típus    Pontos

**Indexfájlok**

x:\cmbroot\ts\index\tmlindex ahol x: a telepítési meghajtó; ha az elérési út nem létezik, akkor létrejön.

**Index munkafájlok**

x:\cmbroot\ts\work\tmlindex ahol x: a telepítési meghajtó; ha az elérési út nem létezik, akkor létrejön.

**Információs bejegyzés**

A Content Manager könyvtárkiszolgáló neve.

Ne módosítsa az ügyfél és a kiszolgáló alapértelmezett DLL neveit.

**AIX rendszereken:**

Név      TMINDEX

**Típus**    **Pontos**

**Indexfájlok**

/home/cltadmin/tsindex/index/tminindex; ha az elérési út nem létezik, akkor létrehozásra kerül.

**Index munkafájlok**

/home/cltadmin/tsindex/work/tminindex; ha az elérési út nem létezik, akkor létrehozásra kerül. Győződjön meg róla, hogy a felhasználó írhatja a könyvtárat.

**Információs bejegyzés**

A Content Manager könyvtárkiszolgáló neve.

8. Kattintson az **OK** gombra.
9. Kattintson duplán a **TMINDEX** elemre a TMINDEX adminisztrációs párbeszédablak megjelenítéséhez.

## Kép keresési adatbázis, katalógus és szolgáltatások létrehozása

A minta szöveg keresési adatok szövegindexének létrehozása után létre kell hozni a minta képi adatok kép keresési adatbázisát és katalógusát.

A kép keresési adatbázis, katalógus és szolgáltatások létrehozása:

1. A következő paranccsal indítsa el a kép keresési kiszolgálót azon a munkaállomáson, amelyre telepítette:  
commsrv
2. Indítsa el a Content Manager rendszeradminisztrációs programját, és jelentkezzen be.
3. A bal felső panel listájából válassza ki a **Kép keresés** elemet.
4. Kattintson a bal oldali panel **Kép keresési kiszolgálók** elemére.
5. Kattintson a **QBICSRV** elemre.  
ahol a QBICSRV a kép keresési kiszolgálónak a telepítés során megadott neve.
6. Kattintson a jobb egérgommbbal a bal oldali panel **Adatbázisok** elemére, majd válassza az előugró menü **Új adatbázis** menüpontját.
7. Az Új adatbázis ablak **Név** mezőjébe írja be a MINTADB nevet, majd kattintson az **OK** gombra.
8. Kattintson a bal oldali panel **Adatbázisok** ikonjára a MINTADB ikon megjelenítéséhez a bal oldali panelen.
9. Kattintson a MINTADB ikonra.
10. A bal oldali panelen kattintson a jobb egérgommbbal a **Katalógusok** elemre, majd válassza az előugró menü **Új katalógusok** menüpontját.
11. Az Új katalógus ablak **Név** mezőjébe írja be a MINTAKAT nevet, majd kattintson az **OK** gombra.
12. A bal oldali panelen kattintson a **Katalógusok** elemre a MINTAKAT ikon megjelenítéséhez.
13. Kattintson a MINTAKAT elemre.
14. A bal oldali panelen kattintson a jobb egérgommbbal a **Szolgáltatások** ikonra, majd válassza az előugró menü **Új szolgáltatások** menüpontját.
15. Az Új szolgáltatás ablak **Név** mezőjében válasszon ki minden szolgáltatást, majd kattintson az **Alkalmaz** gombra. Mind a négy szolgáltatás kiválasztásakor a következő üzenet jelenik meg:  
Minden lehetséges szolgáltatás bekerült a katalógusba.
16. Kattintson az **OK** gombra.

17. Kattintson a **Mégse** gombra.

## A betöltési program futtatása

A szöveg és kép keresés ellenőrzéséhez tölts be a mintaadatokat.

A minta képi adatok az alábbi fájlokban találhatóak:

### Windows rendszereken:

`x:\cmbroot\samples\java\dl\samples.jar`

### AIX rendszereken:

`/usr/lpp/cmb/samples/java/dl/samples.jar`

A mintabetöltési program betölti az adatokat a Content Manager kiszolgálóra, és elvégzi az indexelést. A program futtatásának szintaxisát a forrásprogram elejében található bevezető tartalmazza. A mintabetöltési programok a következők:

### Windows rendszereken:

`x:\cmbroot\samples\java\dl\LoadSampleTSQBICDL.jar`

### AIX rendszereken:

`/usr/lpp/cmb/samples/java/dl/LoadSampleTSQBICDL.jar`

A mintabetöltési programok futtatása:

1. Bontsa ki a .jar fájlokat a következő paranccsal:

```
jar -xvf samples.jar
```

A fájl a megfelelő könyvtárakba bontja ki magát.

2. A mintabetöltési program lefordításához állítsa be a munkaállomás környezeti változóit az alábbiak szerint:

### Windows rendszereken:

- a. Nyissa meg a `x:\cmbroot\cmbenv71.bat` fájlt egy szövegszerkesztőben, és módosítsa az első három sort a munkaállomás környezeti változóinak beállításához:

```
set CMBROOT = e:\cmbroot
set DB2HOME = e:\sql11b
set JAVAHOME = d:\jdk117
```

- b. Mentse el a `cmbenv71.bat` fájlt, és állítsa be a környezeti változókat a következő paranccsal:

```
cmbenv71
```

### AIX rendszereken:

- a. Lépjen be az `/usr/lpp/cmb/bin/` könyvtárba, és futtassa a beállítások elvégzését a következő paranccsal:

```
./cmbenv71.sh
```

- b. Győződjön meg, hogy az `/usr/lpp/cmb/samples/java/dl/` alatti könyvtárakat és a mintafájlokat minden felhasználó írhatja.

3. Fordítsa le a betöltési programot az alábbi paranccsal (a parancsban a kis- és nagybetűk eltérőnek számítanak):

```
javac LoadSampleTSQBICDL.java
```

4. **Követelmény:** A betöltési program elindítása előtt az alábbi kiszolgálóknak futniuk kell:

- Könyvtárkiszolgáló
- Objektumkiszolgáló
- Szöveg keresési kiszolgáló



- Kép keresési kiszolgáló

Ha a futtatást a Content Manager nemzeti nyelvi változata alatt végzi, akkor a betöltési program futtatása előtt az **FRNDEFLANG** változót állítsa **ENU**-ra. A környezeti változót beállító AIX parancs: **export FRNDEFLANG=ENU**

5. Töltse be a mintaadatokat a betöltési programmal a következő parancs segítségével:  
`java LoadSampleTSQBICDL sampleQBIC.dat load.log frnadmin password LIBSRVRN`

ahol az **frnadmin** a felhasználói azonosító, a **password** a jelszó, a **LIBSRVRN** pedig a könyvtárkiszolgáló.

6. Nézze meg a **load.log** fájlt, és győződjön meg róla, hogy a mintaadatok betöltése sikeresen befejeződött.

A mintaadatok betöltésének befejezése után a Content Manager rendszeradminisztrációs programjával vagy a szöveg keresés parancssori eszközeivel indexelje le a példa szöveges adatait.

## A példa szöveges adatainak indexelése

A mintaadatok indexelése:

1. Indítsa el a Content Manager rendszeradminisztrációs programját, és jelentkezzen be.
2. A bal felső ablakrész található listájából válassza ki a **Szöveg keresés** elemet.
3. Kattintson duplán a **Keresési kiszolgálók** elemre.
4. Kattintson duplán a **TM** elemre. A TM a szöveg keresési kiszolgáló keresési kiszolgáló álneve.
5. Kattintson a jobb egérgombbal az **új szövegindexre**, majd válassza az előugró menü **Tulajdonságok** menüpontját.
6. A Tulajdonságok ablak Explicit lapján kattintson a **Frissítés** gombra.
7. Az **Index számláló** mezőben meg kell jelennie a betöltési programmal betöltött dokumentumok számának.
8. Kattintson az **Index** gombra a fájlok indexeléséhez.
9. Néhány perc után kattintson a **Frissítés** gombra, és nézze meg az **Elsődleges dokumentumindex** mezőben a sikeresen indexelt dokumentumok számát.

Az adatok leindexelése után egy Java mintaalkalmazással lekérdezheti az adatokat, vagy futtathat egy egyszerű keresést az **imlsrch** parancssori eszköz segítségével.



---

## 9. fejezet Dokumentumformátumok

---

### Információbányászat dokumentumformátumai

Ez a függelék írja le az információbányászat által támogatott dokumentumformátumokat.

