SIP Creator Handleiding

# Inleiding/UGent SIP Creator

Dit is een handleiding voor het gebruik van de tool UGent SIP Creator. Deze software kan gebruikt worden voor het verpakken van data, bijhorende metadata en checksums in een containerformaat. Dit containerformaat is geschikt voor transport over een netwerk of opslag op een harde schijf.

De UGent SIP Creator werd ontwikkeld door de Universiteitsbibliotheek Gent in het kader van het project ‘Digitale containers voor preservatie’, gerealiseerd met steun van de Vlaamse overheid.

De UGent SIP Creator kan worden gebruikt voor de volgende doeleinden:

* Hernoemen van digitale bestanden, zodat deze voldoen aan de raadgevingen en richtlijnen voor het archiveren van digitale bestanden.
* Verpakken van data en de bijhorende metadata in een containerformaat, *bag* genaamd, waarbij eveneens checksums van data en metadata in diezelfde bag worden bewaard. Dankzij deze checksums kan de integriteit van de bagworden gecontroleerd na elke verplaatsing (over een netwerk of naar een harde schijf).

Deze handleiding maakt je stap voor stap wegwijs in de installatie en het gebruik van deze software.

* 1. UGent SIP Creator: Wie? Wat? Hoe? Waarom?
* 2. UGent SIP Creator downloaden en opstarten
* 3. Voortraject: data hernoemen en metadatabestanden voorbereiden
* 4. Nieuwe bag aanmaken
* 5. Bestaande bag aanpassen
* 6. Meerdere bags tegelijk verwerken
* 7. BagIt, Renamer of Log?
* 8. Bibliografie

# UGent SIP: Wie? Wat? Hoe? Waarom?

Wie?

De UGent SIP Creator werd ontwikkeld door de Universiteitsbibliotheek Gent in het kader van het project ‘Digitale containers voor preservatie’.

De Universiteitsbibliotheek Gent begon in 2004 met digitaliseringsprojecten: de steekkaartencatalogus (2004), [[http://www.recollectinglandscapes.be| Recollecting Landscapes ]](2005) en [[http://www.antifonarium-tsgrooten.be/| Antifonarium Tsgrooten]] (2008), naast de voortdurende digitalisering van de eigen collectie. Daarnaast was de Universiteitsbibliotheek partner in de projecten BOM-Vlaanderen (2008) en [[http://www.archipel-project.be/| Archipel ]](2010).

Op vraag van de Vlaamse overheid werd dit project verbonden aan het project ‘[[Stappenplan\_Overdracht\_Digitaal\_Archief|Stappenplan Overdracht Digitaal Archief]]’ (werktitel: ‘Naar een digitaal depotmodel voor kleine archiefinstellingen’) van het Archief en Museum voor Vlaams Leven in Brussel (AMVB). Er werd besloten intens samen te werken en beide projecten op elkaar af te stemmen. Hiertoe werd een gezamenlijke stuurgroep samengesteld met vertegenwoordiging uit de volgende organisaties: Faro, Vlaamse Erfgoedbibliotheek, PACKED vzw, Archiefbank, Heemkunde Vlaanderen, VUB – Master na Master Archivistiek, Felixarchief, BAM, ADVN, AMVB en de Universiteitsbibliotheek Gent.

Aan de ontwikkeling van de UGent SIP Creator werkten mee:

* Inge Van Nieuwerburgh, coördinator digitale bibliotheek, projectcoördinator
* Nicolas Franck, projectmedewerker, ontwikkeling software
* Ariane Van de Vyver, projectmedewerker, documentatie

Wat?

SIP staat voor Submission Information Package. Dit is een voorbereidende stap om digitale (container)bestanden te kunnen opslaan in een informatiesysteem voor langetermijn-opslag, ontwikkeld volgens het [[http://www.projectcest.be/index.php/OAIS|OAIS-model]].

De UGent SIP Creator kan je gebruiken voor:

* Hernoemen van digitale bestanden alvorens ze tot een containerbestand te verpakken, om zodoende archiefbestanden te krijgen conform de digitaliseringsrichtlijnen en adviezen inzake het hernoemen van te archiveren bestanden.
* Aanmaken van digitale “pakketjes”: een containerbestand, *bag* genaamd, bestaande uit data, de bijhorende metadata en checksums van de data en metadata. De checksums stellen je in staat om wijzigingen in data of metadata te traceren na transport over een netwerk of opslag op een harde schijf.

Hoe?

