

## EDIT - feature request #9262

### Several open issues for cdmlight export related to Cactaceae and other Caryophyllales

10/22/2020 12:45 PM - Andreas Müller

<b>Status:</b>	Closed	<b>Start date:</b>	
<b>Priority:</b>	New	<b>Due date:</b>	
<b>Assignee:</b>	Katja Luther	<b>% Done:</b>	40%
<b>Category:</b>	cdmadapter	<b>Estimated time:</b>	0:00 hour
<b>Target version:</b>	Release 5.18		
<b>Severity:</b>	normal		
<b>Description</b>			
Mail WGB 2020-10-22:			
<p>1. Fehlerhafter Literatur-Output — zwei Autoren + et al. statt Erstautor + et al., siehe angehängtes Mail dazu Das muss unbedingt gelöst werden, sonst müsste ich das alles manuell in Word nachformatieren, und das sind etliche Referenzen, die das betrifft.</p> <p>2. Fehlerhafte „fide“ Ausgabe. Siehe auch die angehängten Mails dazu. Soweit ich sehe, steht überall falsch fide, es müsste immer „designated by“ heißen. Ich verstehe auch nicht ganz, warum das nur bei Lektotypen gelten soll, Neotypen und Epitypen werden schließlich auch designiert? WGB: ja klar, Lekte-, Neo- und Epitypen, hatte ich unten vergessen zu ergänzen. Noch mal für die Ausgabe: Das „fide“ gehört zu den Source-Referenzen für Typen, das „designated by“ (das früher bei allen Referenzen stand) gehört zu den Type-Status-Referenzen bei Lekte-, Neo- und Epitypen. WGB (telefonisch): Designated by gehört direkt hinter die Sammlung, nicht in die eckige Klammer, diese ist nur für fide-Referenzen Das betrifft auch den Portal output. =&gt; Portal muss noch umgestellt werden auf TypeDesignationSetManager</p> <p>3. Ausgabe -title cache generation not implemented-. Das kommt, soweit ich sehe, nur bei specimens vor, und zwar bei allen, aber nicht bei Typen. Specimens gibt es nur in bestimmten Gattungen, z.B. in Lepismium, Rhipsalis, Epiphyllum, Disocactus. WGB: Ich habe das im Output ohne Specimen als factual data nicht gesehen, ist also für den jetzt anstehenden Artikel noch nicht relevant. Das kann aber an anderer Stelle (Amaranthaceae, Vanessas Artikel, hoffentlich auch noch dieses Jahr!) relevant werden. Andreas fragte, wo denn diese Specimen stehen – ich nehme an, dass das in den von Nadja genannten Gattungen specimen als factual data am Taxon sind (?). moved to new ticket <a href="#">#9309</a>.</p> <p>4. Bug-Klammerautoren bei Gattungstypen hatte Katja ja schon gelöst (im nächsten Release).</p> <p>5. Einschluss von syn.-sec.-Referenzen (Kurzzitat) in der homotypischen Gruppe, in HomotypicGroup.csv (und Taxon.csv, siehe unten) Ich gebe derzeit zwei Dokumente aus, eine mit den syn.-secs aber ohne homotypische Gruppe und dann das eigentliche Format mit homotypischen Gruppen. Das verwirrt die Reviewer, von denen wir oben auch Input zu den syn.-sec.-refs. wollen, obwohl wir das normalerweise letztendlich nur auf dem Website veröffentlichen. Ich hätte also gern ein Feld, in dem nach jedem Namen die syn.-sec.-Kurzzitatreferenz mit ausgegeben wird.  Also ein neues Feld HomotypicGroupStringWSec in der HomotypicGroup.csv.</p> <p>6. Homotypical Group in Taxon.csv Ich stoße bei den großen Datenbanken (Euro + Med, Caryophyllales) an die Grenzen der Datenverarbeitung mit Access. Knackpunkt ist die Konkatenierung von Strings aus komplexen Strukturen. Das wirkt sich z.Zt. vor allem an zwei Stellen aus, distributions (derzeit nicht so dringend) und homotypical groups: Um den Taxonnamen incl. nom. Zitat mit der Sec.-Referenz getrennt von seinen homotypischen Synonymen ausgeben zu können, muss ich mir diese group selbst (ohne den Taxonnamen) zusammenstellen. Im verarbeiteten Output sieht das dann so aus: =  Achyranthes bidentata Blume, Bijdr.: 545. 1826. Sec. WFO Import Amara Cheno-Polyg-Dianthus June 2018 = Centrostachys bidentata (Blume) Standl. in J. Wash. Acad. Sci. 5(3): 75. 1915. = Achyranthes fauriei var. tomentosa Honda, Nom. Pl. Jap. ed. emend.: 373. 1957. Achyranthes bidentata var. tomentosa (Honda) H.Hara, Fl. E. Himal.: 635. 1966. =  Der erste Absatz ist der akzeptierte Name mit Sec.-Ref. (kein Problem) Der 2. Absatz ist die homotypische Gruppe des akzeptierten Namens OHNE den Namen selbst (hier liegt das Problem—</p>			

