

# Spesifikasjon for beskrivelse av kvalitet på datasett

Versjon 1.0

# Innholdsfortegnelse

|                                                                                             |    |
|---------------------------------------------------------------------------------------------|----|
| Omfang og avgrensning .....                                                                 | 2  |
| Normative referanser .....                                                                  | 3  |
| Begreper og definisjoner .....                                                              | 4  |
| Standardisert metode for å beskrive kvalitet på datasett .....                              | 7  |
| Vedlegg A - Forslag til kvalitetsdimensjoner, kvalitetsdeldimensjoner og kvalitetsmål ..... | 8  |
| Vedlegg B - Forslag til Implementering i DCAT-AP-NO .....                                   | 18 |
| Implementering i DCAT-AP-NO .....                                                           | 18 |
| Kvalitetsdimensjon Nøyaktighet (accuracy) .....                                             | 20 |
| Kvalitetsdimensjon Kompletthet/Dekning (completeness/coverage) .....                        | 21 |
| Kvalitetsdeldimensjon Underdekning (undercoverage/omission) .....                           | 21 |
| Kvalitetsdeldimensjon Overdekning (overcoverage/commission) .....                           | 24 |
| Kvalitetsdeldimensjon Selektivitet (selectivity) .....                                      | 25 |
| Kvalitetsdeldimensjon Dublett (redundancy/duplicate) .....                                  | 25 |
| Kvalitetsdimensjon Aktualitet (currentness/timeliness) .....                                | 26 |
| Kvalitetsdimensjon Samsvar (compliance/conformity) .....                                    | 26 |
| Kvalitetsbeskrivelse I samsvar med (conforms to) .....                                      | 27 |
| Kvalitetsdimensjon Tilgjengelighet (availability) .....                                     | 27 |
| Kvalitetsdimensjon Relevans (relevancy) .....                                               | 28 |
| Ikke-kvantitativ/fritekst beskrivelse Bruksformål (specific usage) .....                    | 28 |
| Ikke-kvantitativ/fritekst beskrivelse Egnethet (usability) .....                            | 29 |
| Vedlegg C - URIer/navnerom som er brukt .....                                               | 31 |
| Vedlegg D - Referanser .....                                                                | 32 |
| Vedlegg E - Arbeidsgruppen .....                                                            | 33 |
| Beskrivelse av arbeidsgruppens oppgave .....                                                | 33 |
| Sammensetning av arbeidsgruppen .....                                                       | 33 |
| Arbeid i arbeidsgruppen .....                                                               | 34 |

**Status:** utgått

**Publisert:** september 2017

**Oppdatert:** 3. november 2020



Denne spesifikasjonen er utgått. Se [DQV-AP-NO](#) som implementerer kravene i denne spesifikasjonen.

# Omfang og avgrensning

Dokumentet inneholder en spesifikasjon til hvordan man skal beskrive kvalitet på datasett i en datakatalog.

Så langt det er mulig er det i dette dokumentet også tatt med veiledning/eksempler på hvordan denne spesifikasjonen i praksis kan brukes. Veiledningen/eksemplene er ikke normative.

I vedleggene til dette dokumentet er det tatt med forslag til hoveddimensjoner, deldimensjoner og kvalitetsmål av kvalitet på datasett, samt forslag til hvordan disse kan implementeres i standarden DCAT-AP-NO. Disse vedleggene er ikke normative.

Dokumentet inneholder ikke spesifikasjon eller veiledning til hvordan man skal holde god kvalitet på datasett, heller ikke hvordan man skal måle datakvalitet.

Spesifikasjonen er utarbeidet av en arbeidsgruppe (se [vedlegg E](#)) ledet av Kartverket ved Morten Borrebæk.

# Normative referanser

Difi, DCAT-AP-NO, Standard for beskrivelse av datasett og datakataloger, 11.10.2016.

W3C Data on the Web Best Practices: Data Quality Vocabulary (DQV), <sup>[1]</sup> W3C Working Group Note 15 December 2016, <https://www.w3.org/TR/vocab-dqv/>.

W3C Data on the Web Best Practices: Dataset Usage Vocabulary (DUV), <sup>[2]</sup> W3C Working Group Note 15 December 2016, <https://www.w3.org/TR/vocab-duv/>.

W3C Simple Knowledge Organization System (SKOS), <sup>[3]</sup> <http://www.w3.org/2004/02/skos/>.

W3C Web Annotation Data Model, W3C Recommendation 23 February 2017, <https://www.w3.org/TR/annotation-model/>.

---

[1] DQV (Data Quality Vocabulary) er i skrivende stund ikke en anbefaling fra W3C ennå, men kun et “arbeidsgruppenotat”.

[2] DUV (Dataset Usage Vocabulary) er i skrivende stund ikke en anbefaling fra W3C ennå, men kun et “arbeidsgruppenotat”.

[3] EU har i “Commission Implementing Decision (EU) 2017/1358 of 20 July 2017” besluttet å kunne bruke SKOS i offentlige anskaffelser. <http://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1500883656509&uri=CELEX:32017D1358>

# Begreper og definisjoner

Med “kvalitet” (f.eks. i ordene “kvalitetsdimensjon” og “kvalitetsmål”) menes det i dette dokumentet “datakvalitet”, “kvalitet på datasett” og lignende, men ikke “kvalitet” generelt.

Med “mål” (f.eks. i ordet “kvalitetsmål”) menes det i dette dokumentet “måling”/”beregning” (*metric*), men ikke “målsetning”.