De UGent SIP Creator maakt een dergelijk containerbestand aan op basis van de standaarden BagIt en METS.   
BagIt is een hiërarchische bestandsstructuur (map met onderliggende submappen en databestanden), waarbij data, metadata en de bijhorende checksums worden opgeslagen in eenzelfde map.   
METS staat voor Metadata Encoding and Transmission Standard, en is een metadatastandaard om beschrijvende, administratieve en structurele metadata van digitale objecten uit een collectie te beschrijven. De voertaal daarbij is XML. De UGent SIP Creator zal deze beide verenigen tot één containerbestand.

Indien je de metadata aanlevert in een voor de UGent SIP Creator herkenbaar formaat, dan wordt deze metadata in de bag opgeslagen in een METS.xml-bestand. De aanvaarde bestanden zijn een geldig XML-bestand, een CSV-file of een bag-info.txt.

Meer informatie over de werking van de software, kan je vinden in de softwaredocumentatie.

Waarom?

UGent SIP Creator is een plug & play software. Hiermee kunnen kleinere archiefinstellingen zonder eigen ICT-afdeling meteen aan de slag gaan om archiefdata en metadata voor te bereiden voor opslag in een e-depot.

# UGent SIP downloaden en opstarten

## Systeemvereisten

UGent SIP Creator werkt op Windows, Mac, Linux en op alle andere platformen die Java 6 ondersteunen.

## Systeemgeheugen

Om vlot te kunnen werken met de UGent SIP Creator, dient je computer te beschikken over een processor met een minimale kloksnelheid van 1 GHz, en minstens 1GB aan RAM-geheugen.

Deze bepalen de rekenkracht van de processor, welke belangrijk is bij de berekening van de checksums. Wanneer het gaat om zeer grote of zeer veel bestanden, zullen de vereisten mogelijks hoger liggen, bij kleinere bestanden of kleine hoeveelheden lager.

In Windows kan je de kloksnelheid en het RAM geheugen van je processor nagaan door via *Start* naar *Configuratiescherm* te gaan, en vervolgens op *Systeem* te klikken. Onder *Processor* vindt je de kloksnelheid (bvb 2.66GHz) en het RAM geheugen (bvb 4GB).

Java 6

De UGent SIP Creator vereist minimaal Java 6. De laatste versie van Java kun je downloaden via deze [[http://www.java.com/nl/|link]].

Er zijn verschillende manieren om te controleren welke versie van Java geïnstalleerd is op je computer. Windows-, Mac- en Linuxgebruikers die hiermee vertrouwd zijn, kunnen gebruik maken van de command line in CommandPrompt of Terminal. Het commando:

java –version

geeft als resultaat bijvoorbeeld:

java version "1.6.x\_xx"

Een andere manier om uit te zoeken welke versie van Java je hebt, is je systeem testen via deze [[http://www.java.com/nl/download/installed.jsp|online tool]] op de Java-website. Druk op de knop ***Java-versie controleren*** en de website controleert welke versie op jouw computer geïnstalleerd is. Als je versie ouder is dan 1.6.x, kun je meteen de laatste versie van Java downloaden en installeren.

Downloaden en opstarten

De UGent SIP Creator software kan je downloaden via deze link: http://adore.ugent.be/local/ugent-sip-creator.jar . Kies “Opslaan”. Het kan enkele seconden duren voor de download start. Het maakt in principe niet uit waar je dit bestand opslaat, maar maak bij voorkeur gebruik van de volgende locaties:

* Windows: \Program Files
* Mac: /Applications
* Linux: /usr/local/src

Start de software door te dubbelklikken op *ugent-sip-creator.jar*

[printscreen invoegen van de startpagina van de grafische interface]

# Voortraject: data hernoemen en metadatabestanden voorbereiden

## 3.1 Data: Bestanden hernoemen

Alvorens bestanden te archiveren, moeten deze mogelijks worden hernoemd om te voldoen aan de eisen voor data in een bag:

* Gebruik enkel letters (a-z), cijfers (0-9) en de underscore (\_). Geen spaties of andere (lees)tekens!

Daarnaast zijn er nog een reeks andere gouden regels[[1]](#footnote-1) voor het archiveren van digitale bestanden:

* Betekenisvolle of betekenisloze bestandsnamen
* Maximum 30 karakters
* Noteer eventuele data als JJJJMMDD
* Zorg ervoor dat bestandsnamen uniek zijn
* Pas dezelfde structuur overal consequent toe
* Gebruik extensies van drie tekens
* Gebruik voorloopnullen
* Gebruik kleine letters voor zowel de naam als de extensie

In een reeks bestanden kunnen veel van deze richtlijnen worden bereikt door bestanden in bulk te hernoemen.

Hiervoor kan je gebruik maken van de Renamer-software die in de UGent SIP Creator zit verwerkt.