CDM light liefert mir HomotypicGroup.HomotypicGroupString = Achyranthes bidentata Blume; Centrostachys bidentata (Blume) Standl.

Der 3. Absatz ist die homotypische Gruppe eines heterotypischen Synonyms. Das entspricht direkt der Ausgabe HomotypicGroup.HomotypicGroupString in CDM light, also auch kein Problem.

Die Lösung hier wäre also ein zusätzlicher String ohne den akzeptierten Namen — das Problem ist aber natürlich, dass der akzeptierte Namen aus einer homotypical group ja zwischen verschiedenen Klassifikationen variieren kann. Man müsste das also in die Taxon.csv stellen, als HomotypicSynonymGroupString und dann natürlich auch HomotypicSynonymGroupStringWSec.

AM: Da wir ja nur *eine* Klassifikation ausgeben, würde ich davon ausgehen, dass es je homotypischer Gruppe eindeutig ist, welches Taxon das akzeptierte ist. Wir könnten das also schon auch in csv für homotypische Gruppen aufnehmen statt in Taxon.csv (deswegen heißt es ja edm\*light\*).

Oder übersehe ich da was oder ist es besser aufgehoben in der Taxon.csv?

WGB: Ja, da habe ich wohl zu generell gedacht — wenn wir bei der Limitation bleiben, ist das korrekt. Falls es zu Publikationen mit mehreren Klassifikationen kommt, muss man die Word-Dokumente dann vor dem Indizieren zusammenführen

#### 7. Excluded und Unplaced in Sort Order [Mail vom 20.10.]

beim Füllen des Felds SortOrder in der Taxon.csv werden die als excluded oder als unplaced markierten Datensätze korrekt ans Ende gesetzt, aber leider werden die beiden Kategorien gemischt. Das bringt meinen Output dieser Gruppen durcheinander – können wir bitte zuerst die unplaced und dann die excluded am Ende einordnen?

mail KL:

Ich habe mir jetzt nur die unplaced und excluded Taxa, die direkt unter Cactaceae hängen angeschaut, aber da ist der Sortindex korrekt. Gibt es dazu ein Beispiel, wo die Sortierung gemischt ist?

WGB: merkwürdig, die kamen bei mir durcheinander raus – aber erst am Ende. Ich kümmere mich darum.

#### 8. Autonome in homotypischen Gruppen

wenn Autonome in homotypischen Gruppen eingeordnet sind, wird dahinter gar kein Trennzeichen gesetzt, hier mal ein Beispiel = Echinocactus taltalensis Werderm. in Notizbl. Bot. Gart. Berlin Dahlem 10(97): 763. 1929. Copiapoa taltalensis subsp. taltalensis Copiapoa humilis subsp. taltalensis (Werderm.) Doweld in Sukkulenty 4: 49. 2002. Copiapoa humilis var. taltalensis (Werderm.) A.E.Hoffm., Cact. Fl. Silvestre Chile: 118. 1989.

=> not an issue anymore if 10. will implemented

#### 9. Chronologische Anordnung der Synonyme in homotypischen Gruppen

[wir wollen die Kakteen in Willdenowia publizieren, das ist also eilig]

ERS: Im jetzigen Word-Output erscheinen die Synonyme vollkommen unsortiert, und zwar sowohl innerhalb der homotypischen Gruppen (abgesehen davon, dass das Basionym immerhin an erster Stelle steht) als auch außerhalb.