## *Kvalitetsdimensjon:*

- Anbefalt term: kvalitetsdimensjon
- Alternativ term: dimensjon
- Definisjon: kriterier relevant for evaluering av kvalitet på datasett
  - Engelsk original: represents criteria relevant for assessing quality.
    - Kilde til definisjon: [dqv:Dimension i W3C DQV](#)
- Merknad: Tilsvare [dqv:Dimension](#)
- Eksempler: [Nøyaktighet](#), [Kompletthet/Dekning](#)

## *Kvalitetsdeldimensjon:*

- Anbefalt term: kvalitetsdeldimensjon
- Alternativ term: deldimensjon
- Definisjon: en mer detaljert inndeling av en [Kvalitetsdimensjon](#):
- Merknad: Datakvalitetsvokabularet DQV fra W3C har ikke definert kvalitets\*del\*dimensjon som et eget begrep (klasse), men kun kvalitetsdimensjon.
  - For å uttrykke kvalitetsdeldimensjon, anbefaler vi å bruke “skos:broader”
  - En kvalitetsdeldimensjon inneholder dermed alltid en “skos:broader”-relasjon til den kvalitetsdimensjonen som denne deldimensjonen tilhører.
  - Arbeidsgruppen har hittil identifisert behov for kun et nivå av kvalitetsdeldimensjon.
- Merknad: Tilsvare [dqv:Dimension](#)
- Eksempler: [Kvalitetsdeldimensjon Underdekning \(undercoverage/omission\)](#) og [Kvalitetsdeldimensjon Overdekning \(overcoverage/commission\)](#) som kvalitetsdeldimensjoner under kvalitetsdimensjonen [Kompletthet/Dekning](#).

## *Kvalitetsmål:*

- Anbefalt term: kvalitetsmål
- Alternativ term: beregningsmetode for et kvalitetsmål, definisjon av et kvalitetsmål
- Definisjon: definisjon/beskrivelse av måten å beregne/oppgi kvalitetsmål
  - Engelsk original: Represents a standard to measure a quality dimension.
    - Kilde til definisjon: [dqv:Metric](#)
- Merknad: Tilsvare [dqv:Metric](#)
- Merknad: I henhold til DQV skal selve måleverdi oppgis ved å bruke [dqv:QualityMeasurement](#)

- Engelsk original: An observation (instance of `dqv:QualityMeasurement`) assigns a value in a given unit to a Metric.
- Eksempler: [Andel manglende elementer](#) og [Antall manglende elementer](#) som kvalitetsmål langs kvalitetsdeldimensjonen [Kvalitetsdeldimensjon Underdekning \(undercoverage/omission\)](#).

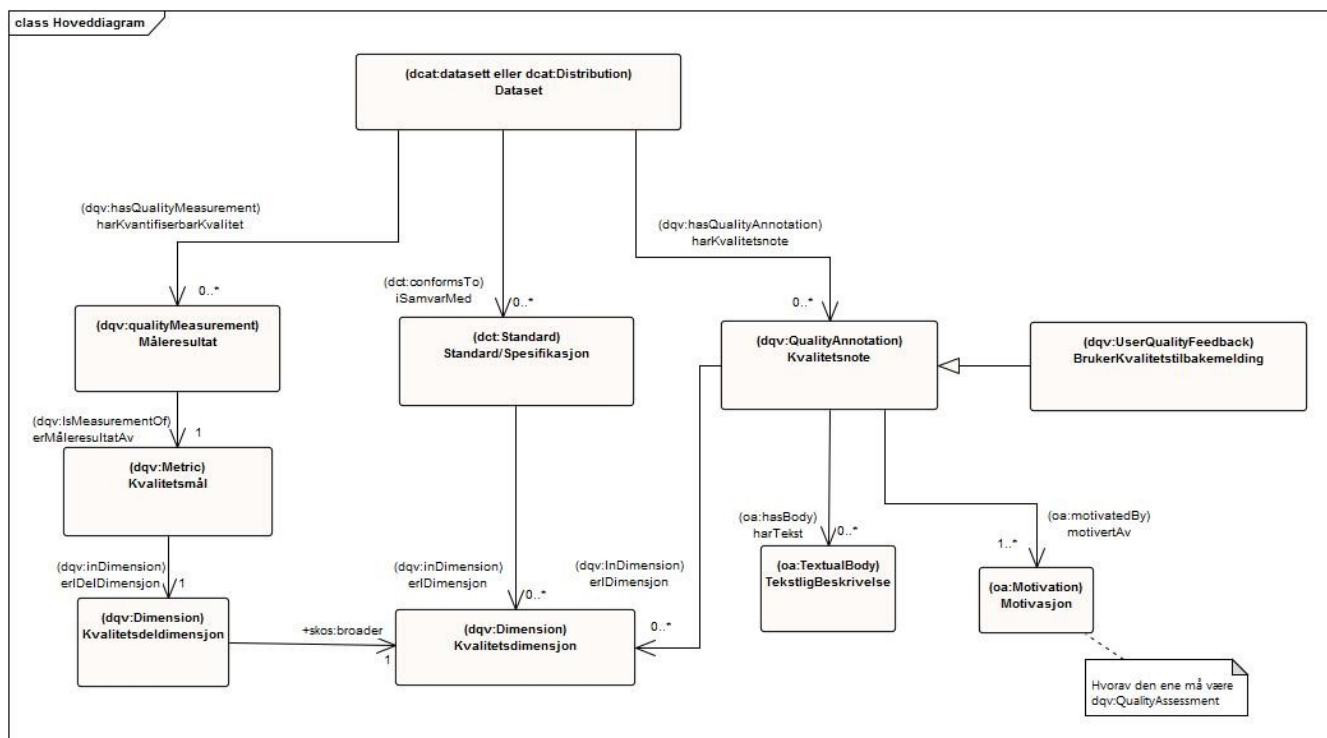
#### *Måleresultat:*

- Anbefalt term: måleresultat
- Alternativ term: kvalitetsmåleresultat
- Definisjon: konkret resultat av [kvalitetsmål](#)
  - Engelsk original: Represents the evaluation of a given dataset (or dataset distribution) against a specific quality metric.
    - Kilde til definisjon: [dqv:QualityMeasurement](#)
- Merknad: Tilsvarende [dqv:QualityMeasurement](#)
- Eksempler: Verdi “0,10” som resultat av “Andel manglende elementer”; verdi “10” som resultat av “Antall manglende elementer”.

