

# E-mail-archiválási tutorial (Networkshop 2021, Kalcsó Gyula, Szűcs Kata Ágnes)

A tutorial első felében a résztvevők átfogó képet kapnak az e-mail-archiválás történetéről, a kialakult nemzetközi gyakorlatokról. Kitérünk az e-mail mint szabványos formátum alapkérdéseire. A [Digital Preservation Coalition](#), valamint az [CLIR Email Archives Task Force](#) legújabb összefoglalói alapján bemutatjuk az e-mailek kezelésének legfontosabb problémáit (kontextuális információk, csatolmányok, linkek, beágyazott, érzékeny tartalmak kezelése). A digitális archiválásban nemzetközi szinten is a legelterjedtebbnek számító [OAIS-modell](#) alapján számba vesszük az e-mailek kezelésének a lehetőségeit. Egy vadonatúj, az [Email Archives Task Force](#) által [kidolgozott ajánlásra](#) is kitérünk: újabban javasolják a PDF-konverziót mint az e-mailek kezelése körüli nehézségek áthidalásának egyik lehetőségét.

A tutorial második fele gyakorlati jellegű lesz. Alapvetően kétféle módszert mutatunk be az e-mail-archiválásra:

- Szotveres megoldás (kliensprogrammal): a [DArcMail](#) használata
- E-mail-konverzió és csomagolás (EML, MBOX, XML csomagolása OAIS-kompatibilis módon [RODA-Innel](#))

**A résztvevőket kérjük, hogy telepítsék fel a saját rendszerüknek megfelelő változatban az alábbi szoftvereket:**

**E-mail-archiváláshoz:**

**DArcMail.** <https://siarchives.si.edu/what-we-do/digital-curation/email-preservation-darcmail>  
<https://siarchives.si.edu/sites/default/files/DArcMail/DArcMail-v1.2-2018-03-07.zip>

Az első hivatkozásban megadott oldal alján található két link, az egyik a felhasználói kézikönyv PDF-ben, a másik maga a szoftver ZIP-ben (ez utóbbit fentebb külön is megadtuk):

## Current Approach

The Archives focuses its email collections development on the accounts of identified record creators. These include senior Institution managers, museum directors and some senior curators, however noteworthy individuals are found at many different levels of the Institution. Therefore, some selection of accounts occurs on a case-by-case basis. This approach has become more popularly known as Capstone. For more details, consult our functional records schedule.

## Challenges and Solutions

Email by its nature presents challenges to archival preservation, most notably: 1) the variety of email message formats; 2) message components hidden by default by the email client; 3) capturing the internal account organization, often nested, given to the email by the account owner; 4) capturing the interrelationships between messages and attachments within an account; and 5) variety of file formats of embedded attachments.

Very large scale presents another significant challenge. The nature of the Smithsonian's work and the depositors' length of time at the Institution results in accounts that range in volume from tens of thousands emails to, in some cases, half a million emails or more. Depending on any given year, we are likely to acquire roughly fifteen accounts. Many available preservation tools are unable to effectively address accounts of this size.

## DArcMail

Our most recent application was created to accomplish preservation and limited processing and access is DArcMail (*Digital Archives of eMail*) of one or more email accounts. Released in 2018, this open source tool is Python-based and uses either SQLite or MySQL on the backend. Consequently, a full DArcMail implementation no-cost, platform-agnostic and easy to maintain. DArcMail adds a graphical user interface, powerful search and filtering and reduced processing time to the CERP tool's preservation functionality. Read more about using [DArcMail](#) or download the [software](#).