#### Szövegszerkesztők: Általános

<b>ANSI szöveg (7 és 8 bites)</b>	Minden változat
<b>ASCII szöveg (7 és 8 bites)</b>	Minden változat
<b>HTML</b>	3.0 változatig (bizonyos korlátozásokkal)
<b>IBM FFT</b>	Minden változat
<b>IBM Revisable Form Text</b>	Minden változat
<b>Microsoft Rich Text Format (RTF)</b>	Minden változat
<b>Unicode szöveg</b>	Minden változat

#### Szövegszerkesztők: DOS

<b>DEC WPS Plus (DX)</b>	4.0 változatig
<b>DEC WPS Plus (WPL)</b>	4.1 változatig
<b>DisplayWrite 2 &amp; 3 (TXT)</b>	Minden változat
<b>DisplayWrite 4 &amp; 5</b>	2.0 kiadásig
<b>Enable</b>	3.0, 4.0 és 4.5
<b>First Choice</b>	3.0 változatig
<b>Framework</b>	3.0 változat
<b>IBM Writing Assistant</b>	1.01 változat
<b>Lotus Manuscript</b>	2.0 változatig
<b>MASS11</b>	8.0 változatig
<b>Microsoft Word</b>	6.0 változatig
<b>Microsoft Works</b>	2.0 változatig
<b>MultiMate</b>	4.0 változatig
<b>Navy DIF</b>	Minden változat
<b>Nota Bene</b>	3.0 változat
<b>Office Writer</b>	4.0 és 6.0 változat
<b>PC-File Letter</b>	5.0 változatig
<b>PC-File+ Letter</b>	3.0 változatig
<b>PFS:Write</b>	A, B és C változat
<b>Professional Write</b>	2.1 változatig

<b>Q&amp;A</b>	2.0 változat
<b>Samna Word</b>	4.0 változatig
<b>SmartWare II</b>	1.02 változat
<b>Sprint</b>	1.0 változatok
<b>Total Word</b>	1.2 változat
<b>Volkswriter 3 &amp; 4</b>	1.0 változatig
<b>Wang PC (IWP)</b>	2.6 változatig
<b>WordMARC</b>	Composer Plus változatig
<b>WordPerfect</b>	6.1 változatig
<b>WordStar</b>	7.0 változatig
<b>WordStar 2000</b>	3.0 változatig
<b>XyWrite</b>	III Plus változatig

## Szövegszerkesztők: Nemzetközi

<b>JustSystems Ichitaro</b>	5.0, 6.0, 8.0, 9.0 és 10.0 változat
-----------------------------	-------------------------------------

## Szövegszerkesztők: Windows

<b>AMI/AMI Professional</b>	3.1 változatig
<b>Corel WordPerfect for Windows</b>	9.0 változatig
<b>JustWrite</b>	3.0 változatig
<b>Legacy</b>	1.1 változatig
<b>Lotus WordPro (Win32 / Intel platformok)</b>	SmartSuite 96, 97 és Millennium
<b>Lotus WordPro (Unix platformok - csak szöveg)</b>	SmartSuite 97 és Millennium
<b>Microsoft Windows Works</b>	4.0 változatig
<b>Microsoft Windows Write</b>	3.0 változatig
<b>Microsoft Word 97</b>	Word 97
<b>Microsoft Word 2000</b>	Word 2000
<b>Microsoft Word for Windows</b>	7.0 változatig
<b>Microsoft WordPad</b>	Minden változat
<b>Novell Perfect Works</b>	2.0 változat
<b>Novell WordPerfect for Windows</b>	7.0 változatig
<b>Professional Write Plus</b>	1.0 változat
<b>Q&amp;A Write for Windows</b>	3.0 változatok
<b>WordStar for Windows</b>	1.0 változat

## Szövegszerkesztők: Macintosh

Microsoft Word	4.0 - 6.0 változat
Microsoft Word 98	Word 98
WordPerfect	1.02 - 3.0 változat
Microsoft Works	2.0 változatig
MacWrite II	1.1 változat

## Táblázatkezelő formátumok

VP Planner 3D	1.0 változat
Enable	3.0, 4.0 és 4.5
First Choice	3.0 változatig
Framework	3.0 változat
Lotus 1-2-3 (DOS és Windows)	5.0 változatig
Lotus 1-2-3 for SmartSuite	SmartSuite 97 és Millennium
Lotus 1-2-3 diagramok (DOS és Windows)	5.0 változatig
Lotus 1-2-3 (OS/2)	2.0 változatig
Lotus 1-2-3 diagramok (OS/2)	2.0 változatig
Lotus Symphony	1.0, 1.1 és 2.0 változat
Microsoft Excel 97	Excel 97
Microsoft Excel 2000	Excel 2000
Microsoft Excel Macintosh	3.0 - 4.0 és 98 változatok
Microsoft Excel Windows	2.2 - 7.0 változatok
Microsoft Excel diagramok	2.x - 7.0 változatok
Microsoft Multiplan	4.0 változat
Microsoft Windows Works	4.0 változatig
Microsoft Works (DOS)	2.0 változatig
Microsoft Works (Mac)	2.0 változatig
Mosaic Twin	2.5 változat
Novell Perfect Works	2.0 változat
QuattroPro for DOS	5.0 változatig
QuattroPro for Windows	9.0 változatig
PFS:Professional Plan	1.0 változat
SuperCalc 5	4.0 változat
SmartWare II	1.02 változat

## Adatbázis formátumok

SmartWare II	1.02 változat
--------------	---------------

<b>Access</b>	2.0 változatig
<b>dBase</b>	5.0 változatig
<b>DataEase</b>	4.x változat
<b>dBXL</b>	1.3 változat
<b>Enable</b>	3.0, 4.0 és 4.5
<b>First Choice</b>	3.0 változatig
<b>FoxBase</b>	2.1 változat
<b>Framework</b>	3.0 változat
<b>Microsoft Windows Works</b>	4.0 változatig
<b>Microsoft Works (DOS)</b>	2.0 változatig
<b>Microsoft Works (Mac)</b>	2.0 változatig
<b>Paradox (DOS)</b>	4.0 változatig
<b>Paradox (Windows)</b>	1.0 változatig
<b>Personal R:BASE</b>	1.0 változat
<b>R:BASE 5000</b>	3.1 változatig
<b>R:BASE System V</b>	1.0 változat
<b>Q &amp; A</b>	2.0 változatig
<b>Reflex</b>	2.0 változat

## Szabványos grafikus formátumok

### **PNG - Portable Network Graphics (Internetes formátum)**

1.0 változat

### **Bináris G3 Fax**

Minden változat

### **BMP (beleértve az RLE, ICO, CUR és OS/2 DIB formátumokat)**

Windows

### **CDR (TIFF kép beágyazása esetén)**

Coral Draw 2.0 – 9.0 változatok

### **CGM - Computer Graphics Metafile**

ANSI, CALS, NIST, 3.0 változat

### **CMX - Corel Clip Art Format**

5.x - 6.x változat

### **DCX (többoldalas PCX)**

Microsoft Fax

### **DRW - Micrografx Designer**

3.1 változat

### **DRW - Micrografx Draw**

4.0 változatig

### **DXF - AutoCAD Drawing Interchange Format (bináris és ASCII)**

14-es változatig

### **EMF**

Windows Enhanced Metafile

### **EPS Encapsulated PostScript**

TIFF kép beágyazása esetén

### **FMV - FrameMaker graphics**

Vektoros és raszterformátumok az 5.0 változatig

### **FPX - Kodak Flash Pix**

Nincs egyedi formátum

<b>GDF - IBM Graphics Data Format</b>	1.0 változat
<b>GEM - Graphics Environment Manager Metafile</b>	Bitkép és vektoros
<b>GIF - Graphics Interchange Format</b>	Compuserve
<b>GP4 - Group 4 CALS formátum</b>	Type I és Type II
<b>HPGL - Hewlett Packard Graphics Language</b>	2.0 változat
<b>IMG - GEM Paint</b>	Nincs egyedi változat
<b>JFIF (nem TIFF formátumú JPEG)</b>	Minden változat
<b>JPEG - Joint Photographic Experts Group formátuma</b>	Minden változat
<b>MET - OS/2 PM Metafile</b>	3.0 változat
<b>PBM - Portable Bitmap</b>	Nincs egyedi változat
<b>Kodak Photo CD</b>	1.0 változat
<b>PCD - PCX Bitmap</b>	PC Paintbrush
<b>Perfect Works (Draw)</b>	Novell 2.0 változat
<b>PGM - Portable Graymap</b>	Nincs egyedi változat
<b>PIC - Lotus 1-2-3 Picture File Format</b>	Nincs egyedi változat
<b>PICT1 &amp; PICT2 (Raszter)</b>	Macintosh szabvány
<b>PIF - IBM Picture Interchange Format</b>	1.0 változat
<b>PNTG</b>	MacPaint
<b>PPM - Portable Pixmap</b>	Nincs egyedi változat
<b>Progresszív JPEG</b>	Nincs egyedi változat
<b>PSP - Paintshop Pro (csak Windows)</b>	5.0 és 5.0.1
<b>RND - AutoShade Rendering File Format</b>	2.0 változat
<b>SDW Ami Draw Snapshot (Lotus)</b>	Minden változat
<b>SRS - Sun Raster File Format</b>	Nincs egyedi változat
<b>Targa</b>	Truevision
<b>TIFF</b>	6-os változatig
<b>TIFF CCITT Group 3 &amp; 4</b>	Fax Systems
<b>VISIO (4-es változat esetén csak oldal előkép mód), Visio 4, 5, 2000</b>	Visio 4, 5, 2000
<b>WMF</b>	Windows Metafile

<b>WordPerfect Graphics [WPG és WPG2]</b>	2.0 változatig
<b>XBM - X-Windows Bitmap</b>	x10 kompatibilis
<b>XPM - X-Windows Pixmap</b>	x10 kompatibilis
<b>XWD - X-Windows Dump</b>	x10 kompatibilis

## Professzionális grafikus formátumok

<b>PSD - Adobe Photoshop fájlformátum</b>	4.0 változat
<b>AI - Adobe Illustrator fájlformátum</b>	7.0 változatig
<b>CDR - Corel Draw</b>	8.0 változatig
<b>DSF - Micrografx Designer</b>	Windows 95, 6.0 változat
<b>DWG - AutoCAD Native Drawing Format</b>	12 - 14 változatok
<b>IGES - Initial Graphics Exchange Specification</b>	5.1 változat
<b>PDF - Portable Document Format</b>	Acrobat 2.1, 3.0 és 4.0, beleértve a japán PDF fájlokat is
<b>PS - Postscript</b>	Level 2