#### *Pre-definerte kvalitetsmål:*

- Anbefalt term: pre-definert kvalitetsmål
- Definisjon: [kvalitetsmål](#) (`dqv:Metric`) som finnes tilgjengelig for gjenbruk, og som kan refereres ved en unik identifikator (URI)
- Merknad: Arbeidsgruppen foreslår å etablere en felles oversikt over pre-definerte kvalitetsmål.
- Eksempler: Annex D i ISO 19157:2013 (E) “List of standardized data quality measures” inneholder en rekke pre-definerte kvalitetsmål.

Nedenfor er en forenklet UML-modell som viser de viktigste begrepene/klasse og relasjoner mellom disse. Tegningen illustrerer også forslag til implementering i DCAT-AP-NO, som er informativt (se [vedlegg B](#) for mer detaljert beskrivelse av forslaget). Det er valgt å bruke norske navn på klassene og rollene, med navn på de vokabularer som skal brukes i implementasjonen i parentes. For å få en full oversikt over modellen må en også se på spesialiseringer og subtyper av andre vokabularer som benyttes i dqv.



Figur 1: Kvalitetsbeskrivelser ved å bruke DQV.

Figuren under illustrerer kvalitetsdimensjoner og kvalitetsdeldimensjoner og relasjon mellom disse, som begge er spesialiseringer av `dqv:Dimension`. Forskjellen er at en kvalitetsdeldimensjon peker til en kvalitetsdimensjon ved hjelp av `skos:broader`. Det gjøres oppmerksom på at de enkelte kvalitetsdimensjoner og kvalitetsdeldimensjoner som er vist i figuren er kun eksempler og ikke normative.

[Eksempler kvalitetsdimensjoner] | *Eksempler\_kvalitetsdimensjoner.png*

Figur 2: Eksempler på kvalitetsdimensjoner og kvalitetsdeldimensjoner.

NB! De enkelte kvalitetsdimensjoner og kvalitetsdeldimensjoner er i dette dokumentet informative eksempler, mens måten å relatere en kvalitetsdeldimensjon til den tilhørende kvalitetsdimensjon er normativ (ved å bruke `skos:broader`).



# Standardisert metode for å beskrive kvalitet på datasett

Arbeidsgruppen foreslår følgende når det gjelder beskrivelse av kvalitet på datasett i en datakatalog:

1. Det er anbefalt (= “skal, når det finnes”) å beskrive kvalitet på datasett i en datakatalog.
2. Bruk [W3C DQV](#) som et standardisert vokabular til å beskrive kvalitet på datasett.
3. Kvalitet på datasett beskrives langs et fåtalls kvalitetsdimensjoner (se [vedlegg med forslag til dimensjonene](#)) - i listen under betyr parentes valgfritt:
  - a. Kvantiserbar kvalitet beskrives i form av “Kvalitetsdimensjon” → “Kvalitetsdeldimensjon” → “Kvalitetsmål” → “Måleresultat”
  - b. Ikke-kvantiserbar kvalitet beskrives i form av (“Kvalitetsdimensjon” →) “Kvalitetsnote”
  - c. Kvalitet som er i samsvar med gitt(e) standarder/spesifikasjoner beskrives i form av (“Kvalitetsdimensjon” →) “Standard/Spesifikasjon”
  - d. Brukertilbakemeldinger knyttet til kvalitet beskrives i form av (“Kvalitetsdimensjon” →) “Brukerkvalitetstilbakemelding” Kvantiserbar kvalitet beskrives dessuten ved å referere <sup>[1]</sup> til et pre-definert kvalitetsmål. Definisjonene skal være i henhold til [W3C DQV](#).

Se også datamodellen i kapittelet foran.

Resten av dokumentet inneholder:

- Et informativt vedlegg med forslag fra arbeidsgruppen til kvalitetsdimensjoner, kvalitetsdeldimensjoner og kvalitetsmål.
- Et informativt vedlegg med forslag til implementering av kvalitetsmål i DCAT-AP-NO, med eksempler på beskrivelser/definisjoner uttrykt ved hjelp av DQV.
- Et informativt vedlegg med URIer/navnerom som er brukt i dette dokumentet.
- Et informativt vedlegg med referanser som ikke er nevnt under normative referanser.
- Et informativt vedlegg med informasjon om arbeidsgruppen som har utarbeidet denne spesifikasjonen.

[1] For å kunne referere til pre-definerte kvalitetsmål, foreslås det å etablere en “Felles oversikt over definisjoner av kvalitetsmål” (dermed også et kontrollert vokabular). Ved behov kan den enkelte etaten definere nye kvalitetsmål, som så legges inn i felles oversikt, for gjenbruk av andre.

# Vedlegg A - Forslag til kvalitetsdimensjoner, kvalitetsdeldimensjoner og kvalitetsmål

Informativ del av dette dokumentet.

Arbeidsgruppen foreslår å starte med følgende kvalitetsdimensjoner <sup>[1]</sup>:

1. Nøyaktighet (Accuracy, [ISO/IEC 25012](#), DQV Kap. 7.2)
2. Kompletthet/Dekning
  - a. Dekning er bredere enn Kompletthet (Completeness, [ISO/IEC 25012](#), DQV Kap. 7.2). Dekning inkluderer bl.a. “overdekning”. Termen “kompletthet” brukes allikevel her fordi den er allerede i bruk av flere fagmiljøer.
3. Aktualitet (Currentness, [ISO/IEC 25012](#), DQV Kap. 7.2)
4. Samsvar (Compliance, [ISO/IEC 25012](#), DQV Kap. 7.2)
5. Tilgjengelighet (Availability, [ISO/IEC 25012](#), DQV Kap. 7.2)
6. Relevans (Relevancy, <https://www.w3.org/TR/vocab-dqv/#bib-ZaveriEtAl>, DQV Kap. 7.3)