* Klik op *Renamer* in de taakbalk of klik op *Open* Renamer bij **1. Bestanden hernoemen (optioneel)**.  
  Er opent een dialoogvenster gelijkend op een Windows Verkenner. Dit kan enkele seconden duren. Blader naar de map waarin de te hernoemen bestanden staan. Selecteer deze map. De inhoud van de map verschijnt in de rechterhelft van het scherm.
* Bestanden hernoemen: Vervolgens kunnen er verscheidene acties worden toegepast om bestanden te hernoemen.   
  Deze acties zijn optioneel; niet alle acties moeten worden uitgevoerd.
* *Templates*: verwijder leestekens, verwijder de spaties of vervang de spaties door een underscore (\_)in de geselecteerde bestandsnamen. Door één van deze templates te kiezen, worden de nodige reguliere expressies ingevuld bij de velden *Vervang* en *Door*.   
  Overschrijf deze niet! Om verwarring te voorkomen, voor je andere hernoemingen best in een tweede bewerking uit.
* *Vervang* en *Door*: geef hier aan welk onderdeel van de bestandsnaam hernoemd moet worden, en waardoor dit onderdeel van de bestandsnaam hernoemd moet worden.
* *Negeer grootte letters*: door deze optie aan te vinken worden zowel hoofdletters als kleine letters hernoemd.
* *Pas toe op onderliggende bestanden en mappen*: door deze optie aan te vinken, worden de gekozen acties uitgevoerd op alle onderliggende mappen en data.
* *Voorvoegsel*: voeg voor de bestandsnaam een voorvoegsel toe
* *Navoegsel*: voeg na de bestandsnaam een navoegsel toe
* *Reguliere expressie*: door dit vakje aan te vinken wordt het mogelijk om reguliere expressies te gebruiken bij *Vervang* en *Door*.
* *Bij fout*: selecteer welke actie de Renamer moet ondernemen wanneer er een fout optreedt bij het hernoemen. Standaard staat hier *alles ongedaan maken* geselecteerd.
* *Ok/Simuleer/Reset*: Met *Simuleer* kan je testen hoe de bestanden er zullen uitzien na het hernoemen. Gebruik deze functie ook om fouten te voorkomen (bvb bestanden die door het hernoemen dezelfde naam zouden krijgen).   
  LET OP! Een hernoeming kan niet ongedaan gemaakt worden, dus een simulatie is aangeraden!   
  Wanneer de simulatie ok is, klik op *ok*. Wanneer de simulatie niet naar wens is, klik op *Reset*. Dan worden alle velden van de *Renamer* terug naar hun beginpositie gezet.   
  LET OP! Deze functie herstelt geen reeds hernoemde bestanden naar hun oorspronkelijke naam!
* Bestanden hernummeren: Daarnaast zijn er ook een aantal functies om bestanden te hernummeren. Deze acties zijn optioneel; niet alle acties moeten worden uitgevoerd.
* *Start teller*: dit is het cijfer dat aan het eerste te hernoemen bestand wordt toegevoegd
* *Tel per*: geeft aan met hoeveel het bovenstaand cijfer vermeerderd zal worden
* *Voorvoegsel*: wordt voor het toe te voegen cijfer weergegeven
* *Navoegsel*: wordt na het toe te voegen cijfer weergegeven
* *Positie*: geeft aan waar in de bestandsnaam het cijfer wordt opgenomen. *Absoluut* duidt op een exacte plaats in de bestandsnaam, dus kies vervolgens op welke plaats de nummering moet worden ingevoerd. *Relatief* duidt op een plaats in functie van (een deel van) de bestandsnaam, dus kies *begin* of *einde*.
* *Aantal voorloopnullen*: voorloopnullen garanderen een correcte sortering van bestanden. Zo vermijdt je sorteringen als

File1.txt

File10.txt

File2.txt

Dit probleem kan je verhelpen door voorloopnullen toe te voegen:

File0001.txt

File0002.txt

File0010.txt

Neem een voldoende aantal voorloopnullen. Doorgaans volstaan 3 tot 4 voorloopnullen, zoals in bovenstaand voorbeeld (4 voorloopnullen).

* *Bij fout*: welke actie de Renamer moet ondernemen wanneer er een fout optreedt bij het hernoemen. Standaard staat hier *alles ongedaan maken* geselecteerd.
* *Ok/Simuleer/Reset*: Met *Simuleer* kan je testen hoe de bestanden er zullen uitzien na het hernummeren. Gebruik deze functie ook om fouten te voorkomen (bvb bestanden die door het hernummeren dezelfde naam zouden krijgen).   
  LET OP! Een hernummering kan niet ongedaan gemaakt worden, dus een simulatie is aangeraden!  
  Wanneer de simulatie ok is, klik op *ok*. Wanneer de simulatie niet naar wens is, klik op *Reset*. Dan worden alle velden terug naar hun beginpositie gezet.   
  LET OP! Deze functie herstelt geen reeds hernoemde bestanden naar hun oorspronkelijke naam!