PDF ↗ ZIP ↗

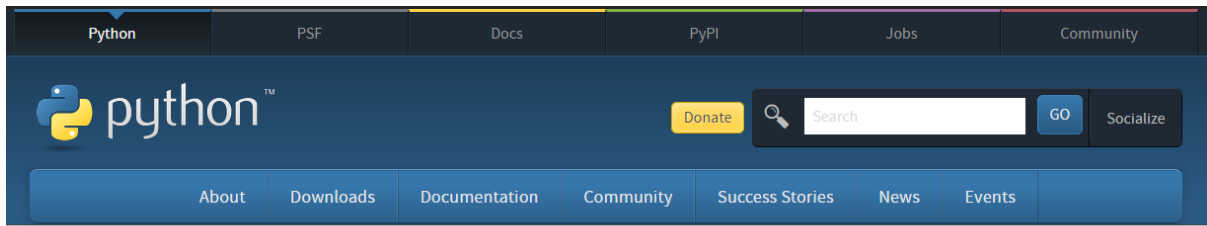
Sign up for email updates on our amazing collection

A ZIP-et letöltjük, tetszőleges helyre kicsomagoljuk, létrejön egy **DArcMail-v1.2-2018-03-07** nevű mappa, abban található a futtatható Python-állományok (kettőre lesz szükségünk: **DArcMail.py** és **DArcMailXml.py**).

Ahhoz, hogy ezeket futtatni tudjuk, fel kell telepítenünk a Python programnyelv 2.7.6-os változatát a rendszerünkre:

**Python 2.7.6.** <https://www.python.org/downloads/release/python-276/>

A letöltési oldalon ki kell választanunk a rendszerünknek megfelelő linket:



## Python 2.7.6

**Release Date:** Nov. 10, 2013

Note: Python 2.7.6 has been superseded by [Python 2.7.8](#).

Python 2.7.6 was released on November 10, 2013. This is a 2.7 series bugfix release. Most importantly, it resolves [an issue](#) that caused the interactive prompt to crash on OS X 10.9. It also includes [numerous bugfixes](#) over 2.7.5.

### Download

This is a production release. Please [report any bugs](#) you encounter.

We currently support these formats for download:

- [Windows x86 MSI Installer \(2.7.6\) \(sig\)](#)
- [Windows x86 MSI program database \(2.7.6\) \(sig\)](#)
- [Windows X86-64 MSI Installer \(2.7.6\) \[1\] \(sig\)](#)
- [Windows X86-64 MSI program database \(2.7.6\) \[1\] \(sig\)](#)
- [Windows help file \(sig\)](#)
- [Mac OS X 64-bit/32-bit x86-64/i386 Installer \(2.7.6\) for Mac OS X 10.6 and later \[2\] \(sig\)](#). [You may need an updated Tcl/Tk install to run IDLE or use Tkinter, see note 2 for instructions.]
- [Mac OS X 32-bit i386/PPC Installer \(2.7.6\) for Mac OS X 10.3 and later \[2\] \(sig\)](#).
- [XZ compressed source tar ball \(2.7.6\) \(sig\)](#)
- [Gzipped source tar ball \(2.7.6\) \(sig\)](#)

Töltsük le a telepítőfájlt, és futtassuk. Valószínűleg rendszergazdai jogosultságra lesz szükségünk. Ha más Python-verzió is van a gépünkön (pl. a 3-as verzió valamely változata), szükség lesz arra, hogy a DArcMail fájljait társítsuk ezzel a Python-verzióval:


A DArcMail mappájában a **DArcMail.py** fájlra jobb egérgombbal kattintva válasszuk a Társítás, majd az Alapértelmezett program választása vagy a Másik alkalmazás választása lehetőséget. A Tallózás vagy a További alkalmazások (szükség esetén a Másik alkalmazás keresése ezen a gépen) feliratra kattintva keressük meg a telepített Python 2.7. mappáját (**python27** lesz a neve), és jelöljük ki a **python.exe** alkalmazást, majd nyomjunk OK-t. Ezután a DArcMail fájljai már Python-logóval lesznek megjelölve, ha jól csináltunk mindent.

Szükség van még a Pythonhoz egy kényelmesebb grafikus felület telepítésére:

**wxPython 3.0.2.** <https://sourceforge.net/projects/wxpython/files/wxPython/3.0.2.0/>

Itt is ki kell választanunk a rendszerünknek megfelelő telepítőfájlt:

Home / Browse / Development / Frameworks / wxPython / Files



# wxPython

A wrapper for the cross-platform, GUI API toolkit wxWidgets  
Brought to you by: [robind](#)

Summary | **Files** | Reviews | Support | Wiki | Donate | Code

The interactive file manager requires Javascript. Please enable it or use [sftp](#) or [scp](#).  
You may still *browse* the files here.