Tabellen under inneholder arbeidsgruppens forslag til kvalitetsdimensjoner, kvalitetsdeldimensjoner og kvalitetsmål som bør tas med i Felles datakatalog. Når det senere arbeides med å pre-definere kvalitetsdimensjoner, kvalitetsdeldimensjoner og kvalitetsmål, vil forslagene i denne tabellen gås gjennom nøye før de endelig fastsettes.

| Kilde/Etat (med ref) samt evt gruppering) | Kvalitetsdimensjon     | Kvalitetsdeldimensjon | Kvalitetsmål | Beskrivelse                                                                                                                                                               |
|-------------------------------------------|------------------------|-----------------------|--------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| DQV                                       | Nøyaktighet (Accuracy) |                       |              | <i>DQV: The degree to which data has attributes that correctly represent the true value of the intended attribute of a concept or event in a specific context of use.</i> |
| KARTV                                     |                        |                       |              | nøyaktighet av kvantitative egenskaper, riktighet av ikke-kvantitative egenskaper og objektenes klassifisering og relasjoner                                              |

| Kilde/Etat (med ref) samt evt gruppering) | Kvalitetsdimensjon | Kvalitetsdeldimensjon                 | Kvalitetsmål | Beskrivelse                                                                              |
|-------------------------------------------|--------------------|---------------------------------------|--------------|------------------------------------------------------------------------------------------|
|                                           |                    | Klassifikasjonsriktighet              |              | sammenligning mellom anvendt klassifisering og virkelighet                               |
|                                           |                    | Ikke-kvantitativ egenskapsnøyaktighet |              | hvor nær ikke-kvantitative verdier er sanne verdier eller verdier akseptert som sanne    |
|                                           |                    | Kvantitativ egenskapsnøyaktighet      |              | hvor nær kvantitative verdier er sanne verdier eller verdier akseptert som sanne         |
| KARTV                                     |                    |                                       |              | hvor godt stedfestingen til et objekt samsvarer med virkeligheten/fasit                  |
|                                           |                    | Absolutt stedfestingsnøyaktighet      |              | hvor nær stedfestet posisjon er sann posisjon eller posisjon akseptert som sann          |
|                                           |                    | Nabonøyaktighet                       |              | hvor bra stedfestet posisjon samsvarer med andre stedfestede posisjoner                  |
|                                           |                    | Posisjonsnøyaktighet i rasterdata     |              | hvor bra posisjon i raster samsvarer med sann posisjon eller posisjon akseptert som sann |

| Kilde/Etat (med ref) samt evt gruppering) | Kvalitetsdimensjon | Kvalitetsdeldimensjon    | Kvalitetsmål                          | Beskrivelse                                                                                                   |
|-------------------------------------------|--------------------|--------------------------|---------------------------------------|---------------------------------------------------------------------------------------------------------------|
| KARTV                                     |                    | Stedfestingspålitelighet |                                       | uttrykk for hvor sterkt mulig gjenværende grove feil i materialet for stedfestingen påvirker slutt-resultatet |
| KARTV                                     |                    |                          |                                       | kvaliteten til egenskaper som definerer tid eller tidsavhengigheter mellom objekter                           |
|                                           |                    | Tidsnøyaktighet          |                                       | hvor nær angitte tidsverdier er sanne verdier eller verdier akseptert som sanne                               |
| SSB (BLUE-ETS)                            |                    |                          |                                       | Hvor nært objekter og variable ligger de korrekte verdiene og i hvilken grad data er til å stole på           |
|                                           |                    | Identifiserbarhet        | Antall (med problem). Andel i prosent | Objekter med ugyldige identifikasjonsnøkler (feil syntaks)                                                    |
|                                           |                    | Autensitet               | --”--                                 | Objekter med gale identifikasjonsnøkler (men korrekt syntaks)                                                 |
|                                           |                    | Konsistens               | --”--                                 | Om objektene er innbyrdes konsistente                                                                         |
|                                           |                    | Mistenkelige verdier     | Antall                                | Knyttet til objekter                                                                                          |

| Kilde/Etat (med ref) samt evt gruppering) | Kvalitetsdimensjon     | Kvalitetsdeldimensjon                                                                                                        | Kvalitetsmål          | Beskrivelse                                                                                                                                                                     |
|-------------------------------------------|------------------------|------------------------------------------------------------------------------------------------------------------------------|-----------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                           |                        | Feil knyttet til variable, f.eks. Validitet, rapporteringsfeil, registreringsfeil, bearbeidingsfeil og mistenkelige verdier) | Varians/standardavvik | Knyttet til variable, her bare summarisk gjengitt. Kan detaljere dette                                                                                                          |
| SSB (CoP)                                 |                        |                                                                                                                              |                       | Hvor langt fra sann verdi ligger resultatet? Kan man stole på det?                                                                                                              |
|                                           |                        | Utvalgsfeil                                                                                                                  | Varians/standardavvik | Feil i statistikk grunnet at den eventuelt er basert på utvalg]                                                                                                                 |
|                                           |                        | Andre feil (dekning, frafall, målefeil, bearbeidingsfeil, modellfeil)                                                        | Varians/standardavvik | En rekke feiltyper utgjør dimensjonene, her bare summarisk gjengitt                                                                                                             |
| DQV                                       | Fullstendighet/Dekning |                                                                                                                              |                       | DQV completeness): The degree to which subject data associated with an entity has values for all expected attributes and related entity instances in a specific context of use. |
| KARTV                                     |                        |                                                                                                                              |                       | beskrivelse av hvilke enheter som er med i et datasett i forhold til de som burde vært med.                                                                                     |
|                                           |                        | (Manglende data)                                                                                                             |                       | data som mangler i et datasett                                                                                                                                                  |