In de rechter schermhelft zie je steeds wat het resultaat van een eventuele hernoeming/hernummering zou zijn. Wanneer de bestanden hernoemd zijn, of indien je de bestanden toch niet wil hernoemen, sluit gewoon het dialoogvenster en je komt terug in de BagIt interface.

## 3.2 Metadatabestanden voorbereiden

De UGent SIP Creator maakt containerbestanden waarin data, metadata en checksums samen worden bewaard. Dit is voornamelijk met het oog op het transporteren van de bestanden over een netwerk, en/of het bewaren van de bestanden in een (online) repository.

We willen een containerbestand creëren dat zo generiek mogelijk is, en dus door zoveel mogelijk repositories kan worden aanvaard. Hiertoe dient de metadata in een bepaald formaat te worden omgezet: een *METS.XML-bestand*.

De gebruiker hoeft dit niet zelf te doen, de UGent SIP Creator genereert dit mets.xml-bestand zelf, indien je de metadata aanlevert in een herkenbaar formaat. De mogelijkheden om metadata aan te leveren zijn:

* XML-bestand dat wordt geëxporteerd uit de eigen catalogus/het eigen archiefbeheerssysteem
* CSV-bestand
* Bag-info.txt bestand, waarin een minimale set aan Dublin Core elementen zitten
* PREMIS-bestand

### XML-bestand

Het geniet de voorkeur om metadata toe te voegen door middel van een XML-bestand. Dankzij XML kunnen diverse metadatastandaarden worden ondersteund: MARC, MODS, EAD, MADS, DC volgens OAI, AMD, TextMD, VMD, Premis, etc. Zolang deze maar in een [http://www.loc.gov/standards/| geldig XML-schema] zijn gegoten.

Bij voorkeur bevat de XML een namespace. Dit is niet verplicht, maar wel ten zeerste aangeraden, dit om het succesvol mappen van de metadata mogelijk te maken.

Indien de metadata reeds beschreven zijn in een archiefbeheerssysteem waaruit XML-exports mogelijk zijn, kan de metadata op deze manier aan de bag worden toegevoegd.

Indien het XML-bestand aan geen van deze standaarden voldoet, krijg je een waarschuwing dat de metadata een zéér beperkte ondersteuning kent in repositories, en daardoor bij een export niet zal worden meegenomen.

### CSV-bestand

[Nog steeds geen informatie over CSV-bestand]

* Waar te vinden?
* Hoe te genereren
* Welke voorwaarden?

### Bag-info.txt-bestand

Een bag-info.txt bestand is een zeer toegankelijke manier om metadata toe te voegen. Je hoeft hiervoor helemaal geen kennis te hebben van informatica. Wel veel tijd, want dit vergt manueel werk.

Een bag-info.txt bestand bevat een vlakke beschrijving van de bag:

Bag-Size: 758.7 KB

Bagging-Date: 2011-01-09

DC-AccessRights: ugent

DC-Creator: Claus, Hugo

DC-Identifier: rug01: 000009017

DC-Title: Het verdriet van België

Daarbij worden de Bag-Size en de Bagging-Date gegenereerd door de UGent SIP Creator, de overige beschrijvende metadata kan je zelf toevoegen. Een overzicht van de Dublin Core Elements vind je [[DC,\_DCMES|hier]].

Wat er aan beschrijvende metadata wordt opgenomen, daar ben je zeer vrij in: zeer uitgebreid of zeer beperkt. Zo lang de structuur *DC-Element:* maar gerespecteerd wordt.

De voorwaarden waaraan een bag-info.txt bestand moet voldoen zijn de volgende:

* Het bestand moet een .txt-file zijn
* Het bestand moet [[DC,\_DCMES|Dublin Simple Elements]] bevatten
* De filename hoeft niet bag-info.txt te zijn, maar met deze bestandsnaam geef je wel te kennen dat dit bestand de beschrijvende metadata bevat.

Doordat het bestand DC elementen bevat, kan de UGent SIP Creator metadata mappingen maken op basis hiervan.

### PREMIS-bestand aanmaken

[Hoe gaat dit in zijn werk???]

[Antw Nicolas:

druk op de knop “Premis” in de tabel van metadata.