## Bemutató formátumok

<b>Microsoft PowerPoint for Macintosh</b>	4.0 és 98 változatok
<b>Corel Presentations</b>	8.0 és 9.0 változat
<b>Novell Presentations</b>	3.0 és 7.0 változat
<b>Harvard Graphics for DOS</b>	2.x és 3.x változat
<b>Harvard Graphics</b>	Windows változatok
<b>Freelance 96</b>	Freelance 96
<b>Freelance for Windows 95</b>	SmartSuite 97 és Millennium
<b>Freelance for Windows</b>	1.0 és 2.0 változat
<b>Freelance for OS/2</b>	2.0 változatig
<b>Microsoft PowerPoint for Windows</b>	7.0 változatig
<b>Microsoft PowerPoint 97</b>	PowerPoint 97
<b>Microsoft PowerPoint 2000</b>	PowerPoint 2000

## Tömörített és kódolt formátumok

<b>ZIP (PKWare)</b>	2.0g változatig
<b>GZIP</b>	Nincs egyedi változat
<b>LZA önkicsomagoló tömörítés</b>	Nincs egyedi változat

<b>LZH tömörítés</b>	Nincs egyedi változat
<b>Microsoft Iratgyűjtő</b>	7.0 és 97
<b>MIME (szöveges levél)</b>	Nincs egyedi változat
<b>UUEncode</b>	Nincs egyedi változat
<b>UNIX tömörítés</b>	Nincs egyedi változat
<b>UNIX TAR</b>	Nincs egyedi változat

## Egyéb

<b>vCard elektronikus névjegykártya</b>	2.1 változat
<b>Futtatható (EXE, DLL)</b>	Nincs egyedi változat
<b>Windows NT futtatható</b>	Nincs egyedi változat
<b>MSG (csak szöveg)</b>	Microsoft Outlook levélformátum
<b>Microsoft Project (csak szöveg)</b>	Project 98





---

## 10. fejezet Jogok kezelése

Ez a fejezet mutatja be az EIP jogkezelési szolgáltatását, és írja le a jogok kezelésének alapelveit. Elmagyarázza továbbá a tulajdon védelmére használható megjelölési technikákat.

---

### Szellemi tulajdon védelme

A digitális formában tárolt multimédia objektumok a vállalat szellemi tulajdonát képezik. Az ilyen objektumok védelme kritikus az üzletmenet szempontjából, különösen akkor, ha ezek az objektumok az Interneten is megtalálhatók, ahol a másolásuk viszonylag könnyű. A Content Manager megjelölési technológiájának segítségével megakadályozhatja a vállalat szellemi tulajdonának jogosulatlan felhasználását a digitális multimédia objektumok megjelölésével (védelem) vagy ujjlenyomattal való ellátásával (azonosítás).

Az értékes objektumok megjelölésére több tényező miatt is szükség lehet:

- Forrás azonosítása a jogosulatlan másolás vagy újrafelhasználás megakadályozása érdekében. Ezt másképp vízjelnek is nevezik, és általában látható.
- A tartalom címzettjének azonosítása a jogosulatlan másolás vagy újrafelhasználás megakadályozása érdekében. Ezt más néven ujjlenyomattal való ellátásnak is nevezik, és jellemzően nem látható.
- Kapcsolattartó biztosítása további információk beszerzéséhez.
- További információk, például dátum és idő biztosítása az értéknövelt terjesztési láncokban.

A digitális objektumok megjelölhetők, mielőtt eljutnának a vásárlókhoz. A vízjelek és ujjlenyomatok alkalmazhatók a szállítás előtt. Az ujjlenyomatok szállítás előtt alkalmazása viszont feltételezi, hogy az ismert, ezért szükség lehet arra, hogy a jelölés a szállítási folyamat részeként, dinamikusan kerüljön alkalmazásra. A jelölésnek a saját (ellenőrzött) környezetben, a szállítás előtt végzett alkalmazása növeli a biztonságot, mivel csökkenti az objektum megbolygatására tett kísérletek kockázatát.

A jelölések a kezelési és szállítási folyamatban több helyen is alkalmazhatók. A megfelelő megoldást az adott helyzet ismeretében lehet meghatározni. A megjelölések a folyamat alábbi pontjain alkalmazhatók:

- Az objektum tárolása előtt  
Ha az objektumhoz általános megjelölést kell használni (például egy tulajdonost azonosító vízjel esetén), akkor ennek alkalmazására a leginkább praktikus időpont az objektum tárolása. A Content Manager rendszerben lehetőség van az eredeti (jelölés nélküli) és a megjelölt objektum tárolására is. Ennek alternatívájaként megoldható, hogy csak a megjelölt objektum legyen tárolva, míg a jelöletlen objektum különálló lerakatba kerüljön.
- Az objektum tárolása után  
Ha meg kívánja jelölni a Content Manager rendszeren tárolt objektumokat, akkor ez megoldható úgy, hogy visszakeresi az objektumot, megjelöli azt, majd vagy lecseréli a jelöletlen objektumot a megjelölt változattal, vagy a megjelölt objektumot új elemként tárolja.
- Az objektum visszakeresésekor  
Ha a jelölés a címtől függően változhat, akkor lehetőség van az objektumok visszakeresése utáni dinamikus megjelölésre is. Ilyenkor az eredeti objektum helyett a megjelölt objektum terjeszthető.

Ha a rendszeren sok jelöletlen örökölt objektum található, és nem kíván időt vagy erőforrásokat szánni az összes objektum megjelölésére, akkor szintén a visszakereséskor alkalmazott megjelölés jelenti a jó megoldást.

---

## Megjelölési technikák alkalmazása

Többféle tartalom megjelölési technika közül lehet választani. Minden technika egy-egy adott problémát céloz meg, és eltérő ellenállást nyújt az eltávolítással és módosítással szemben.

A jelöléseket az alábbiak szerint jellemezhetjük:

- Szállított információk

**Vízjel** Azonosítja a tartalom forrását. Tartalmazhatja például az objektum tulajdonosát és verziószámát.

### Ujjlenyomat

Azonosítja a tartalom címzettjét. Tartalmazhatja például az objektum kézbesítési célját és helyét.

- Láthatóság

### Látható

A megjelölés látható és észrevehető.

### Láthatatlan

A megjelölés rejtve van az objektumon belül.

- Integritás

### Gyenge

A megjelölés bármilyen módosítás esetén megsérül.

### Erőteljes

A megjelölés ellenáll a megjelölt objektumot ért módosításoknak, például átméretezésnek, tömörítésnek, forgatásnak és kivágásnak.

- Alkalmazás időpontja
  - Az objektum létrehozásakor
  - Az objektum tárolásakor
  - Az objektum terjesztési céllal végzett visszakeresésekor
  - Az objektum fogadásakor a címzett munkaállomásán
- Hely
  - Ha egy látható jelzés az illegális újrafelhasználást hivatott megakadályozni, akkor alkalmazható például egy kép nagyobb területére. A megjelölés elhelyezhető a kép egy részletgazdagabb területén, ahonnan még nehezebb eltávolítani.
  - Láthatatlan megjelölés használata esetén a kép részletgazdag területe lehetővé teszi az adatok elrejtését a kép apró részletei között.
  - Ha egy látható jelzés a tulajdonjog jelzését szolgálja, akkor elhelyezhető a kép sarkába, ahol nem zavarja a látványt.
  - Látható és láthatatlan jelölés együttes alkalmazásakor először a láthatót kell alkalmazni.
- Formátum

### Bináris

Az ilyen jelölés lehet egy véletlen bitsorrend, amely a kép egészén keresztül ismétlődik. Ez a véletlen bitsorrend a kép megjelöléséhez vagy jelölésének eltávolításához használható kulcs.

A jelölés kép is lehet.

### **Szervezett adatok**

A jelölés tartalmazhat beágyazott szöveges adatokat.

## **Látható megjelölés**

A látható megjelölés a kép fölé helyezett olyan átlátszó maszk, amely esetén a kép és a maszk is látható. A nehezen eltávolítható látható megjelölések hatékonyan elrettenthetik a felhasználókat az objektumok illegális felhasználásától.

Látható jelzést érdemes használni az alábbi helyzetekben:

- Képek biztosítása a vásárlóknak áttekintés céljából úgy, hogy ne legyen kedvük újrafelhasználni az áttekintési példányokat
- Képek elhelyezése hirdetésben vagy az Interneten

## **Láthatatlan megjelölés**

A láthatatlan jelölések a képen oly módon elrejtett adatok, amelyek nem utalnak a kép módosítására. A jelölés alkalmazásához, felismeréséhez és visszafejtéséhez külön alkalmazás szükséges.

Láthatatlan jelzést érdemes használni az alábbi helyzetekben:

- Információk beágyazása a tulajdonjog azonosításához és a megjelölt objektumok illegális másolásának megakadályozásához (vízjel)
- Információk beágyazása a terjesztési útvonal nyomon követéséhez (ujjlenyomat)
- Feljegyzés vagy a kép címének beágyazása



---

## 11. fejezet Kisegítő lehetőségek

A termék több szolgáltatást is biztosít a kihívásokkal küzdő emberek segítéséhez. Ilyen szolgáltatások például a következők:

- A termék minden funkciója elérhető a billentyűzet segítségével is, egér használata nélkül.
- Bővített megjelenítési tulajdonságok támogatása.
- Kompatibilitás a segítő technológiákkal.
- Kompatibilitás az operációs rendszer kisegítő lehetőségeivel.
- Elérhető dokumentációs formátumok.