| Kilde/Etat (med ref) samt evt gruppering) | Kvalitetsdimensjon               | Kvalitetsdeldimensjon | Kvalitetsmål       | Beskrivelse                                                                                                      |
|-------------------------------------------|----------------------------------|-----------------------|--------------------|------------------------------------------------------------------------------------------------------------------|
|                                           |                                  | (Overskytende data)   |                    | data som ikke skal være i et datasett                                                                            |
| SSB                                       |                                  |                       |                    | Angir om innholdet i datasettet er fullstendig i forhold til SSBs bruk, at det dekker akkurat det det skal dekke |
|                                           |                                  | <b>Underdekning</b>   | Antall eller andel | Manglende enheter i datasettet                                                                                   |
|                                           |                                  | <b>Overdekning</b>    | Antall eller andel | Forekomst av ikke-enheter i datasettet                                                                           |
|                                           |                                  | <b>Selektivitet</b>   | Tekst (forklar)    | Datasettet inneholder bare deler av den statistiske populasjonen                                                 |
|                                           |                                  | <b>Dubletter</b>      | Antall eller andel | Forekomst av dubletter (enheter som er registrert flere ganger)                                                  |
| SSB (CoP)                                 |                                  | Se Accuracy           |                    | Statistikk som ikke bygger på fullstendige data vil være unøyaktig og ikke til å stole på - Se ellers Accuracy   |
| <b>DQV</b>                                | <b>Aktualitet (Currenttness)</b> |                       |                    | <i>DQV: The degree to which data has attributes that are of the right age in a specific context of use.</i>      |

| Kilde/Etat (med ref) samt evt gruppering) | Kvalitetsdimensjon | Kvalitetsdeldimensjon                                                                           | Kvalitetsmål                   | Beskrivelse                                                                                                                                                |
|-------------------------------------------|--------------------|-------------------------------------------------------------------------------------------------|--------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------|
| SSB (BLUE-ETS)                            |                    | <b>Aktualitet</b>                                                                               | Tid (dager eller uker)         | Tid mellom slutten av kildens referansetidspunkt og SSB kan bruke data                                                                                     |
|                                           |                    | <b>Punktlighet</b>                                                                              | Andel som er punktlig. Prosent | Mulig forsinkelse mellom lovet og realisert tidspunkt når SSB kan få data                                                                                  |
|                                           |                    | <b>Tidsdifferanse</b>                                                                           | Tid (dager eller uker)         | Tid fra slutten av kildens referansetid til SSB konkluderer med at vi kan starte jobben med data, pga. etterslep i registeret, f.eks. sene flyttemeldinger |
|                                           |                    | *Enheters dynamikk og variables stabilitet<br>*(Dynamics of objects and stability of variables) | Tekst                          | Bla. endringer i koder mellom referansetidspunkt og SSBs bruk                                                                                              |
| SSB (CoP)                                 |                    | <b>Aktualitet</b>                                                                               | Tid (dager eller uker)         | Tid fra slutten av statistikkens referanseperiode til den publiseres                                                                                       |
|                                           |                    | <b>Punktlighet</b>                                                                              | Andel som er punktlig. Prosent | Avvik fra publiseringstidspunkt som skal være varslet 3 måneder på forhånd                                                                                 |

| Kilde/Etat (med ref) samt evt gruppering) | Kvalitetsdimensjon          | Kvalitetsdeldimensjon                                                           | Kvalitetsmål | Beskrivelse                                                                                                                                                                                |
|-------------------------------------------|-----------------------------|---------------------------------------------------------------------------------|--------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <i>DQV</i>                                | <i>Samsvar (Compliance)</i> |                                                                                 |              | <i>DQV: The degree to which data has attributes that adhere to standards, conventions or regulations in force and similar rules relating to data quality in a specific context of use.</i> |
| KARTV                                     |                             | <b>ConformanceResult</b>                                                        |              | A conformance result is the outcome of comparing the value or set of values obtained from applying a                                                                                       |
| SSB (BLUE-ETS)                            |                             |                                                                                 |              | Måler hvorvidt kilden kan gi god statistikk kombinert med andre kilder                                                                                                                     |
|                                           |                             | <b>Sammenlignbarhet av objekter</b><br>(Comparability and alignment of objects) |              | Dreier det seg om de samme objektene ved kobling av datasett?                                                                                                                              |
|                                           |                             | *Koblingsvariabel<br>*(Linking variable)                                        |              | Eventuelle problemer med koblingsvariabel                                                                                                                                                  |
|                                           |                             | <b>Sammenlignbarhet av variable</b><br>(Comparability of variables)             |              | Er variablene definert eller gruppert på samme måte?                                                                                                                                       |



| Kilde/Etat (med ref) samt evt gruppering) | Kvalitetsdimensjon                    | Kvalitetsdeldimensjon    | Kvalitetsmål                                        | Beskrivelse                                                                                                                                                                                      |
|-------------------------------------------|---------------------------------------|--------------------------|-----------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| SSB (CoP)                                 |                                       | <b>Sammenheng</b>        | Tekst, blant annet om bruk av standard grupperinger | Henger statistikken sammen med annen statistikk på samme eller tilgrensende områder? Brukes f.eks. Samme begreper/definisjoner, grupperinger og andre internasjonale eller nasjonale standarder? |
|                                           |                                       | <b>Sammenlignbarhet</b>  | Tekst                                               | Kan resultatene sammenlignes over tid og sted/geografi?                                                                                                                                          |
| <b>DQV</b>                                | <b>Tilgjengelighet (Availability)</b> |                          |                                                     | <i>DQV: The degree to which data has attributes that enable it to be retrieved by authorized users and/or applications in a specific context of use.</i>                                         |
| SSB (BLUE-ETS)                            |                                       |                          |                                                     | Teknisk brukbarhet av datasettet og data i settet                                                                                                                                                |
|                                           |                                       | Lesbarhet                | Tekst                                               | Om filen kan leses (ikke problemer med formater, ødelagt fil, uvanlig karaktersett, umulig å dekode)                                                                                             |
|                                           |                                       | Samsvar i fildeklarasjon | Andel?                                              | Manglende metadata, avvik fra beskrivelse                                                                                                                                                        |