[Download Latest Version](#)  
wxPython-pdb-4.0.0a3-py36-win\_amd64.zip (63.4 MB) [Get Updates](#)

[Home / wxPython / 3.0.2.0](#)

Name	Modified	Size	Downloads / Week
↕ Parent folder			
<a href="#">wxWidgets3.0-devel-win64-3.0.2.0.exe</a>	2014-11-28	86.8 MB	39
<a href="#">wxWidgets3.0-devel-win32-3.0.2.0.exe</a>	2014-11-28	84.6 MB	9
<b>64</b> <a href="#">wxPython3.0-win64-3.0.2.0-py27.exe</a>	2014-11-28	16.8 MB	123
<a href="#">wxPython3.0-win64-3.0.2.0-py26.exe</a>	2014-11-28	16.8 MB	3
<a href="#">wxPython3.0-win32-docs-demos-3.0.2.0.exe</a>	2014-11-28	39.4 MB	28
<b>32</b> <a href="#">wxPython3.0-win32-3.0.2.0-py27.exe</a>	2014-11-28	15.3 MB	186
<a href="#">wxPython3.0-win32-3.0.2.0-py26.exe</a>	2014-11-28	15.3 MB	1
<a href="#">wxPython3.0-osx-docs-demos-3.0.2.0-cocoa-py2.7.dmg</a>	2014-11-28	36.8 MB	2
<a href="#">wxPython3.0-osx-docs-demos-3.0.2.0-carbon-py2.7.dmg</a>	2014-11-28	36.5 MB	2

64 bites Windows esetén a képen piros 64-es felirattal megjelölt, 32 bites rendszer esetében pedig a piros 32-essel megjelölt telepítő lesz a megfelelő. Töltsük le, és futtassuk (rendszergazdai jogosultság szükséges). A telepítés során minden ablakban nyugodtan kattintsunk a Next feliratra, ill. a második ablakban fogadjuk el a License Agreement (I accept the agreement). Az utolsó ablakban mindenhol vegyük ki a pipát, és kattintsunk a Finish gombra.

Szükség van még egy olyan Python-komponensre, amely alkalmassá teszi a Python-t az adatbázissal való együttműködésre:

**MySQL Connector/Python** <https://downloads.mysql.com/archives/c-python/>

A Product Version legördülő listából válasszuk a **2.0.4**-et.

Válasszuk ki a rendszerünket az alatta található újabb legördülő listából (Operating System és OS Version):

## MySQL Product Archives

MySQL Connector/Python (Archived Versions)

**⚠ Please note that these are old versions. New releases will have recent bug fixes and features!**  
To download the latest release of MySQL Connector/Python, please visit [MySQL Downloads](#).

Product Version:   
Operating System:

Windows (Architecture Independent), MSI Installer Python 3.3 (mysql-connector-python-2.0.4-py3.3.msi)	Apr 1, 2015	139.0K	<a href="#">Download</a>
Windows (Architecture Independent), MSI Installer Python 3.4 (mysql-connector-python-2.0.4-py3.4.msi)	Apr 1, 2015	139.0K	<a href="#">Download</a>
Windows (Architecture Independent), MSI Installer Python 2.7 (mysql-connector-python-2.0.4-py2.7.msi)	Apr 1, 2015	139.0K	<a href="#">Download</a>

Windows esetén ügyeljünk arra, hogy a Python 2.7-nek megfelelő verziót töltsük le (a képen legalul). Kattintsunk jobb oldalt a Download gombra.

Töltsük le a telepítőt, és futtassuk (rendszergazdai jogosultság szükséges).

Ezután gépünk már alkalmas a DArMail futtatására, a program mappájában dupla kattintással el tudjuk azt indítani.