---

### Billentyűzetről végzett bevitel és navigáció

A billentyűzettel végzett adatbevitelnél és a navigációnál a következő lehetőségek állnak rendelkezésre:

#### Billentyűzetről végzett adatbevitel

A termék működtetéséhez az egér helyett a billentyűzet is használható.

A menüpontok és vezérlőelemek gyorsbillentyűket biztosítanak, amelyek segítségével a menüpontok és vezérlőelemek a billentyűzetről is kiválaszthatók. A gyorsbillentyűket a menüpontokban és vezérlőelemekben aláhúzás jelöli.

#### Billentyűzet fókusz

A Windows alapú rendszereken a billentyűzet beviteli pontja ki van emelve, jelezve az ablak aktív területét, illetve a billentyűlételek érvényesülésének helyét.

#### Válaszidő beállítások

A Windows alapú rendszereken a Vezérlőpulton beállíthatók a válaszidők.

---

### Kisegítő megjelenítőkkel kapcsolatos lehetőségek

Az ügyfelek több szolgáltatást is biztosítanak a csökkent látású felhasználók számára a felület kibővítéséhez. Ilyen bővítések például a nagy kontrasztú beállítások és a testre szabható betűtípus tulajdonságok.

#### Nagy kontrasztú mód

Az ügyfelek támogatják az operációs rendszer által biztosított nagy kontrasztú módot. Ez a szolgáltatás erősebb kontrasztot biztosít az előtér- és háttérszínek között.

#### Betűtípus beállítások

A Windows alapú rendszereken beállíthatja a menük és párbeszédablakok megjelenítéséhez használt betűtípusokat, illetve ezek színét és méretét. Az ügyfél lehetővé teszi a dokumentumlisták betűtípusának beállítását.

#### Színfüggetlenség

A termék funkcióinak használatához nincs szükség színek megkülönböztetésének képességére.

---

### Kompatibilitás a kisegítő technológiákkal

Az ügyfelek kompatibilisek a képernyő felolvasó alkalmazásokkal, például a Narrator és Via Voice termékekkel. Az ügyfelek rendelkeznek a kisegítő alkalmazások által a megjelenő információk vizuálisan korlátozott felhasználók számára történő elérhetővé tételéhez megkövetelt tulajdonságokkal.

---

## Hozzáférhető dokumentáció

A termékdokumentáció PDF formátumban áll rendelkezésre. A PDF fájlok HTML vagy szöveges fájlakká alakíthatók az [access.adobe.com](https://access.adobe.com) címen ingyenesen biztosított segédeszközökkel. Ez lehetővé teszi a felhasználóknak, hogy a dokumentációkat a böngészők megjelenítési beállításai szerint jelenítsék meg. Emellett biztosítja a képernyő felolvasó és egyéb kisegítő technológiák használatát.

---

## Megjegyzések

Ezek az információk az Egyesült Államokban forgalmazott termékekre és szolgáltatásokra vonatkoznak.

Elképzelhető, hogy a dokumentumban szereplő termékeket, szolgáltatásokat vagy lehetőségeket az IBM más országokban nem forgalmazza. Az adott országokban rendelkezésre álló termékekről és szolgáltatásokról a helyi IBM képviselők szolgálnak felvilágosítással. Az IBM termékekre, programokra vagy szolgáltatásokra vonatkozó utalások sem állítani, sem sugallni nem kívánják, hogy az adott helyzetben csak az IBM termékeit, programjait vagy szolgáltatásait lehet alkalmazni. Minden olyan működésében azonos termék, program vagy szolgáltatás alkalmazható, amely nem sérti az IBM szellemi tulajdonjogát. A nem IBM termékek, programok és szolgáltatások működésének megítélése és ellenőrzése természetesen a felhasználó felelőssége.

A dokumentum tartalmával kapcsolatban az IBM-nek lehetnek bejegyzett, vagy bejegyzés alatt álló szabadalmai. Jelen dokumentum nem ad semmiféle jogos licencet ezen szabadalmakhoz. A licenckérelmeket írásban a következő címre küldheti:

IBM Director of Licensing  
IBM Corporation  
North Castle Drive  
Armonk, NY 10504-1785  
U.S.A.

Ha duplabyte-os (DBCS) információkkal kapcsolatban van szüksége licencre, akkor lépjen kapcsolatban az országában az IBM szellemi tulajdon osztályával, vagy írjon a következő címre:

IBM World Trade Asia Corporation  
Licensing  
2-31 Roppongi 3-chome, Minato-ku  
Tokyo 106, Japan

**A következő bekezdés nem vonatkozik az Egyesült Királyságra, valamint azokra az országokra, amelyeknek jogi szabályozása ellentétes a bekezdés tartalmával:** AZ INTERNATIONAL BUSINESS MACHINES CORPORATION JELEN KIADVÁNYT "ÖNMAGÁBAN", BÁRMIFÉLE KIFEJEZETT VAGY VÉLELMEZETT GARANCIA NÉLKÜL ADJA KÖZRE, IDEÉRTVE, DE NEM KIZÁRÓLAG A JOGSÉRTÉS KIZÁRÁSÁRA, A KERESKEDELMI ÉRTÉKESÍTHETŐSÉGRE ÉS BIZONYOS CÉLRA VALÓ ALKALMASSÁGRA VONATKOZÓ VÉLELMEZETT GARANCIÁKAT. Bizonyos államok nem engedélyezik egyes tranzakciók kifejezett vagy vélelmezett garanciáinak kizárását, így elképzelhető, hogy az előző bekezdés Önre nem vonatkozik.

Jelen dokumentum tartalmazhat technikai, illetve szerkesztési hibákat. Az itt található információk bizonyos időnként módosításra kerülnek; a módosításokat a kiadvány új kiadásai tartalmazzák. Az IBM mindennemű értesítés nélkül fejlesztheti és/vagy módosíthatja a kiadványban tárgyalt termékeket és/vagy programokat.

A kiadványban a nem IBM webhelyek megjelenése csak kényelmi célokat szolgál, és semmilyen módon nem jelenti ezen webhelyek előnyben részesítését másokhoz képest. Az ilyen webhelyeken található anyagok nem képezik az adott IBM termék dokumentációjának részét, így ezek használata csak saját felelősségre történhet.

Az IBM belátása szerint bármilyen formában felhasználhatja és továbbadhatja a felhasználóktól származó adatokat anélkül, hogy a felhasználó felé ebből bármilyen kötelezettsége származna.

A programlicenc azon birtokosainak, akik információkat kívánnak szerezni a programról (i) a függetlenül létrehozott programok vagy más programok (beleértve ezt a programot is) közti információcseréhez, illetve (ii) a kicserélt információk kölcsönös használatához, fel kell venniük a kapcsolatot az alábbi címmel:

IBM Corporation  
J46A/G4  
555 Bailey Avenue  
San Jose, CA 95141-1003  
U.S.A.

Az ilyen információk bizonyos feltételek és kikötések mellett állnak rendelkezésre, ideértve azokat az eseteket is, amikor ez díjfizetéssel jár.

A dokumentumban tárgyalt licencprogramokat és a hozzájuk tartozó licenc anyagokat az IBM az IBM Vásárlói megállapodás, vagy a felek azonos tartalmú megállapodása alapján biztosítja.

A dokumentumban található teljesítményadatok ellenőrzött környezetben kerültek meghatározásra. Ennek következtében a más működési körülmények között kapott adatok jelentősen különbözhetnek a dokumentumban megadottaktól. Egyes mérések fejlesztői szintű rendszereken kerültek végrehajtásra, így nincs garancia arra, hogy ezek a mérések azonosak az általánosan hozzáférhető rendszerek esetében is. Továbbá bizonyos mérések következtetés útján kerültek becslésre. A tényleges értékek eltérhetnek. A dokumentum felhasználóinak ellenőrizni kell az adatok alkalmazhatóságát az adott környezetben.

A nem IBM termékekre vonatkozó információk a termékek szállítóitól, illetve azok publikált dokumentációiból, valamint egyéb nyilvánosan hozzáférhető forrásokból származnak. Az IBM nem tesztelte ezeket a termékeket, így az IBM a nem IBM termékek esetében nem tudja megerősíteni a teljesítményre és kompatibilitásra vonatkozó, valamint az egyéb állítások pontosságát. A nem IBM termékekkel kapcsolatos kérdéseivel forduljon az adott termék szállítóhoz.

Az IBM jövőbeli tevékenységére vagy szándékaira vonatkozó állításokat az IBM mindennemű értesítés nélkül módosíthatja, azok csak célokat jelentenek.

Az információk között példaként napi üzleti tevékenységekhez kapcsolódó jelentések és adatok lehetnek. A valóságot a lehető legjobban megközelítő illusztráláshoz a példákban egyének, vállalatok, márkák és termékek nevei szerepelnek. Minden ilyen név a képzelet szüleménye, és valódi üzleti vállalkozások neveivel és címeivel való bármilyen hasonlóságuk teljes egészében a véletlen műve.

#### SZERZŐI JOGI LICENC:

A könyv forrásnyelvi alkalmazásokat tartalmaz, amelyek a programozási technikák bemutatására szolgálnak a különböző működési környezetekben. A példaprogramokat tetszőleges formában, az IBM-nek való díjfizetés nélkül másolhatja, módosíthatja és terjesztheti fejlesztési, használati, eladási vagy a példaprogram operációs rendszer alkalmazásprogram illesztőjének megfelelő alkalmazásprogram terjesztési céllal. Ezek a példák nem kerültek minden körülmények között tesztelésre. Az IBM így nem tudja garantálni a megbízhatóságukat, javíthatóságukat vagy a program funkcióit. A példaprogramokat tetszőleges formában, az IBM-nek való díjfizetés nélkül másolhatja,



módosíthatja és terjesztheti fejlesztési, használati, eladási vagy az IBM alkalmazásprogram illesztőjének megfelelő alkalmazásprogram terjesztési céllal.