| Kilde/Etat (med ref) samt evt gruppering) | Kvalitetsdimensjon          | Kvalitetsdeldimensjon | Kvalitetsmål | Beskrivelse                                                                                                                                   |
|-------------------------------------------|-----------------------------|-----------------------|--------------|-----------------------------------------------------------------------------------------------------------------------------------------------|
|                                           |                             | Konverterbarhet       | Tekst        | Feil som umuliggjør konvertering til de formater SSB bruker                                                                                   |
| SSB (CoP)                                 |                             |                       |              | Om statistikken er lett tilgjengelig og forklart for brukerne                                                                                 |
|                                           |                             | Tilgjengelighet       | Tekst        | Lett tilgjengelig og godt presentert statistikk: Hvor finnes den, henvisninger, bla. Til Statistikkbank. Eksempel på god tilgjengelighet: API |
|                                           |                             | Klarhet               | Tekst        | Dokumentasjon og metadata, eks. "Om statistikken" og annen dokumentasjon                                                                      |
| <b>DQV</b>                                | <b>Relevans (Relevancy)</b> |                       |              | <i>DQV: Relevancy refers to the provision of information which is in accordance with the task at hand and important to the users' query.</i>  |
| KARTV                                     |                             |                       |              | det som ikke er dekket av øvrige kategorier og aggregert kvalitet basert på flere kvalitetselementer                                          |

| Kilde/Etat (med ref) samt evt gruppering) | Kvalitetsdimensjon | Kvalitetsdeldimensjon                  | Kvalitetsmål             | Beskrivelse                                                                                                   |
|-------------------------------------------|--------------------|----------------------------------------|--------------------------|---------------------------------------------------------------------------------------------------------------|
|                                           |                    | <b>Egnethet</b>                        |                          | det som ikke er dekket av øvrige kategorier og aggregert kvalitet basert på flere kvalitetselementer          |
|                                           |                    | <b>AggregertKvalitet</b>               |                          | Summering av kvalitet fra andre kvalitetselementer .                                                          |
| KARTV                                     |                    |                                        |                          | provenance, source(s) and production process(es) used in producing a resource                                 |
|                                           |                    | <b>Lineage</b>                         |                          | Beskrivelse av hvordan produktet er fremkommet, og beskriver kilde(r) og ulike skritt i produksjonsprosessen. |
| KARTV                                     |                    | <b>Målestokksfaktor</b><br>Denominator |                          | Målestokksfaktor                                                                                              |
| KARTV                                     |                    | <b>Bruk</b>                            |                          | brief description of the resource and/or resource series usage.                                               |
| SSB (CoP)                                 |                    | <b>Relevans</b>                        | <i>Treff på websider</i> | Er statistikken relevant for brukeren? Også omtale av brukerundersøkelser og rutiner for brukerkontakt        |

[1] Med noe ulik ordvalg, er dimensjonene også definert i Eurostats oppdaterte definisjonsdatabase [RAMON](#), som i stor grad er basert på SDMX som er referert til i [StatDCAT-AP](#).

# Vedlegg B - Forslag til Implementering i DCAT-AP-NO

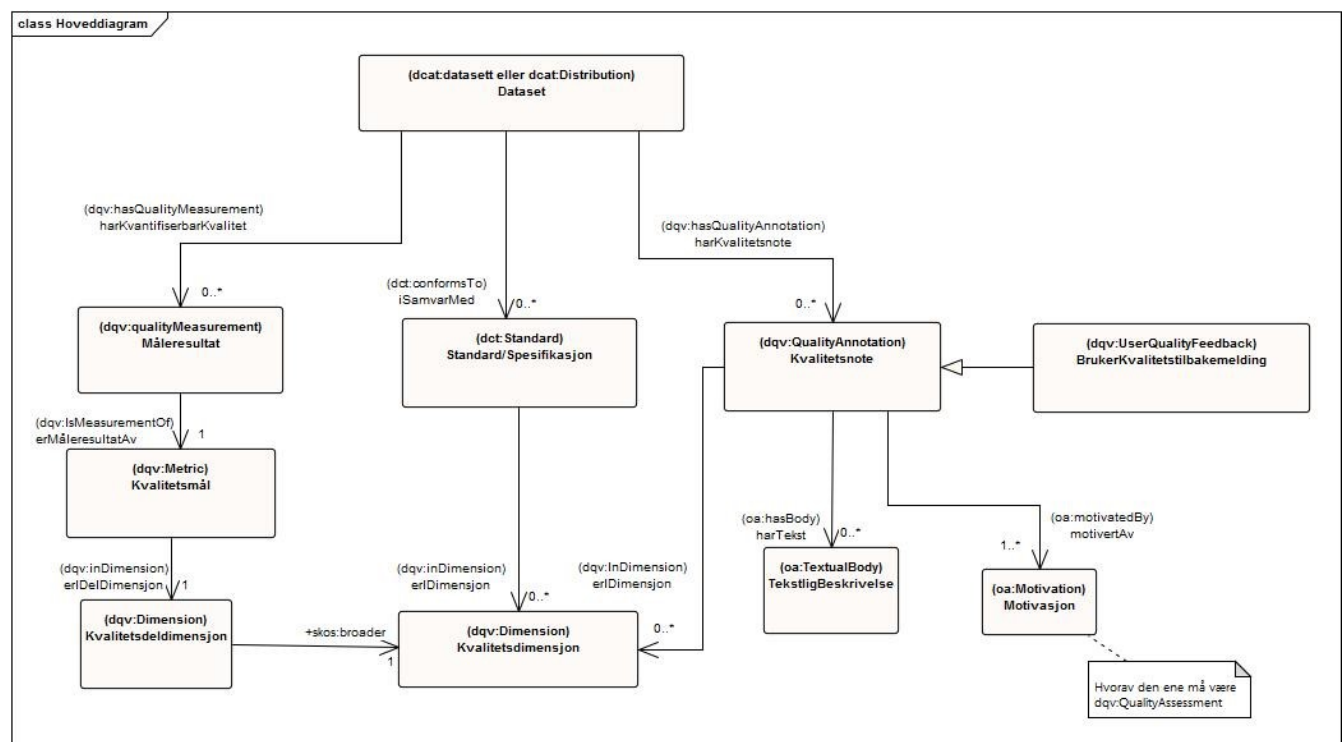
Informativ del av dette dokumentet.