Man könnte nun entweder alphabetisch sortieren (so ist es in der Nepenthes-Liste umgesetzt, wobei allerdings sogar die invaliden Namen der alphabetischen Sortierung unterworfen sind, was sehr unüblich ist — normalerweise stehen diese am Ende), oder chronologisch. Ich bevorzuge eine chronologische Sortierung, soweit ich weiß, ist das auch Willdenowia-Standard. So haben wir es auch damals im Cichorieen-Portal umgesetzt, siehe z.B.

[http://cichorieae.e-taxonomy.net/portal/cdm\\_dataportal/taxon/0c17975e-14d9-432b-846a-9f290c1a88fe/synonymy](http://cichorieae.e-taxonomy.net/portal/cdm_dataportal/taxon/0c17975e-14d9-432b-846a-9f290c1a88fe/synonymy). Das Schöne dabei ist, dass bei fehlenden Jahreszahlen wiederum alphabetisch sortiert wurde. Also, im Cichorieenportal ist es schon sehr ausgefeilt, das hat viel Arbeit gekostet, wie ich mich erinnere.

Wenn es das schon mal im Portal gibt, ist ja die Implementierung zumindest schon mal spezifiziert — wir möchten generell Willdenowia-Standard folgen.

Die Anordnung der heterotypischen Synonyme („außerhalb“) erfolgt im GDM2Word-Programm, mit den invaliden Namen am Ende. Ich werde versuchen, das zu einer Anordnung nach (Basionym-) Jahr mit nachgestellten invaliden Namen zu ändern.

WGB: Dazu müsste ich aber den Gruppenstring jeweils immer parsen, um das älteste Jahr herauszubekommen und dabei dann noch invalide etc. designations ausschließen, das ist kompliziert. Daher also als zusätzliches GDM-light-Feld:

HomotypicGroupYear — Das älteste Jahr eines validen Namens in der homotypischen Gruppe wurde durch sortindex gelöst

#### 10. Homotypische Gruppen: Identitätszeichen vor jedem Synonym

ERS: Eine Bemerkung zum Layout: In der Willdenowia ist es Standard, dass auch innerhalb der homotypischen Gruppen das Identitätszeichen vor jedem Synonym ein Identitätszeichen steht (siehe z. B. Willdenowia 49: 423. 2019, unter Drimia anthericoides). Ich finde, das verbessert die Übersicht. Die Namen alle ohne weitere Trennzeichen so aneinanderzureihen, wie es jetzt der Fall ist, finde ich nicht so schön, aber das mag Geschmacks- bzw. Gewöhnungssache sein. Im Caryophyllales-Paper wurde es übrigens auch nach Willdenowia-Standard umgesetzt.

Dies umzusetzen müsste einfach sein — und falls für einen anderen Publikationsort das nicht gewünscht wird, ist es auf jedem Fall einfacher, das wieder herauszunehmen, als es einzufügen.

→ sollte man das dann nicht in der Word-Generierung machen, entweder mit Identitätszeichen oder ohne, je nachdem wie es für die entsprechende Publikation gebraucht wird. Der edmLight-Export sollte ja möglichst generisch bleiben.

WGB: Ich sehe gerade, dass das in Willdenowia (im von Eckhard genannten Artikel) tatsächlich so gemacht wird. Ich würde ja nach jedem Synonymzeit einen Punkt setzen. Egal, wenn wir Willdenowia folgen, müssen die anderswo gemachten Punkte weggelassen werden, also bleibt etwas zu tun

#### 11. Homotypische Gruppe: Komma

Vor dem nomenklatorischen Status, nach der Jahreszahl sollte ein Komma stehen, das ist wohl auch Willdenowia-Standard,

fand Nadja auch:

WB: Außerdem finde ich es eigentlich nicht schön, dass der nom. status. nach einem Punkt hinter der Jahreszahl kommt — ich fände da ein Komma besser; was meinst Du? Z.B.:

≙ *Turbinacarpus pseudomacrolele* var. *lausseri* Diers & Gerhart Frank in Kakteen And. Sukk. 42(5): 108. 1991.  
*Kadenicarpus pseudomacrolele* var. *lausseri* (Diers & Gerhart Frank) Doweld in Sukkulenty 1(1): 22. 1998. nom. inval.  
*Pediocactus pseudomacrolele* var. *lausseri* (Diers & Gerhart Frank) Halda in Acta Mus. Richnov., Sect. Nat. 5: 22. 1998.  
nom. inval. *Turbinacarpus pseudomacrolele* f. *lausseri* (Diers & Gerhart Frank) Panar. in Pianta Grasse 18: 120. 1999.  
nom. inval. *Kadenicarpus pseudomacrolele* subsp. *lausseri* (Diers & Gerhart Frank) Doweld in Sukkulenty 3: 37. 2000.  
nom. inval. *Turbinacarpus kranzianus* f. *lausseri* (Diers & Gerhart Frank) D. Donati, Revis. Tasson. Gen. Turbinacarpus: 21. 2003. nom. inval.