---

## Védjegyek

Az alábbi kifejezések az International Business Machines Corporation védjegyei az Egyesült Államokban és/vagy más országokban:

IBM	DisplayWrite	PowerPC
400	e-business	PTX
Advanced Peer-to-Peer Networking	HotMedia	QBIC
AIX	Hummingbird	RS/6000
AIXwindows	ImagePlus	SecureWay
APPN	IMS	SP
AS/400	Micro Channel	VideoCharger
C Set ++	MQSeries	Visual Warehouse
CICS	MVS/ESA	VisualAge
DATABASE 2	NetView	VisualInfo
DataJoiner	OS/2	WebSphere
DB2	OS/390	
DB2 Universal Database	PAL	

Az Approach, a Domino, a Lotus, a Lotus 1-2-3, a Lotus Notes és a SmartSuite a Lotus Development Corporation védjegye vagy bejegyzett védjegye az Egyesült Államokban és/vagy más országokban.

Az Intel és a Pentium az Intel Corporation védjegye vagy bejegyzett védjegye az Egyesült Államokban és/vagy más országokban.

A Microsoft, a Windows és a Windows NT a Microsoft Corporation bejegyzett védjegye az Egyesült Államokban és/vagy más országokban.

A Java és minden Java alapú védjegy és logó a Sun Microsystems, Inc. védjegye vagy bejegyzett védjegye az Egyesült Államokban és/vagy más országokban.

Az UNIX a The Open Group bejegyzett védjegye az Egyesült Államokban és/vagy más országokban.

Más cégek, termékek és szolgáltatások nevei mások védjegyei vagy szolgáltatásjegyei lehetnek.



---

## Szószedet

A szószedet a rendszerrel kapcsolatos fogalmakat és rövidítéseket írja le. A *dőlt* betűs kifejezések meghatározására a szójegyzék más részén kerül sor.

### A

**absztrakt osztály.** Olyan objektumorientált programozási *osztály*, amely egy alapvető képvisel, és amelynek a származtatásai jelentik az alapvető megvalósításait. Absztrakt osztályból nem készíthető objektum, vagyis ezek alapján nem hozhatók létre példányok.

**adatsere.** Képek importálása vagy exportálása a Content Manager ImagePlus for OS/390 rendszer indexeivel egy másik ImagePlus rendszerre *általános adatsere fájl* vagy *általános adatsere egység* használatával.

**adatformatum.** Lásd: *MIME típus*.

**adattároló.** (1) Általános fogalom adatok tárolására szolgáló helyekre (például: adatbázisrendszer, fájl vagy könyvtár). (2) Az alkalmazásprogramok szempontjából egy *tartalomkiszolgáló* virtuális megjelenése.

**ADSM.** Lásd: *Tivoli tárolókezelő*.

**alkalmazásprogram illesztő (API).** Olyan szoftveres illesztő, amely lehetővé teszi az alkalmazások közti kommunikációt. Az API programnyelvi szerkezetekből vagy utasításokból összeállított készlet, amely az alkalmazásokba kódolva bizonyos funkciókat vagy szolgáltatásokat vehet igénybe más licenccprogramoktól.

**állandó azonosító (PID).** Az *objektumokat* a tárolásuk helyétől függetlenül meghatározó egyedi azonosító. A PID az elem azonosítójából és helyéből áll.

**alosztály.** Egy másik osztályból származtatott *osztály*. Az osztály és az alosztály között több osztály is lehet.

**általános adatsere egység (CIU).** Az általános adatsere fájl (CIF) átviteltől független egysége. A fogadó adatbázishoz való viszonyt meghatározó CIF része. Egy CIF több CIU-t tartalmazhat.

**általános adatsere fájl (CIF).** Egy ImagePlus Interchange Architecture (IPIA) adatfolyamot tartalmazó fájl.

**API.** Lásd: *alkalmazásprogram illesztő*.

**átfedés.** Előre meghatározott adatok, például vonalak, árnyalások, szöveg, mezők vagy logók olyan gyűjteménye, amely a nyomtatás során az oldalon változó adatokkal kerül össze.

**átjáró.** Két, eltérő hálózati architektúrát alkalmazó számítógépes hálózatot összekapcsoló funkcionális egység. Az

átjáró különböző architektúrára épülő hálózatokat vagy rendszereket kapcsol össze. A híd azonos vagy hasonló architektúrájú hálózatok vagy rendszerek összekapcsolását végzi.

**Átlapolt audio/video (AVI).** Olyan *Erőforráscsere fájlformátum* (RIFF) fájlmeghatározás, amely lehetővé teszi az audio- és videojelek átlapolását egy fájlban belül. Az egyes sávok váltakozó szeletekben érhetők el lejátszás és rögzítés esetén, fenntartva a fájllezköz sorrendi elérését.

**attribútum.** Olyan adategység, amely az elemek egy bizonyos jellemzőjét vagy tulajdonságát (például nevét, címét, korát, stb.) írja le, és amely felhasználható az adott elem kikeresésére. Az attribútumok rendelkeznek egy típussal, amely az attribútum által megadott információk értelmezési tartományát jelenti, illetve rendelkeznek egy értékkel, amely az előbbi értelmezési tartományba esik. Egy multimédia fájlrendszerben egy fájlra vonatkozó információk például a cím, a futási idő vagy a kódolás típusa (MPEG1, H.263, stb.). Az Enterprise Information Portal esetén lásd még: *egyesített attribútum* és *eredeti attribútum*.

**AVI.** Lásd: *Átlapolt audio/video*.

**azonosító.** Egy objektumot ábrázoló karaktersorozat, amelyet a rendszer az objektum visszakereséséhez használ.

### B

**bináris nagy objektum (BLOB).** Olyan byte-sorozat, amelynek mérete 0 byte és 2 GB között lehet. A karaktersorozathoz nincs társított kódlap és karakterkészlet. A kép, audio és video objektumokat BLOB-ok tárolják.

**BLOB.** Lásd: *bináris nagy objektum*.

### C

**CGI.** Lásd: *Common Gateway Interface*.

**CGI parancsfájl.** Olyan számítógépes program, amely egy webkiszolgálón futva a *Common Gateway Interface (CGI)* használatával olyan feladatokat hajt végre, amelyeket a webkiszolgálók általában nem végeznek (például adatbázisok kezelése, úrlapok feldolgozása). A CGI parancsfájlok olyan CGI programok, amelyek valamilyen parancsfájl nyelven, például Perlben íródtak.

**CIF.** Lásd: *általános adatsere fájl*.

**CIU.** Lásd: *általános adatsere egység*.

**Common Gateway Interface (CGI).** A webkiszolgálók, és a hozzájuk képest külső programok közti információcsere egyik szabványa. A külső programok bármely programozási nyelven

megírhatók, amelyet a webkiszolgálót futtató operációs rendszer támogat. Lásd: *CGI parancsfájl*.

**csatolás.** Két *elem*, a forrás és a cél közötti irányított viszony. Több csatolás segítségével 1-N típusú társítások modellezhetők. Vesd össze: *hivatkozás*.

**csomag.** Hozzáférés felügyeletet és névtér kezelést biztosító, egymással kapcsolatban álló *osztályok* és illesztők gyűjteménye.

## D

**DDO.** Lásd: *dinamikus adatobjektum*.

### **dinamikus adatobjektum (DDO).**

Alkalmazásprogramokban a tárolt objektumok általános ábrázolása, amelyet a rendszer a tárolóba helyezéshez, illetve onnan eltávolításhoz használ.

**dokumentum.** Tárolható, visszakereshető, és a Content Manager rendszerek és felhasználók között külön egységként cserélhető *elem*. A dokumentum *szemantikai típusával* rendelkező elemek elvileg a dokumentumot alkotó információkat tartalmazzák, de nem feltétlenül utalnak rá, hogy ez a Content Manager dokumentummodell megvalósítása.

A dokumentum besorolással rendelkező elemtípusokból (a Content Manager dokumentummodell egy adott megvalósítása) létrehozott elemeknek tartalmazniuk kell dokumentumrészeket. A dokumentumként besorolt elemtípusokból dokumentum vagy mappa szemantikai típusú elemek hozhatók létre.

A dokumentumrészek többféle tartalomtípust is magukban foglalhatnak, például szöveget, képeket és táblázatokat.

**dokumentumtípus meghatározás (DTD).** Egy adott osztályba tartozó XML dokumentumok felépítését meghatározó szabályok. A DTD a szerkezetet elemekkel, attribútumokkal és jelölésekkel határozza meg, továbbá megadja az adott dokumentumosztály egyes elemeinek, attribútumainak és jelöléseinek felhasználására vonatkozó korlátozásokat. A DTD analóg az adatbázis sémákkal abban az értelemben, hogy a DTD teljesen meghatározza egy adott leírónyelv szerkezetét.

**DTD.** Lásd: *dokumentumtípus meghatározás*.

## E

**egyesített adattároló.** Tetszőleges számú *tartalomkiszolgáló* (például: Content Manager) virtuális megjelenési formája.

**egyesített attribútum.** Olyan Enterprise Information Portal metaadat kategóriák, amelyek egy vagy több *tartalomkiszolgáló eredeti attribútumaira* vannak leképezve. Például a *kötvényszám* egyesített attribútumot le lehet képezni a *kötvény\_sz* attribútumra a Content Manager rendszeren, és a *kötvényazonosító* attribútumra a Content Manager ImagePlus for OS/390 kiszolgálón.