– selecteer een xml bestand met premis. Dit bestand wordt gevalideerd volgens schema van Premis.

– In de mets.xml wordt een log bijgehouden van alle wijzingen in een premis-record. De aangeleverde xml wordt hiermee gemerged. Maar dat gebeurd op zo'n manier dat de loggegevens nooit overschreven worden.

– afspraak voor wie zo'n premis aanlevert:

“Object” met xmlID “bagit”

Wordt ingevuld door SIP-creator. Indien aanwezig in de aanlevering, wordt die verwijderd.

“Event” met eventType “bagit” wordt ingevuld door SIP-creator. Indien aanwezig in de aanlevering, wordt die verwijderd.

Telkens de premis wordt aangepast via deze knop, worden alle events, agents, objects en rights weggegooid in het huidige, behalve de gegevens die door de SIP-creator zelf ingevuld worden zoals hierboven uitgelegd.

Zodoende blijven de gegevens zich niet opstapelen.]

# Nieuwe bag aanmaken

* Ga naar *BagIt* en kies *Nieuwe bag* (Ctrl+N) of klik op het icoontje ***Nieuwe bag***
* Geef in het veld *Bag-Id* een identifier op voor deze bag.   
  De UGent SIP Creator genereert automatisch een voorstel. Je kan deze overschrijven en zelf een identifier toekennen, maar we raden aan de voorgestelde identifier te kiezen, om uniciteit te garanderen.   
  **LET OP!** De identifier moet beginnen met een letter (geen cijfer). Indien ongeldig, zal het tekstvak rood kleuren en is het niet mogelijk een bag aan te maken.

## 4.1 Bestanden toevoegen

* Klik op het groene icoontje *Voeg toe*. Er opent een dialoogvenster gelijkend op een Windows Verkenner. Dit kan enkele seconden duren.

Blader naar de map/bestanden die je wil verpakken tot een bag. Selecteer de map(pen) en/of bestanden die je wil toevoegen.   
LET OP! De bestanden dienen wel geldige bestandsnamen te hebben.  
LET OP! Bij het samenvoegen van bestanden uit diverse bronmappen: eenzelfde bag kan geen meerdere bestanden met dezelfde bestandsnaam bevatten.

* De toegevoegde bestanden verschijnen in de linker helft van het scherm. Mappen zijn uitklapbaar, zodat onderliggende mappen en bestanden zichtbaar gemaakt kunnen worden.
* Indien een bestand of map verwijderd moet worden uit de selectie: selecteer het betreffende bestand of de betreffende map, en klik op het icoontje *Verwijder selectie*

## 4.2 Metadata toevoegen

Voor het toevoegen van metadata zijn er twee opties:

* Laad de metadata op vanuit een bestand
* Voeg de metadata manueel toe door middel van invulvelden

### Metadata opladen vanuit een bestand

* **Metadata opladen vanuit een XML bestand**

Voeg bij voorkeur metadata toe door middel van een XML-bestand. Indien de metadata reeds beschreven zijn in een catalogussysteem waaruit XML-exports mogelijk zijn, kan de metadata aan de bag worden toegevoegd vanuit deze XML-export. Meer informatie over het hoe en het waarom van XML-imports kan je hier vinden.

* Ga naar **3. Metadata toevoegen** en kies *Vanuit bestand*
* Klik op *Voeg toe*. Een drop down menu verschijnt. Selecteer *XML*
* Blader naar de map waarin het XML-bestand zit opgeslagen. Selecteer het gewenste XML-bestand en klik op *Open*. Het geselecteerde XML-bestand verschijnt in de lijst van toegevoegde metadatabestanden. Door op de toegevoegde lijn te klikken, wordt de inhoud van het XML-bestand zichtbaar. Het type XML wordt normaal gezien herkend door de bagger (en dus automatisch ingevuld). In andere gevallen kan je bij “label” aangeven welk soort metadata de XML bevat.  
  Voeg géén mets.xml bestand toe. De opgeladen XML zal worden omgezet naar een mets.xml-bestand. Zo ontstaat een mets.xml in een mets.xml, en dat geeft verwarring.
* Indien je een bestand wenst te verwijderen uit de lijst van toegevoegde metadatabestanden, selecteer het bestreffende metadatabestand en klik op *Verwijder selectie*.
* **XML Mapping**

Met de UGent SIP Creator is het mogelijk om een XML-export uit de eigen catalogus/archiefbeheerssysteem te mappen naar andere XML-formaten. Deze mogelijkheid wordt aangeboden voor het geval men de bags wil opslaan in een (online) repository. Het is namelijk mogelijk dat een repository slechts één bepaald XML-formaat ondersteund.