Andererseits steht vor der Jahreszahl ja auch ein Punkt ...

=> Verwendung des fullTitleCache

## 12. NameTypeDesignations werden nicht exportiert

### 13. NK: Semikolon zwischen zwei Notes

Zwei Notes werden mit einem Semikolon zwischen Ihnen ausgegeben, zum Beispiel

*Hariota* DC. is a homonym of *Hariota* Adans. (1763) and was therefore replaced by the name *Hariota* Britton & Rose.;

Phylogenetics:

Die Notes sind normalerweise in ganzen Sätzen geschrieben, deswegen ist Semikolon da überflüssig.

Ich fände es auch für die Lesbarkeit besser, wenn zwischen zwei Notes ein Absatz eingefügt werden würde => KL: Das Problem mit den Notes sollte wahrscheinlich eher bei der Word Generierung gelöst werden, da könnte man ja ein „;“ durch ein „./n“ oder ähnliches ersetzt werden.

### 14. Doppelte Autoren bei nomenklatorischen Referenzen. CDM: *Hariota* Britton & Rose in L.H.Bailey, Standard Cycl. Hort.: 1432-1915

Word: *Hariota* Britton & Rose in L.H.Bailey, **L.H.Bailey**, Standard Cycl. Hort. 1432. 1915 => gefixt im cdmlight output

### 15. Inautoren mit Initials: CDM: *Rhipsalis bambusoides* F.A.C.Weber in Bois, Dict. Hort.: 1048. 1898

Word: *Rhipsalis bambusoides* F.A.C.Weber in Bois\*, *D.G.J.M.\**, Dict. Hort. 1048. 1898 => muss in Word gefixt werden, ggf. hinzufügen des TaxonName.fullTitleCaches, der dies richtig beinhaltet. Auch Punkt 11. würde mit dem fullTitleCache gelöst -> fulltitleCache wird verwendet

### 16. Doubtful synonyms should be displayed with "?"

(issues 3 and 4 removed as 3 is related to titleCache generation for DNA-Samples and 4 is solved already)

#### Related issues:

Related to EDIT - feature request #7700: use TypeDesignationSetManager for cr...

Closed

Related to EDIT - feature request #9309: Improve handling of specimen without...

New

#### Associated revisions

##### Revision 30190957 - 10/22/2020 06:42 PM - Andreas Müller

ref #9262 fix fide and designated by handling for cdmlight export and TypeDesignationSetManager

##### Revision 9eab4a32 - 10/22/2020 07:14 PM - Andreas Müller

ref #9262 fix et. al position in cdmlight

##### Revision 0ad005e1 - 10/23/2020 11:55 AM - Andreas Müller

ref #9262 move designated by out of brackets

##### Revision 38cdb834 - 10/23/2020 03:25 PM - Andreas Müller

ref #9262 move lectotype citation for name types in stringify(NameTypeDesignation) method

##### Revision 742c4c12 - 10/23/2020 05:38 PM - Andreas Müller

ref #9262 add test for lecto-nametype with designated by and fide

##### Revision bdb70df7 - 10/23/2020 05:43 PM - Andreas Müller

ref #9262 javadoc link to ticket

##### Revision 085f44ab - 10/23/2020 08:48 PM - Andreas Müller

ref #9262 unify buildString methods in TypeDesignationSetManager and unify some minor formatting issues

**Revision ef013705 - 10/23/2020 10:49 PM - Andreas Müller**

ref #9262 fix failing test by adding lectotype to name type status in test terms

**Revision f9f1f4a3 - 10/23/2020 10:55 PM - Andreas Müller**

ref #9262 further cleanup citation handling in TypeDesignationSetManager

**Revision f36cf20b - 10/23/2020 11:00 PM - Andreas Müller**

ref #9262 further cleanup citation handling in TypeDesignationSetManager (cont.)