**egyesített gyűjtemény.** Az *egyesített keresés* eredményeként kapott objektumcsoport.

**egyesített keresés.** Az Enterprise Information Portal termékből kiadott lekérdezés, amely egyszerre egy vagy több *tartalomkiszolgálón* keres adatokat.

**egyesített példány.** Olyan Enterprise Information Portal metaadat objektum, amely *egyesített attribútumokból* áll, és egy vagy több *egyesített szövegindex* rendelhető hozzá.

**egyesített szövegindex.** Olyan Enterprise Information Portal metaadat objektum, amely legalább egy *tartalomkiszolgáló eredeti szövegindexeire* van leképezve.

**egységes erőforrás helymeghatározás (URL).** Egy számítógép vagy hálózat, például az internet információs erőforrásait ábrázoló karaktersorozat. A karaktersorozat tartalmazza az információs erőforrás eléréséhez szükséges protokollt, illetve a protokoll által az információs erőforrás helyének meghatározásához használt információkat. Az Internet szöveggörnyezetében például a különféle információs erőforrások eléréséhez használt protokollok rövidített nevei a következők: http, ftp, gopher, telnet és news.

**egyszerűsített ügyfél.** Olyan ügyfél, amelyre kevés szoftver van telepítve, vagy egyáltalán nem rendelkezik telepített szoftverrel, de hozzáfér a hozzá csatlakozó hálózati kiszolgálók által kezelt és biztosított szoftverekhez. Az egyszerűsített ügyfél az összes funkciót tartalmazó ügyfél, például egy munkaállomás alternatívája.

**elem.** Általános kifejezés az Enterprise Information Portal által felügyelt információk legkisebb egységére. Minden elem rendelkezik egy azonosítóval. Elem lehet például egy *mappa* vagy egy *dokumentum*.

**elemtípus.** A hasonló *elemek* meghatározásához és későbbi kikereséséhez használt sablon, amely egy *gyökér összetevőből*, nulla vagy több *leszármazott összetevőből* és egy besorolásból áll.

**elemtípus besorolás.** Az *elemtípuson* belüli kategorizálás, amely részletesebben azonosítja az adott elemtípushoz tartozó *elemeket*. Egy elemtípus valamennyi eleme azonos elemtípus besorolással rendelkezik.

A Content Manager rendszer a következő elemtípus besorolásokat biztosítja: *mappa*, *dokumentum*, objektum, video, kép és szöveg. Ezek mellett a felhasználók saját elemtípus besorolásokat is meghatározhatnak.

**eredeti attribútum.** Egy adott *tartalomkiszolgálón* kezelt objektum egy olyan jellemzője, amely az adott tartalomkiszolgálóra jellemző. A *kötvényszám kulcsmező* például egy Content Manager tartalomkiszolgáló eredeti attribútuma lehet, míg a *kötvény\_azon* egy Content Manager OnDemand tartalomkiszolgálóé.

**eredeti példány.** Egy *tartalomkiszolgálón* kezelt objektum, amely *eredeti attribútumokból* áll. A Content Manager *indexosztályok* például Content Manager *kulcsmezőkből* összeálló eredeti példányok.

**eredeti szövegindex.** Egy adott *tartalomkiszolgálón* kezelt szöveges *elemek* indexe. Ilyen például egy Content Manager tartalomkiszolgáló egy adott szöveg keresési indexe.

**Erőforráscsere fájlformátum (RIFF).** Más típusú számítógépes berendezéseken való lejátszásra vagy megjelenítésre tervezett hang és grafikai tárolási formátum.

**erőforráskezelő.** A Content Manager rendszernek az *objektumokat* kezelő összetevője. Ezekre az objektumokra a *könyvtárkiszolgálón* tárolt *elemek* hivatkoznak.

**erőforráskezelő ideiglenes tároló.** Az *erőforráskezelő* munkaterülete. *Megjelenítési területnek* is hívják.

## F

**fájlrendszer.** A merevlemezek tárolóterületre particionálásának módja az AIX rendszereken.

**feladat.** A korábbi Content Manager munkafolyamatoknál és Enterprise Information Portal speciális munkafolyamatoknál bármilyen aktív tevékenység a *munkafolyamaton* belül.

**feladat állapot.** Egy egyedi *feladat*, *dokumentum* vagy *mappa* állapota.

**feladatcsoport.** Az Enterprise Information Portal 7.1 terméknél egyik helyről a másik helyre irányított *dokumentum* gyűjtemény. A felhasználók a *feladatlistákon* keresztül férnek hozzá a feladatcsoportokhoz.

**feladatlista.** Egy felhasználóhoz rendelt *feladatok*, *dokumentumok* vagy *mappák* összessége.

**felfüggesztés.** Egy *objektum* eltávolítása a *munkafolyamatból*, és az aktiválásához szükséges felfüggesztési feltételek meghatározása. Az objektum aktiválása lehetővé teszi a feldolgozásának folytatását.

**felhasználó.** Az Enterprise Information Portal esetén bárki, akit az Enterprise Information Portal adminisztrációs program azonosít.

**felhasználó leképezés.** Az Enterprise Information Portal felhasználói azonosítók és jelszavak társítása a tartalomkiszolgálók megfelelő felhasználói azonosítóihoz és jelszavaihoz. A felhasználó leképezés egységes bejelentkezést tesz lehetővé az Enterprise Information Portal rendszerbe és több *tartalomkiszolgálóra*.

**felhasználói csoport.** Legalább egy meghatározott egyéni *felhasználóból* álló csoport, amelyet egy csoportnév azonosít.

**felhasználói kilépési pont.** Olyan pont az IBM programokban, ahol egy felhasználói kilépési rutin veheti át a vezérlést.

**felhasználói kilépési rutin.** A felhasználók által írt olyan rutin, amely a előre meghatározott *felhasználói kilépési pontokban* átveszi a vezérlést.

**felszabadítás.** A felfüggesztési feltételek eltávolítása egy *elemről*. A felfüggesztett elemek felszabadítására akkor kerül sor, amikor teljesülnek a felszabadítási feltételek, vagy egy megfelelő jogosultsággal rendelkező felhasználó felülbírálja a feltételeket, és manuálisan felszabadítja azokat.

**folyam adatok.** Egy hálózati kapcsolaton a megadott sebességgel küldött adatok. Az adatfolyam egy típusból vagy különféle típusok kombinációjából állhat. A bit/másodpercenként megadott adatsebességek a különféle típusú folyamatok és hálózatok szerint változnak.

**folyamat állapota.** A teljes *munkafolyamat* állapota.

## H

**hálózati tábla fájl.** A Content Manager rendszer valamennyi csomópontjának rendszerre jellemző konfigurációs információit tartalmazó szöveges fájl. A rendszer minden csomópontjának rendelkeznie kell egy hálózati tábla fájjal, amely azonosítja a csomópontot, és felsorolja azokat a csomópontokat, amelyekhez csatlakoznia kell.

A hálózati tábla neve FRNOLINT.TBL.

**helyettesítő karakter.** Egy vagy több karakter ábrázolására használható speciális karakter, például egy csillag (\*) vagy kérdőjel (?). A helyettesítő karakterek tetszőleges karaktert vagy karaktersorozatot jelenthetnek.

**helyi hálózat (LAN).** Olyan hálózat, amelyben az eszközök kommunikációs céllal össze vannak kötve egymással, és amely nagyobb hálózathoz csatlakoztatható.

**Hiperszöveg leírónyelv (HTML).** Az SGML szabványnak megfelelő leírónyelv, amelyet elsősorban hiperhivatkozásokat tartalmazó szöveges és képi információk online megjelenítésének támogatására fejlesztettek ki.

**hivatkozás.** Egyirányú, 1-1 típusú társítás egy gyökér vagy *leszármazott összetevő* és egy másik *gyökér összetevő* között. Vesd össze: *csatolás*.

**hozzáférés felügyelet.** Az a folyamat, amely biztosítja, hogy bizonyos funkciókhoz és tárolt *objektumokhoz* csak a feljogosított felhasználók férhessenek hozzá, és csak a megadott módon.

**hozzáférés felügyeleti lista.** Legalább egy felhasználói azonosítóból vagy felhasználói csoportból, és a hozzájuk tartozó *jogosultságokból* álló lista. A hozzáférési listákkal lehet felügyelni az Enterprise Information Portal rendszer *keresési sablonjaira* vonatkozó hozzáférést.

**HTML.** Lásd: *Hiperszöveg leírónyelv*.

## I

**ideiglenes tároló.** A főtárnál kisebb és gyorsabb speciális célú puffer, amely a gyakran használt adatok másolatát tárolja. Az ideiglenes tároló használata lecsökkentheti a hozzáférési időt, viszont magasabb memóriakövetelményeket támaszt.

**Image Object Content Architecture (IOCA).** Képek cseréjéhez és megjelenítéséhez használt szerkezetgyűjtemény.

**index.** Egy adott *elemet* vagy *objektumot* a későbbi visszakeresés céljából azonosító attribútumértékek.

**indexosztály.** Lásd: *elemtípus*.

**indexosztály nézet.** A Content Manager rendszerek korábbi változataiban az *indexosztály* *részalmazatokra* az API-k esetén használt kifejezés.

**indexosztály részalmaz.** A Content Manager rendszerek korábbi változataiban az alkalmazások által a mappák és objektumok tárolásához, visszakereséséhez és megjelenítéséhez használt *indexosztályok* egy nézete.

