**Gunnhild syr sammen musit- og corema-dump-filer**

Ny branch i git hub: «specimen-type-merge»

1. gi corema-dump-filer nye navn (se tabell nedenfor – jeg har gjort dette manuelt. Eirik: automatiser i DownloadDump)
2. src\utils\stitchFiles.js eller stitchFiles\_musit.js – jeg kjører denne i konsollen med «node stitchFiles.js», så lages nye filer som leses i setCollection() i fileread.js
   1. hovedfil leses, legger alle data i objekter i array
   2. legger til items til objektene som én property med array med alle itemtypene. Legger videre til flere properties med arrays med item-uuid’er, item-datoer etc. Rekkefølgen på elementer i arrayene viser hvilke data som hører til hvilket item. Til slutt gjøres arrayene om til string med data separert med komma.
3. I object.js: for hvert item, legg til rad i tabellen for hvert item, og hvert items property
4. I paginateAndRender.js: list opp items i siste kolonne i tabellen med søketreff

**Dokumentasjon stitchFiles.js**

stithcFiles.js inneholder følgende funksjoner:

**readDumpFile(filename, callback)**

Beskrivelse: leser en corema- eller musitdumpfil og gjør om innholdet til JSON

Input: filename – Tekststreng, navn på dumpfila som skal leses

Output: callback-funksjon – sender JSON objektet til funksjonen som kaller denne

**readline.createInterface({}).on(line, function ()).on(close, function()**

Beskrivelse:

Leser en hoved-dumpfil og gjør om innholdet til JSON som så legges i en array med objekter, ett objekt for hvert individ.

Kaller readDumpFile() som leser andre dumpfiler og gjør om til JSON,

og legger disse dataene til objektene i den første arrayen.

Filer i dumpen

* fra corema; de som bare er i corema

|  |  |  |  |
| --- | --- | --- | --- |
| **Filnavn i dump** | **Brukes i koden 2.7.20** | **Skal hete** | **Inneholder** |
| occurrence.txt | x | birds\_occurrence.txt | Hoveddata for alle items |
| amplification.txt | x | dna\_birds\_amplification.txt | referanse til genetiske sekvenser |
| materialsample.txt | x | birds\_materialsample.txt | Item-type, konsentrasjon (på DNA-ekstrakt) og enhet for konsentrasjon |
| multimedia.txt | x | birds\_multimedia.txt | bildeinformasjon |
| permit.txt |  |  | tilhørende tillatelser; type, status og forklaring |
| preparation.txt | X | dna\_birds\_preparation.txt | informasjon om ekstrahert DNA; type (f.eks. “gDNA”), ekstraksjonsmetode, utøver og dato |
| preservation.txt | X | birds\_preservation.txt | hvordan item’et er oppbevart (f.eks. etanol, tørt, queens buffer) |
| resourcerelationship.txt |  | birds\_resourcerelationship.txt | relasjon mellom item’er og voucher i trad. samling, bold process id o.a. |

* samme for «mammals», «fishHerps» og «other»

Fra musit:

|  |  |  |  |
| --- | --- | --- | --- |
| **Filnavn etter prosessering i DownloadDump** | **Brukes i koden 2.7.20** | **Skal hete** | **Inneholder** |
| sopp\_occurrence.txt | x | sopp\_occurrence.txt | Hoveddata for alle objekter |

* samme for lav, karplanter, entomologi

Fra corema, de som har organismeinfo i musit:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Filnavn i dump** | **Brukes i koden 2.7.20** | **Skal hete** | **Inneholder** |  |
| occurrence.txt | x | dna\_fungi\_lichens\_occurrence.txt | Hoveddata for alle items | én linje per item |
| amplification.txt | x | dna\_fungi\_lichens \_amplification.txt | referanse til genetiske sekvenser | ubegrenset linjer per item |
| materialsample.txt | x | dna\_fungi\_lichens \_materialsample.txt | Item-type, konsentrasjon (på DNA-ekstrakt) og enhet for konsentrasjon | én linje per item |
| multimedia.txt |  |  | bildeinformasjon |  |
| permit.txt |  |  | tilhørende tillatelser; type, status og forklaring |  |
| preparation.txt | X | dna\_fungi\_lichens \_preparation.txt | informasjon om ekstrahert DNA; type (f.eks. “gDNA”), ekstraksjonsmetode, utøver og dato | én linje per item |
| preservation.txt | X | dna\_fungi\_lichens \_preservation.txt | hvordan item’et er oppbevart (f.eks. etanol, tørt, queens buffer) | én linje per item |
| resourcerelationship.txt | x | fungi\_lichens \_resourcerelationship.txt | relasjon mellom item’er og voucher i trad. samling, bold process id o.a. | ubegrenset linjer per items |

* samme for «entomology» og «plants»