**Revision d5603cc5 - 10/29/2020 11:13 PM - Katja Luther**

ref #9262: add homotypic group strings

**Revision f5eab879 - 10/30/2020 10:47 AM - Katja Luther**

ref #9262: fix homotypical group name comparator

**Revision 5a8468de - 11/04/2020 03:15 PM - Katja Luther**

ref #9262: smaller changes

**Revision 6043d78e - 11/04/2020 04:20 PM - Andreas Kohlbecker**

ref #9262 adapting dataportal to lowercase typedesignation status

**Revision 9270c708 - 11/11/2020 03:36 PM - Katja Luther**

ref #9262: add homotypic group sortindex

**Revision 44bef338 - 11/18/2020 03:51 PM - Katja Luther**

ref #9262: add possibility to export condensed distribution string

**Revision c842935a - 11/18/2020 03:58 PM - Katja Luther**

ref #9262: add possibility to export condensed distribution string - new preference predicate

**Revision 5d62d265 - 11/18/2020 04:50 PM - Katja Luther**

ref #9262: minor

**Revision cb6764ac - 11/19/2020 01:32 PM - Katja Luther**

ref #9262: adapt cdmlight preference handling to renamed predicate

## History

---

**#1 - 10/22/2020 12:56 PM - Andreas Müller**

- File *picture597-1.png* added

- Description updated

**#2 - 10/22/2020 06:35 PM - Andreas Müller**

- Related to feature request #7700: use *TypeDesignationSetManager* for creating *typedesignation* string in *cdmlightExport* added

**#3 - 10/22/2020 06:49 PM - Andreas Müller**

- Description updated

**#4 - 10/22/2020 07:15 PM - Andreas Müller**

- Description updated

**#5 - 10/23/2020 10:56 AM - Andreas Müller**

- Description updated

**#6 - 10/23/2020 10:57 AM - Andreas Müller**

- Description updated

**#7 - 10/23/2020 11:58 AM - Andreas Müller**

- Description updated

**#8 - 10/23/2020 12:22 PM - Andreas Müller**

- Description updated

- % Done changed from 0 to 10

**#9 - 10/23/2020 12:24 PM - Andreas Müller**

- Description updated

**#10 - 10/26/2020 12:40 PM - Katja Luther**

- Description updated

**#11 - 10/27/2020 10:05 AM - Andreas Müller**

- Description updated

**#12 - 10/27/2020 10:12 AM - Andreas Müller**

- Description updated

**#13 - 10/27/2020 11:16 AM - Katja Luther**

- Description updated

**#14 - 10/28/2020 12:16 PM - Katja Luther**

- Description updated

**#15 - 11/02/2020 01:30 PM - Katja Luther**

- Description updated

**#16 - 11/02/2020 01:53 PM - Katja Luther**

- Description updated

**#17 - 11/02/2020 02:55 PM - Andreas Müller**

- Description updated

- % Done changed from 10 to 40

**#18 - 11/13/2020 09:05 AM - Katja Luther**

- Description updated

**#19 - 11/13/2020 09:06 AM - Katja Luther**

- Status changed from New to In Progress

**#20 - 11/13/2020 09:15 AM - Katja Luther**

- Description updated

**#21 - 11/13/2020 09:19 AM - Katja Luther**

- Description updated

**#22 - 11/13/2020 09:27 AM - Andreas Müller**

- Description updated

**#23 - 11/17/2020 10:19 AM - Katja Luther**

- Description updated

**#24 - 11/19/2020 01:53 PM - Katja Luther**

- Description updated

**#25 - 11/19/2020 01:56 PM - Katja Luther**

- Related to feature request #9309: Improve handling of specimen without titlecache in cdmLight added

**#26 - 11/19/2020 01:56 PM - Katja Luther**

- Description updated

- Status changed from In Progress to Resolved

**#27 - 11/19/2020 01:58 PM - Katja Luther**

all issues are implemented except the specimen without titleCache, therefore a new ticket exist ([#9309](#))

**#28 - 11/25/2020 10:48 PM - Andreas Müller**

so can we close the ticket?

**#29 - 11/30/2020 12:06 PM - Katja Luther**

- Status changed from Resolved to Closed

Andreas Müller wrote:

so can we close the ticket?

yes.

## Files

---

picture597-1.png	17.5 KB	10/22/2020	Andreas Müller
------------------	---------	------------	----------------