**információbányászat.** Automatizált folyamat, a kulcsinformációk kinyerése a szövegekből (összegzés), gyakori témák keresése a dokumentum gyűjteményekben (kategorizálás) és fontos dokumentumok keresése hatékony, rugalmas lekérdezésekkel.

**IOCA.** Lásd: *Image Object Content Architecture*.

**iterátor.** Olyan osztály vagy szerkezet, amelyen a dokumentumgyűjtemények tagjai egyesével haladnak át.

## J

**JavaBeans.** Platformfüggetlen szoftverösszetevő technológia többször felhasználható Java összetevők, úgynevezett "beanek" írására. A JavaBeaneket az összeállítás után más szoftverfejlesztők illetve Java alkalmazások is használhatják. A JavaBeanek használatával a szoftverfejlesztők a beaneket grafikus fogd-és-vidd fejlesztői környezetben állíthatják össze.

**jellemző.** Kép keresési kiszolgálón tárolt vizuális tartalominformációk. Vizuális vonalakat is jelent, amelyeket az alkalmazások az egyezések meghatározásához használnak. A négy *QBIC* jellemző: szín, hisztogram szín, pozíció szín és kitöltés.

**jogosultság.** Egy adott *objektum* egy adott módon végrehajtott elérésére vonatkozó jog. A jogosultságokba egyebek között a rendszeren tárolt objektumok létrehozása, törlése és kiválasztása tartozik bele. A jogosultságokat az adminisztrátor osztja ki.

**jogosultságkészlet.** A rendszer összetevőinek és funkcióinak kezelésére vonatkozó *jogosultságok* gyűjteménye. A jogosultságkészleteket az adminisztrátor osztja ki a felhasználóknak (felhasználói azonosítóknak) és *felhasználói csoportoknak*.

**Joint Photographic Experts Group (JPEG).** (1) A digitalizált folytonos színezésű képek tömörítési szabványát kidolgozó munkacsoport. (2) A csoport által megalkotott állókép szabvány.

**JPEG.** Lásd: *Joint Photographic Experts Group*.

## K

**kapcsolati osztály.** Olyan objektumorientált programozási *osztály*, amely szabványos hozzáférést biztosít az egyes *tartalomkiszolgálók* egyedi alkalmazásprogram illesztőjéhez (API).

**keresési feltétel.** Az Enterprise Information Portal esetén bizonyos mezők, amelyeket az adminisztrátor beállít egy *keresési sablonhoz*, így korlátozva vagy pontosabban meghatározva a *felhasználók* számára rendelkezésre álló választási lehetőségeket.

**keresési sablon.** Egy *keresési feltételekből* álló űrlap, amelyet az adminisztrátor hoz létre egy adott típusú egyesített kereséshez. Az adminisztrátor határozza meg a keresési sablon elérésére feljogosított *felhasználókat* és *felhasználói csoportokat* is.

**kiszolgáló tároló.** A megadott *tartalomkiszolgálók eredeti példányainak* és *eredeti attribútumainak* összefoglaló listája.

**kiszolgálómeghatározás.** Egy adott *tartalomkiszolgálót* az Enterprise Information Portal rendszerben egyedi módon azonosító jellemzők.

**kiszolgálótípus meghatározás.** Az adminisztrátor által megadott jellemzők azon listája, amely az adott típusú egyéni kiszolgálók egyedi azonosításához szükséges az Enterprise Information Portal rendszerben.

**kiterjesztett adatobjektum (XDO).**

Alkalmazásprogramokban a tárolt összetett multimédiás *objektumok* általános ábrázolása, amelyet a rendszer a tárolóba helyezéshez, illetve onnan eltávolításhoz használ. Az XDO-kat leggyakrabban *DDO*-k tartalmazzák.

**Kiterjeszthető leírónyelv (XML).** SGML származtatású illetve SGML részalmaz leírónyelvek meghatározására szolgáló szabványos metanyelv. Az XML kihagyja az SGML összetettebb és kevésbé használt részeit, így sokkal könnyebbé teszi a dokumentumtípusokat kezelő alkalmazások írását, a strukturált információk létrehozását és kezelését, illetve a strukturált információk átvitelét és megosztását különféle számítástechnikai rendszerek között. Az XML használata nem követeli meg az SGML feldolgozáshoz szükséges nagy méretű alkalmazásokat és összetett feldolgozást. Az XML fejlesztése a World Wide Web Consortium (W3C) védnöksége alatt történik.

**kombinált keresés.** *Paraméteres*, szöveg és kép kereséseket kombináló lekérdezés.

**konstruktor.** A programozási nyelvekben az osztály nevével megegyező nevű, az osztály objektumainak létrehozására és inicializálására használt metódus.

**könyvtár ügyfél.** A Content Manager rendszernek az az összetevője, amely alacsonyszintű programozási illesztőt biztosít a könyvtár rendszerhez. A könyvtár ügyfél olyan API-t tartalmaz, amelyek a szoftver fejlesztőkészlet részei.



**könyvtárkiszolgáló.** A Content Manager rendszernek az az összetevője, amely az *elemek* tárolását, kezelését és lekérdezéseinek felügyeletét látja el.

**kötet.** Tényleges fizikai tárolóeszköz vagy egység ábrázolása, amelyen a rendszer objektumai tárolhatók.

**kulcsmező.** Lásd: *attribútum*.

**kurzor.** Megnevezett vezérlési szerkezet, amelyet az alkalmazásprogramok egy több sorból álló halmaz egy adott sorának megjelölésére használnak. A kurzor a halmaz sorainak visszaadására szolgál.

## L

**LAN.** Lásd: *helyi hálózat*.

**lekérdezés képtartalom alapján (QBIC).** Olyan lekérdezési technológia, amely lehetővé teszi a jellemzőknek nevezett vizuális tartalom alapján végrehajtott kereséseket a sima szöveg helyett. A QBIC használatával az objektumokat vizuális jellemzőik, például színük vagy kitöltésük alapján lehet keresni.

**lekérdezési karaktersorozat.** A lekérdezés tulajdonságait és tulajdonságainak értékét meghatározó karaktersorozat. A lekérdezési karaktersorozatokat alkalmazásokban lehet létrehozni, majd át lehet adni azokat a lekérdezéseknek.

**leszármazott összetevő.** Egy *elemtípus* elhagyható második vagy alacsonyabb szintje. Minden leszármazott összetevő közvetlen hozzárendelésben van a felette álló szinttel.

## M

**mappa.** Tetszőleges *elemtípusú elem*, amely besorolásától függően mappa *szemantikai típussal* rendelkezik. A mappa szemantikai típusú elemek a nem erőforrás elemek képességei, illetve az elemtípus besorolás (például *dokumentum* vagy erőforrás elem) által biztosított képességek mellett a Content Manager által biztosított mappa funkcionális tartalmazzák. A mappák tetszőleges számú és típusú elemet tartalmazhatnak, beleértve a dokumentumokat és almappákat is. A mappák indexelése *attribútumok* alapján történik.

**mappakezelő.** Content Manager modell az adatok online dokumentumként és mappaként kezeléséhez. A mappakezelő API-kat elsődleges illesztőként használhatja az alkalmazások és a Content Manager tartalomkiszolgálók között.

**média archiváló.** Audio és video folyamatok tárolására használható fizikai eszköz. A VideoCharger a média archiválók egy típusa.

**médiakiszolgáló.** A Content Manager rendszernek egy olyan AIX alapú összetevője, amely video fájlok tárolására és elérésére használható.

**megjelenítés.** Az offline vagy alacsony prioritású eszközökön tárolt *objektumok* áthelyezésének folyamata egy online vagy magasabb prioritású eszközre, általában egy

felhasználó vagy a rendszer kérésére. Amikor egy felhasználó egy állandó tárolóban található objektumra vonatkozó kérést ad ki, akkor egy munkapéldány kerül kiírásra a *megjelenítési területen*.

**megjelenítési terület.** Az *erőforráskezelő* munkaterülete. Néhol *erőforráskezelő tárolónak* is nevezik.

**metódus.** Java tervezésnél vagy programozásnál egy műveletben meghatározott viselkedést megvalósító szoftver. Jelentése megegyezik a C++ tagfüggvényével.

**MIME típus.** Az Interneten átvitt objektumok típusának meghatározására szolgáló Internetes szabvány. A MIME típusok közé többféle audio, kép és video tartozik. Minden objektum rendelkezik MIME típussal.

**multimédia.** Különböző médiaelemek (szöveg, grafika, audio, állókép, video, animáció) számítógépes megjelenítési és vezérlési céllal történt egyesítése.

**multimédia fájlrendszer.** Video és audio adatok tárolására és kiszolgálására optimalizált *fájlrendszer*.

**munkafolyamat.** Az Enterprise Information Portal esetén *munkalépések* és a lépéseket irányító szabályok sorozata, amelyeken keresztül a *feladatcsoport* a feldolgozás során átvitelre kerül.

Például: a *kárigény elfogadás* azt a folyamatot írhatja le, amelyet az egyes kárigények elfogadásakor követni kell.

**munkalépés.** A *munkafolyamatnak* vagy *dokumentumtovábbítási folyamatnak* olyan diszkrét pontja, amelyen egy adott *feladatnak*, *dokumentumnak* vagy *mappának* át kell haladnia.

## O, Ö

**objektum.** Tetszőleges digitális tartalom, amelyet a felhasználó önálló egységként tárolhat, hívhat le vagy kezelhet. Ilyen például egy *JPEG* kép, egy MP3 audio, egy *AVI* video vagy egy könyv szöveges része.

**Objektumcsatolás és -beágyazás (OLE).** A Microsoft specifikációja az objektumok csatolására és beágyazására, amelyek így más alkalmazásokból is aktiválhatók.