### Konverzióhoz:

**Thunderbird** <https://www.thunderbird.net/hu/>

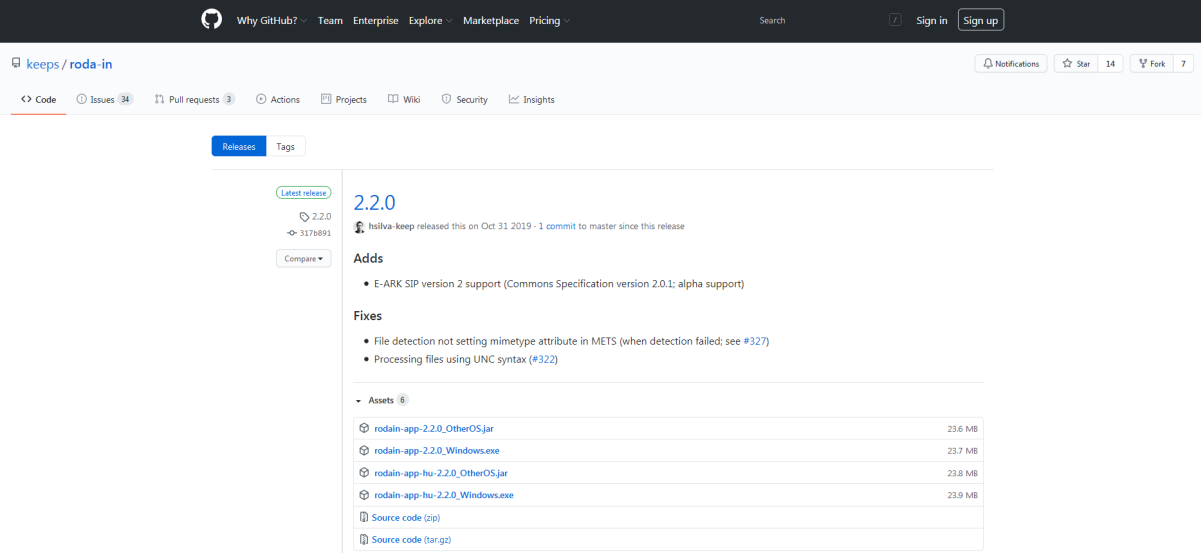
Válasszuk ki a rendszerünknek megfelelő telepítőfájl linkjét:

Töltsük le, és futtassuk (rendszergazdai jogosultság szükséges). A Szokásos összetevőkkel telepíthetjük a programot (2. ablak), valamint az utolsó ablakban megadhatjuk, hogy akarjuk-e alapértelmezett levelezőprogramként használni a Thunderbirdöt (amennyiben nem, vegyük ki a pipát). A telepítés után hozzá kell adnunk egy e-mail fiókot, hogy használni tudjuk a programot e-mailek kezelésére. Az első indításkor még egyszer rákérdez a rendszerrel való integrációra, ha nem akarjuk alapértelmezett programként használni, akkor ne felejtsük el megtenni a megfelelő beállításokat (kivenni a pipákat).

### SIP-csomag előállításához:

## RODA-In SIP Creation Tool <https://rodain.roda-community.org/>

A Download latest release gombra kattinva válasszuk ki a rendszerünknek megfelelő telepítőfájl linkjé (abban a szerencsés helyzetben vagyunk, hogy van magyar verzió, amelynek a fájlnevében szerepel a hu):



The screenshot shows the GitHub repository page for 'keeps/rodain'. The page displays the latest release, version 2.2.0, which was released on October 31, 2019. The release includes several assets for download:

- rodain-app-2.2.0\_OtherOS.jar (23.6 MB)
- rodain-app-2.2.0\_Windows.exe (23.7 MB)
- rodain-app-hu-2.2.0\_OtherOS.jar (23.8 MB)
- rodain-app-hu-2.2.0\_Windows.exe (23.9 MB)
- Source code (zip)
- Source code (tar.gz)

The page also lists 'Adds' and 'Fixes' for this release. The 'Adds' section includes 'E-ARK SIP version 2 support (Commons Specification version 2.0.1; alpha support)'. The 'Fixes' section includes 'File detection not setting mimetype attribute in METS (when detection failed: see #327)' and 'Processing files using UNC syntax (#322)'. The page also shows a 'Compare' button and a 'Tags' tab.

Töltsük le, és futtassuk az alkalmazást (nincs szükség külön telepítésre).