Dette vedlegget inneholder først forslag til hvordan beskrivelse av kvalitet på datasett skal kunne implementeres i DCAT-AP-NO. Vedlegget inneholder deretter eksempler på hvordan selve beskrivelsene kan uttrykkes ved å bruke DQV. Kapittel [Kompletthet/Dekning](#) med [Kvalitetsdeldimensjon Underdekning \(undercoverage/omission\)](#) og kvalitetsmålene der under, viser et relativt komplett eksempel på hva som trengs for å beskrive “Dekning” ved å bruke DQV.

Navnerom (*namespaces*) som er brukt i eksemplene er ikke kvalitetssikret. Navnerom “xxx” er brukt der vi foreløpig ikke fant relevante etablerte vokabular (og dermed bør det opprettes norske vokabularer før de eventuelt kommer inn i relevante internasjonale vokabularer).

## Implementering i DCAT-AP-NO

I løsningsforslaget til implementering i DCAT-AP-NO baserer vi oss på prinsippene for RDFS-modellering. En står derfor fritt til å legge til egenskaper som ikke er omtalt i spesifikasjonen. Størst mulig grad av gjenbruk av eksisterende vokabularer er en selvsagt forutsetning.



Gjengivelse av figur 1 Kvalitetsbeskrivelser ved å -bruke DQV.

Jf. figuren ovenfor som er gjengitt fra kapittel [Begreper og definisjoner](#), foreslår vi følgende:

- Beskrivelse av kvantiserbar kvalitet implementeres i DCAT-AP-NO på følgende måte:
  - Det føyes til et nytt felt i DCAT-AP-NO, med:
    - URI: [dqv:hasQualityMeasurement](#)

- Range: [dqv:QualityMeasurement](#)
- Fritekst-kommentar til et kvantitativt måleresultat beskrives ved å bruke [skos:note](#)
- Beskrivelse av kvalitet som er i samsvar med gitt(e) standard(er)/spesifikasjon(er) implementeres i DCAT-AP-NO på følgende måte:
  - Bruk eksisterende felt i DCAT-AP-NO v.1.1, [Datsett: i samsvar med](#)
  - Der relevant, oppgis hvilke(n) kvalitetsdimensjon(er) standarden/spesifikasjonen dekker, ved å bruke [dqv:inDimension](#).
- Kvalitetsnote (for ikke-kvantiserbar kvalitet) implementeres i DCAT-AP-NO på følgende måte:
  - Det føyes til et nytt felt i DCAT-AP-NO, med:
    - URI: [dqv:hasQualityAnnotation](#)
    - Range: [dqv:QualityAnnotation](#)
  - Selve tekstlig beskrivelse inkluderes (ved å bruke [oa:hasBody](#)) i [dqv:QualityAnnotation](#), som en “body” (som en [oa:TextualBody](#))
  - I henhold til DQV skal det i/fra [dqv:QualityAnnotation](#) oppgis motivasjon (ved å bruke [oa:motivatedBy](#)) hvorav [dqv:QualityAssessment](#) er minimum å oppgi som motivasjon.
  - Der relevant, oppgis hvilke(n) kvalitetsdimensjon(er) kvalitetsnoten dekker, ved å bruke [dqv:inDimension](#).
- Brukertilbakemeldinger relatert til kvalitet på datasett implementeres i DCAT-AP-NO på følgende måte:
  - Kvalitetsrelaterte brukertilbakemeldinger angis ved å bruke samme feltet som foreslått under forrige kulepunkt for Kvalitetsnote, men med [dqv:UserQualityFeedback](#) som range, dvs.:
    - URI: [dqv:hasQualityAnnotation](#)
    - Range: [dqv:UserQualityFeedback](#)
  - [dqv:UserQualityFeedback](#) er i henhold til DQV en subklasse av [dqv:QualityAnnotation](#), og arver dermed egenskapene fra [dqv:QualityAnnotation](#) for å håndtere selve tekstlig beskrivelse og ev. angivelse av kvalitetsdimensjon(er) - jf. kulepunktet ovenfor om Kvalitetsnote.
  - For brukertilbakemeldinger skal det oppgis motivasjon utover [dqv:qualityAssessment](#) i tråd med [vokabular for motivasjon og formål](#) definert i W3Cs Web Annotations Data model. Merk at en her definerer om brukertilbakemeldingen for eksempel er et spørsmål eller et svar (fra andre brukere eller fra utgiver selv).
  - Vi anbefaler (ev. vurderer å gjøre obligatorisk senere) å oppgi skaper/avsender og dato for brukertilbakemeldinger ved å benytte [dct:creator](#) og [dct:created](#)

Eksempel på brukertilbakemelding:

```
:myDataset
  a dcat:dataset ;
  dqv:hasQualityAnnotation :userFeedback .
```

```

:userFeedback
  a dqv:UserQualityFeedback ;
  dct:creator "Donald Duck" ;
  dct:created "2017-09-13"^^xsd:date ;
  oa:hasBody :usabilityComment ;
  oa:motivatedBy dqv:qualityAssessment, oa:commenting .

:usabilityComment
  a oa:TextualBody ;
  rdf:value "Dette datasettet er uegnet for ..." ;
  dct:language "nb" ;
  dct:format "text/plain" .

```

[KlasseneIDQV] | *KlasseneIDQV.png*

*Figur 3: Datamodell som viser de mest relevante klassene i DQV (kilde: [W3C DQV](#)).*

Jf. figuren ovenfor som er kopiert fra W3C DQV, har arbeidsgruppen i første omgang ikke valgt å standardisere/spesifisere implementering av QualityPolicy og Prov (provenance) som også er i DQV-modellen ovenfor. Dette fordi det i denne omgangen ikke er identifisert konkrete behov/brukerhistorier knyttet til disse. Dette utelukker ikke mulighet for den enkelte virksomheten/sektoen å beskrive kvalitet i form av QualityPolicy og Prov (provenance) slik DQV har definert det. Informer gjerne Difi om erfaringer på dette.