Er kunnen meerdere XML-mappings aan één bag worden toegevoegd. Belangrijk is wel dat het geselecteerde XML-bestand voldoet aan de geldigheidsvoorwaarden zoals gedefinieerd door de Library of Congress. Houd er ook rekening mee dat een mapping betekent dat dezelfde informatie meermaals wordt opgeslagen binnen hetzelfde containerbestand, en dus eigenlijk een soort dubbel is.

* Ga naar **3. Metadata toevoegen** en kies *Vanuit bestand*
* Klik op *Voeg toe*. Een drop down menu verschijnt. Selecteer *XML mapping*.
* Blader naar de map waarin het XML-bestand zit opgeslagen. Selecteer het gewenste XML-bestand en klik op *Open*.
* In het dialoogvenster verschijnt bij *van* het type van je bron-XML-bestand. Bij *naar* geef je aan naar welk XML-formaat je het bron-XML-bestand wil mappen. De UGent SIP Creator analyseert de bron-XML en zal aangeven welke mappingen ervoor beschikbaar zijn.
* Indien je een bestand wenst te verwijderen uit de lijst van toegevoegde metadatabestanden, selecteer het bestreffende metadatabestand en klik op *Verwijder selectie.*
* **Metadata opladen vanuit een CSV-bestand**

[Uitleg over CSV]

* Ga naar **3. Metadata toevoegen** en kies *Vanuit bestand*
* Klik op *Voeg toe*. Een drop down menu verschijnt. Selecteer *CSV*.
* Blader naar de map waarin het CSV-bestand zit opgeslagen. Selecteer het gewenste CSV-bestand en klik op *Open*.
* Kies een scheidingskarakter: kies het scheidingskarakter tussen de veldnamen en de veldwaarden (komma, puntkomma, spatie of tab).
* Kies een aanhalingskarakter: door aanhalingstekens kan de software eventuele leestekens in de veldwaarden niet verkeerdelijk interpreteren als een scheidingskarakter. Kies enkele of dubbele aanhalingstekens als aanhalingskarakter.
* Klik op *Voltooien*
* Indien je een bestand wenst te verwijderen uit de lijst van toegevoegde metadatabestanden, selecteer het bestreffende metadatabestand en klik op *Verwijder selectie*.
* **Metadata opladen vanuit een bag-info.txt bestand**

Indien je niet beschikt over een systeem waaruit XML-exports mogelijk zijn, en je hebt ook geen CSV-file ter beschikking, dan is er nog een derde mogelijkheid om vanuit een bestaand bestand op te laden: een bag-info.txt.

Een bag-info.txt bestand bevat een vlakke beschrijving van de bag, uitgedrukt in Dublin Core Elements. Hoe zo een bag er uit moet zien, en wat de restricties zijn, vind je hier.

Doordat het bestand DC elementen bevat, kan de UGent SIP Creator metadata mappingen maken op basis hiervan.

* Ga naar **3. Metadata toevoegen** en kies *Vanuit bestand*
* Klik op *Voeg toe*. Een drop down menu verschijnt. Selecteer *BAG-INFO*
* Blader naar de map waarin het bag-info.txt bestand zit opgeslagen. Selecteer het gewenste bag-info.txt bestand en klik op *Open*.
* Het bag-info.txt bestand zal worden geconverteerd naar Dublin Core (OAI). Klik op *Voltooien* om het bag-info.txt bestand toe te voegen aan de bag.
* Indien je een bestand wenst te verwijderen uit de lijst van toegevoegde metadatabestanden, selecteer het bestreffende metadatabestand en klik op *Verwijder selectie*.
* **Metadata opladen vanuit een PREMIS-bestand**
* Ga naar **3. Metadata toevoegen** en kies *Vanuit bestand*
* Klik op *Premis*. Een drop down menu verschijnt:
* *Toon*:
* *Leeg*:
* *Aanpassen*:

### 

### Metadata manueel toevoegen

Indien de metadata niet kunnen worden opgeladen vanuit een bestand, of indien deze metadata ontoereikend zijn, kan er metadata manueel worden toegevoegd

* Ga naar **3. Metadata toevoegen** en kies *Manueel*
* *Standaard*: door deze aan te vinken, kan je Dublin Core Simple Elements velden toevoegen.  
  Eenzelfde DC-veld mag meermaals voorkomen (behalve Bag-ID). Anderzijds dient niet elk veld ingevuld te worden. Hoe meer velden ingevuld, hoe vollediger de metadataset, maar hoe tijdrovender het toevoegen van de metadata is.
* Door de *Standaard* checkbox uit te vinken, kan je zelf omschrijvingen toekennen en op deze manier metadata toevoegen. Let op dat je in de eerste velden (de velden die de eerste kolom zullen vormen) enkel gebruik maakt van letters, cijfers, koppeltekens en underscores.