**objektumkiszolgáló.** Lásd: *erőforráskezelő*.

**objektumkiszolgáló gyorsítótár.** Lásd: *erőforráskezelő gyorsítótár*.

**OLE.** Lásd: *Objektumcsatolás és -beágyazás*.

**osztály.** Objektumorientált tervezés vagy programozás esetén olyan modell vagy sablon, amelyből közös definícióval rendelkező objektumok példányai hozhatók létre. Az így létrehozott objektumok a közös definíciónak köszönhetően közös tulajdonságokkal, műveletekkel és viselkedéssel rendelkeznek. Az objektumok egy osztály példányai.

**összetevő.** A *gyökér összetevők* és *leszármazott összetevők* általános neve.

## P

**paraméteres keresés.** Olyan *objektum* lekérdezés, amely az objektumok *tulajdonságain* alapul.

**PID.** Lásd: *állandó azonosító*.

## Q

**QBIC.** Lásd: *lekérdezés képtartalom alapján*.

## R

**rang.** Egész érték, amely egy adott rész alkalmazhatóságát jelzi a keresés eredményeiben. A magasabb rangok közelebbi egyezést jelentenek.

**README fájl.** Olyan fájl, amelyet a hozzá tartozó program telepítése vagy használatba vétele előtt kell elolvasni. A README fájlok általában a legfrissebb termékinformációkat, telepítési útmutatásokat vagy használati tippeket tartalmaznak.

**renderelés.** Az a folyamat melynek során az általában nem képi alapú adatok képi megjelenítésre kerülnek. A Content Manager esetén például a szövegszerkesztő dokumentumai megjelenítési célokra képekké alakíthatók.

**rész.** Lásd: *objektum*.

**RIFF.** Lásd: *Erőforráscsere fájlformátum*.

**RMI kiszolgáló.** A Java *Távoli metódushívás (RMI)* osztott objektummodellt megvalósító kiszolgáló.

## S

**számosság.** Az adatbázistábla sorainak száma.

**szemantikai típus.** Egy *elem* használata vagy szabályai. Az alap, a feljegyzés és a megjegyzés a Content Manager rendszerek beépített szemantikai típusai, amelyek mellé a felhasználók saját szemantikai típusokat határozhatnak meg.

**szülőosztály.** Az *osztály*, amelyből egy osztály származtatásra került. Az osztály és a szülőosztály között több osztály is lehet.

## T

**tartalomkiszolgáló.** Multimédia és üzleti adatokat, valamint az ezek használatához szükséges kapcsolódó metaadatokat tároló szoftverrendszer. A Content Manager és a Content Manager ImagePlus for OS/390 például tartalomkiszolgálók.

**Távoli metódushívás (RMI).** Osztott programozást biztosító API készlet. A Java virtuális gép (JVM) objektumai más JVM-ek objektumaiban lévő metódusokat hívhatnak meg.

**tevékenységlista.** A felhasználó által a *folyamatokon* vagy dokumentumtovábbítási folyamatokban végrehajtható tevékenységeket tartalmazó lista, amelyet a munkafolyamat *koordinátora* határoz meg.

**Tivoli tárolókezelő (TSM).** Olyan *ügyfél/kiszolgáló* termék, amely tároláskezelési és adathozzáférési szolgáltatásokat nyújt heterogén környezetekben. Többféle kommunikációs módszert is támogat, adminisztrációs szolgáltatásokat tartalmaz a fájlok mentésének és tárolásának kezeléséhez, továbbá lehetővé teszi a mentési műveletek ütemezését.

**Többcélú Internetes levélkiterjesztések (MIME).** Lásd: *MIME típus*.

**történetnapló.** A *munkafolyamat* tevékenységeinek feljegyzéseit tároló fájl.

**TSM.** Lásd: *Tivoli tárolókezelő*.

**TSM kötet.** A *Tivoli tárolókezelő* által kezelt tár egy logikai területe.

**tulajdonság.** Az *objektumok* egy jellemzője, amely meghatározza az objektumot. A tulajdonságok módosíthatók és változtathatók. Ilyen tulajdonság például a stílus.

## U, Ü

**ügyfél/kiszolgáló.** Kommunikáció esetén az osztott adatfeldolgozási környezetek olyan együttműködési modellje, amelyben az egyik helyen található program egy másik helyen található programnak küldi el a kérését, majd a válaszra vár. A kérést kiadó programot ügyfélnek, a választ biztosító programot pedig kiszolgálónak hívják.

**ügyfélalkalmazás.** Az objektumorientált vagy Internet API-k felhasználásával az Enterprise Information Portal *tartalomkiszolgálók* elérésére írt alkalmazás.

## X

**XDO.** Lásd: *kiterjesztett adatobjektum*.

**XML.** Lásd: *Kiterjeszthető leírnyelv*.



# Tárgymutató

## A, Á

adminisztrációs tartomány 34  
adminisztrációs ügyfél  
létrehozás  
keresési feltétel 26  
keresési sablonok 25  
munkafolyamat 100  
meghatározás  
feladatlisták 99  
tevékenységek 100  
tevékenységlista 100  
adományozható jogosultságkészlet 34

## B

billentyűzet 129

## C

cmbcc2mime.ini 14

## E, É

EIP  
adminisztrációs összetevő 3  
információbányászat lehetőség 4  
Információs központ összetevő 6  
kapcsolatok 4  
kép keresési ügyfél 4  
szöveg keresési ügyfél 4  
tartalom megjelenítő ügyfél 5  
Web Crawler lehetőség 4  
EIP összetevők  
adminisztráció 3  
információbányászat 4  
Információs központ 6  
kapcsolatok 4  
kép keresés 4  
operációs rendszer kompatibilitás 3  
szöveg keresés 4  
tartalom megjelenítő 5  
Web bejáró 4  
Enterprise Information Portal  
létrehozás  
keresési feltétel 26  
keresési sablonok 25  
munkafolyamat 100  
meghatározás  
feladatlisták 99  
tevékenységek 100  
tevékenységlista 100  
erőforráskezelő  
felhasználók hozzárendelése 34  
hozzárendelés tartományhoz 37  
erőforráskezelő, áthelyezés tartományok  
között 39

## F

feladatcsoport, leírás 96  
feladatlista  
leírás 97  
meghatározás 99  
felhasználó 29  
áthelyezés tartományok között 38  
jogosultságkészlet 34  
felhasználói azonosító 29  
felhasználói csoport 34  
áthelyezés tartományok között 38  
fizikai kihívások 129  
folyamat tervező  
folyamat létrehozása 100  
indítás 100  
leírás 98

## GY

gyűjtemény  
hozzárendelés tartományhoz 37

## G

gyűjtemények  
áthelyezés tartományok között 39

## H

hozzáférés felügyeleti lista  
áthelyezés tartományok között 39

## I, Í

IBM Enterprise Information Portal for  
Multitplatforms  
összetevők 2  
indítás  
folyamat tervező 100  
Információbányászat  
célcsoport 44  
használat üzleti környezetben 44  
leírás 41  
osztályozás összeállítása 61  
összetevők 42  
példa 45  
szolgáltatás 41  
támogatott dokumentumformátumok 49  
támogatott nyelvek 49  
Információstruktúrálnálási eszköz  
betanítási dokumentumok kiválasztása 64  
kezdeti lépések 62  
leírás 61  
osztályozás betanítása 70  
osztályozás kiértékelése 67  
osztályozás meghatározása 62  
telepítés 61  
WebSphere alkalmazáskiszolgáló  
használat 61

Információstruktúrálnálási eszköz (Folytatás)  
zárolási mechanizmus 62

## J

jogosultsági csoport 33  
jogosultságkészlet 29, 33, 34  
áthelyezés tartományok között 39  
létrehozás 33

## K

kapcsolatok 4  
katalógus  
átnevezés 62  
betanítás 70  
betanítási dokumentumok hozzáadása 64  
hozzáadás 62  
kiértékelés 67  
törlés 62  
kép keresés  
álnév hozzárendelése 111  
beállítás 110  
kapcsolat ellenőrzése 112  
kép keresés beállítása 110  
kép keresési lehetőség 4  
keresési feltétel  
meghatározás és leképezés 26  
keresési sablonok, létrehozás 25  
kiszolgáló lehetőségek 129  
kiszolgáló tároló 17

## L

LDAP  
beállítás 32  
importálás 32

## M

metaadat tároló  
információbányászat használata 41  
MIME típus fájl  
módosítás kiszolgálókon 14  
MIME típusok testreszabása 14  
mintaadatok betöltése 112  
mintabetöltési program, futtatás 115  
munkafolyamat  
alapelvek 93  
létrehozás 100  
tervezés 96  
munkafolyamat szolgáltatás  
beállítás 93  
összetevők 98

## O, Ó

osztályozás

Információstrukturálási eszköz  
használata 61

## SZ

szöveg és kép keresési dokumentumok  
betöltése 115

## S

szöveg keresés

beállítás 109  
XML támogatás 109

## T

tartalom megjelenítő lehetőség 5

tartalomkiszolgáló  
meghatározás 17

tartomány 39

tartományok  
aladminisztrátor jogosultságok 36  
felsőszintű adminisztrátor  
jogosultságok 36  
létrehozás 35  
magyarázat 35

tervezés

Enterprise Information Portal 2

tevékenységek, meghatározás 100

tevékenységlista  
előre meghatározott tevékenységek 100  
meghatározás 97, 100

## W

Web bejáró

EIP lehetőség 4





Programszám: 5724-B43

Nyomtatva Dániában

SC22-5304-01

