

## Bilag B – Ønsket struktur af en aflevering

dl\_20190708\_rt1/ En aflevering lægges i mappe med navn dl\_yyyymmdd\_rtx . Alle filer for aviser fra 8. juli 2019 ligger i denne folder. Første aflevering navngives "rt1" (roundtrip 1). Senere fejlrettelser lægges i mapper rt2, rt3, etc.

↳ ekstrabladet/ Hver avistitel lægges i sin egen underfolder navngivet med det relevante avis\_id (se bilag A).

↳ pages/ Alle sider lægges i en underfolder ved navn 'pages'

↳ 20190708\_ekstrabladet\_section01\_page001.pdf

For hver avisside afleveres en PDF-fil jvf. krav 2.1.2

Filnavne sikres unikke ved at gentage metadata fra mappernavne

↳ 20190708\_ekstrabladet\_section01\_page001.xml

For hver avisside afleveres en XML-fil med metadata jvf. krav 2.1.8

↳ 20190708\_ekstrabladet\_section01\_page002.pdf

↳ 20190708\_ekstrabladet\_section01\_page002.xml

↳ ...

↳ articles/ Artikler lægges i en underfolder ved navn 'articles'

↳ 20190708\_ekstrabladet\_article\_001.pdf

For hver artikel afleveres en XML-fil med metadata jf. punkt 3.1.2

↳ 20190708\_ekstrabladet\_article\_001.xml

↳ ...

↳ politiken/

↳ pages/

↳ 20190708\_politiken\_section01\_page001.pdf

↳ 20190708\_politiken\_section01\_page001.xml

↳ 20190708\_politiken\_section01\_page002.pdf

↳ 20190708\_politiken\_section01\_page002.xml

↳ ...

↳ articles/

↳ 20160501\_politiken\_article\_001.xml

↳ ...

↳ checksums.txt En liste af checksummer af alle filer i formatet 'md5sum'

Se <https://en.wikipedia.org/wiki/Md5sum>

↳ transfer\_complete En tom fil som altid skrives som den sidste fil.

Når denne fil skrives er afleveringen afsluttet og komplet.