### Overige metadata?

Indien je nog metadata wil toevoegen uit andere bestanden, kan je deze opladen onder *Data toevoegen*. Deze bestanden worden dan mee opgenomen in de bag, maar zullen door de UGent SIP Creator worden beschouwd als *payload*, en niet als metadata. Bijgevolg is deze metadata ook niet doorzoekbaar en kan deze niet gemapt worden naar andere standaarden.

## 4.3 Bag opslaan

* Klik op het icoontje *Opslaan als*  in de balk bovenaan, of klik op het icoontje *Opslaan als..*  bij **4.Opslaan**
* *Bestand*: de software stelt de bag-id voor als bestandsnaam. Klik op Selecteer. Kies (eventueel) een nieuwe bestandsnaam en een locatie waar de bag moet worden opgeslagen.
* *Type opslag*: Kies welk type bag je wil opslaan. Om een bag onder meerdere vormen op te slaan: kies een type en sla de bag op. Heropen de bag en sla deze op als een ander type.
* *Type checksum*: Kies een algoritme waarmee de bag wordt opgeslagen. Ze zijn gerangschikt volgens van eenvoudig naar meer complex. De laatste zijn dus “veiliger”, maar vereisen eveneens meer rekencapaciteit van de processor om te maken of te controleren.

# Bestaande bag aanpassen

* Ga naar BagIt en kies *Open bag* (Ctrl + O)of klik op het icoontje ***Open bag***. Blader naar de gewenste bag.

## 5.1 Bestanden toevoegen

## 5.2 Metadata toevoegen

### Metadata opladen vanuit een bestand

* **Metadata opladen vanuit een XML bestand**
* **XML Mapping**
* **Metadata opladen vanuit een CSV-bestand**
* **Metadata opladen vanuit een bag-info.txt bestand**
* **Metadata opladen vanuit een PREMIS-bestand**

### Metadata manueel toevoegen

### Overige metadata?

## 5.3 Bag opslaan

Er zijn meerdere mogelijkheden om een bewerkte bag op te slaan:

* Klik op het icoontje ***Opslaan*** in de balk bovenaan. De bestaande bag zal overschreven worden. De checksums zullen herberekend worden. De oorspronkelijke bag zal overschreven worden.
* Klik op het icoontje ***Opslaan als*** in de balk bovenaan, of klik op het icoontje ***Opslaan******als..*** bij stap **4.Opslaan**. De software stelt de bag-id voor als bestandsnaam. Kies een nieuwe bestandsnaam en een nieuwe locatie als je wil dat de oorspronkelijke bag niet wordt overgeschreven.
* Klik op het icoontje ***Sluit bag*** in de balk bovenaan. Als je wijzigingen hebt aangebracht in de bag, zal de software vragen of u de huidige bag wil bewaren. Klik *ja* als je de oorspronkelijke bag wil overschrijven. Klik *nee* als je de wijzigingen niet wil bewaren. De oorspronkelijke bag blijft bewaard.

## 5.4 Valideer bag

Met deze optie controleer je of de bag nog exact dezelfde inhoud bevat als op het moment dat deze werd aangemaakt.

- klopt het aantal databestanden en het aantal metadatabestanden nog? Werd er niks toegevoegd of verwijderd?

- klopt de inhoud van de databestanden en metadatabestanden nog? Hiervoor worden de checksums herberekend.

Afhankelijk van de grootte van de bag, kan deze optie flink wat geheugen en rekencapaciteit van de computer vergen.

* Klik in de taakbalk op het groene icoontje *Valideer bag*

## 5.5 Is deze bag volledig?

Deze optie is een afgeslankte optie van ***Valideer bag*:** de software gaat na of alle verplichte bestanden aanwezig zijn en of alle databestanden zoals opgegeven in het manifestbestand aanwezig zijn. Maw, of het aantal bestanden nog hetzelfde is als toen de bag werd gemaakt. Er worden geen checksums herberekend, waardoor deze controle minder volledig is dan *Valideer*, maar eveneens minder rekenkracht vraagt van de pc.

* Klik in de taakbalk op het gele icoontje *Is bag volledig?*

# Bestaande bag exporteren

Deze functie is bedoeld om de aangemaakte bag om te zetten naar een formaat voor een specifiek repository. Momenteel ondersteunt de UGent SIP Creator enkel exports naar DSpace repositories.

* Ga naar *Bagit* en klik op *Open bag* (Ctrl + O) of klik op het icoontje **Open bag** enblader naar de gewenste bag. Selecteer de gewenste bag en klik op *Open*.
* Ga naar *Bagit* en klik op *Exporteer* (Ctrl + E)
* Kies een *bestemmingsmap* en geef een *bestandsnaam* op.
* Kies een exportformaat: DSpace Mets Archive (deze bevat bij voorkeur MODS metadata) of DSpace Simple Archive (deze bevat bij voorkeur OAI metadata) en klik op *Voltooien*.

De exportbestanden die worden gecreëerd in deze bestemmingsmap zijn nu klaar voor import naar een

(online) repository.

# Meerdere bags tegelijk verwerken

## Valideer reeks bags

In plaats van elke bag afzonderlijk te valideren, is het ook mogelijk om een reeks bags tegelijk te controleren. Dit kan bijvoorbeeld handig zijn wanneer je een periodieke validatie wil doen.

* Ga naar *BagIt* en kies *Valideer reeks bags*.
* Klik op *Selecteer* en blader naar de map waar de reeks bags staat opgeslagen. Selecteer de gewenste bags en klik op *Open*.
* Standaard wordt er enkel gecontroleerd of de bags volledig zijn (dwz: zitten alle bestanden uit het manifest-md5.txt bestand nog in de bag). Door *controleer geldigheid* aan te vinken, zal ook de eigenlijke inhoud van de bestanden dmv checksums gecontroleerd worden. Deze optie zal dus meer rekenkracht vergen van de pc dan enkel een controle op volledigheid.
* Klik op *Start validatie*
* Indien je een bag van naderbij wil bekijken, selecteer dan deze bag en klik op *Open selectie*

## Creëer reeks bags

Deze functie is bedoeld om van een reeks mappen met data en eventueel metadata een reeks bags te maken.  
LET OP! De metadata moet zich reeds in de vorm van een XML-bestand bij de data bevinden.

* Ga naar *Bagit* en klik op *Creëer reeks bags*
* Klik op *Selecteer* en blader naar de map waar de reeks data is opgeslagen. Selecteer de gewenste mappen en klik op *Open*
* Indien er reeds metadatabestanden (in XML-formaat) werden aangemaakt, kunnen deze worden toegevoegd. Geef de titel(s) op van de metadatabestanden, gescheiden door een “,”  
  Bvb: marc.xml, mods.xml
* Door op *Behoud metadata* te klikken geef je aan dat het om metadata gaat en niet om data (om verwarring tussen data en metadata uit te sluiten).
* Vink *Voeg DC-record toe*  aan als je nog geen DC-record hebt ingegeven onder *Metadata bestanden*. Op die manier wordt de toegevoegd metadata gemapt naar DC, wat belangrijk is voor export naar een e-depot.
* Door op *Vervang bronmap* te klikken, zal de oorspronkelijke mappenstructuur overschreven worden door de bag-structuur. Indien je de oorspronkelijke mappenstructuur wil behouden, kan je dit vakje uitvinken. Geef dan een *Bestemmingsmap* op.   
  LET OP! Dit moet een nieuwe, lege map zijn. Je kan geen bestaande map opgeven als bestemmingsmap.
* Klik op *OK*

## Controleer manifest

Deze functie is bedoeld voor mappen en data waarvan reeds een manifest bestaat (bvb aangemaakt in een eerder archiveringsfase) te controleren op geldigheid.

* Ga naar *Bagit* en klik op *Controleer manifest*
* Klik op *Selecteer* en blader naar de map(pen) waarin de te controleren manifests zich bevinden
* Klik op *Start controle*.

# BagIt, Renamer, Log of Help?

Via de taakbalk bovenaan navigeer je door de verschillende software-onderdelen:

* *BagIt* brengt je in de bagit-interface, waar de data en metadata worden samengevoegd tot een containerbestand.
* *Renamer* brengt je in de renamer-interface. Hier kan je bestanden hernoemen en hernummeren, zodat deze voldoen aan de richtlijnen van het archiveren van digitale bestanden.
* *Log* brengt je in het logbestand dat de UGent SIP Creator software bijhoudt. In geval er iets misloopt, kan je in dit logbestand nagaan wat er is foutgelopen.

# Bibliografie

* softwaredocumentatie

1. Boudrez (Filip), Digitaal Archiveren: richtlijn & advies nr. 3, Mappenstructuur en bestandsnamen voor digitale documenten, pp. 4-5.

   <http://www.projectcest.be/index.php/FARO_Handboek_digitaliseren_van_documenten> (pp 71-72) [↑](#footnote-ref-1)