Arbeidsgruppen antar at det vil være behov for å beskrive kvalitet både knyttet til datasett (dcat:dataset) og for distribusjoner (dcat:distribution). Arbeidsgruppen anbefaler at kvalitetsbeskrivelser knyttes til datasett (dcat:dataset) i størst mulig grad. Unntaket er når beskrivelsene eksplisitt gjelder leveranse-kvalitet, for eksempel beskrivelse av tilgjengelighet, aktualitet eller "conformance" for det ulike leveransene (distribusjonene) av datasettet.

## Kvalitetsdimensjon Nøyaktighet (accuracy)

**Merknad:** Arbeidsgruppen fant ut at dette var en av de kvalitetsdimensjonene som det allerede er etablert flere parallelle (internasjonale) kvalitetsdeldimensjoner og kvalitetsmål, som også til dels er nokså fagspesifikke. Arbeidsgruppen valgte derfor ikke å bruke tid på å gå dypere langs denne dimensjonen med å eksemplifisere mer i form av kvalitetsdeldimensjoner og kvalitetsmål.

### Definsjon:

- I hvilken grad datasettet korrekt representerer virkeligheten, for en spesifikk brukskontekst.
  - Engelsk original: The degree to which data has attributes that correctly represent the true value of the intended attribute of a concept or event in a specific context of use.
  - Kilde: [W3C DQV](#).

### Uttrykt i DQV:

```

@prefix dqv: <https://www.w3.org/TR/vocab-dqv/> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .

```

```
:accuracy
  a dqv:Dimension ;
  skos:prefLabel "accuracy"@en ;
  skos:prefLabel "nøyaktighet"@nb ;
  skos:definition "the degree to which data has attributes that correctly represent
the true value of the intended attribute of a concept or event in a specific context
of use"@en .
```

## Kvalitetsdimensjon Kompletthet/Dekning (completeness/coverage)

**Merknad:** Arbeidsgruppen fant ut at dette er en av de kvalitetsdimensjonene som det er mulig å bli enig om noen få felles kvalitetsdeldimensjoner og der under kvalitetsmål.

**Merknad:** “Dekning” er bredere enn “Kompletthet” (completeness fra [ISO/IEC 25012](#)). “Dekning” inkluderer bl.a. “Overdekning”. Termen “kompletthet” er allikevel tatt med fordi den allerede er tungt brukt i fagmiljøene.

### Definisjon:

- I hvilken grad datasettet inneholder forventede opplysninger, for en spesifikk brukskontekst.

### Uttrykt i DQV:

```
@prefix dqv: <https://www.w3.org/TR/vocab-dqv/> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .

:coverage
  a dqv:Dimension ;
  skos:prefLabel "coverage"@en ;
  skos:prefLabel "dekning"@nb ;
  skos:altLabel "kompletthet"@nb ;
  skos:definition "i hvilken grad datasettet inneholder forventede opplysninger, for
en spesifikk brukskontekst"@nb .
```

## Kvalitetsdeldimensjon Underdekning (undercoverage/omission)

**Merknad:** med noe redaksjonelt avvik tilsvarer dette ISO 19157:2013(E) Annex D.2.2.

### Definisjon:

- I hvilken grad det mangler elementer som forventes å være med, for en spesifikk brukskontekst.

### Uttrykt i DQV:

```

@prefix dqv: <https://www.w3.org/TR/vocab-dqv/> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix xxx: <https://ikke.eksisterer.enda/pre_def_kvalitetsmaal/> ;
skos:note "antar at det blir etablert en oversikt over pre-definerte
kvalitets(del)dimensjoner og kvalitetsmål som kan refereres med en URI"@nb .

:undercoverage
  a dqv:Dimension ;
  skos:prefLabel "undercoverage"@en ;
  skos:prefLabel "underdekning"@nb ;
  skos:altLabel "omission"@en ;
  skos:definition "the degree to which required information is missing in a
particular dataset"@en ;
  skos:broader xxx:coverage # antar at "coverage" er definert.

```

## Kvalitetsmål Manglende elementer (missing items)

**Merknad:** ISO 19757:2013(E) Table D.5 - Missing item definerer et kvalitetsmål på hvorvidt et gitt/spesifikt element mangler, mens det som omhandles her i dette avsnittet er et kvalitetsmål på hvorvidt det mangler noen (uspesifikke) elementer i datasettet, derfor "elementer"/"items" i flertall.

### Definisjon:

- Hvorvidt det mangler noen elementer i datasettet.

### Uttrykt i DQV:

```

@prefix dqv: <https://www.w3.org/TR/vocab-dqv/> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix xsd: <https://www.w3.org/TR/xmlschema11-2/> .
@prefix xxx: <https://ikke.eksisterer.enda/pre_def_kvalitetsmaal/> ;
skos:note "antar at det blir etablert en oversikt over pre-definerte
kvalitets(del)dimensjoner og kvalitetsmål som kan refereres med en URI"@nb .

#definisjon av kvalitetsmål

:missingItemsMetric
  a dqv:Metric ;
  skos:prefLabel "missing items metric"@en ;
  skos:prefLabel "hvorvidt det mangler elementer"@nb ;
  skos:definition "whether there are some items missing in a particular dataset"@en ;
  dqv:expectedDataType xsd:boolean ;
  dqv:inDimension xxx:undercoverage # antar at "undercoverage" er definert .

#eksempel på angivelse av måleresultat "true" (ja, det mangler noe)

:measurementMissingItems
  a dqv:QualityMeasurement ;
  dqv:isMeasurementOf :missingItemsMetric ;

```



```
dqv:value "true"^^xsd:boolean .
```

## Kvalitetsmål Antall manglende elementer (number of missing items)

### Definisjon:

- Antall elementer som ikke er i datasettet men som forventes å være med.

### Uttrykt i DQV:

```
@prefix dqv: <https://www.w3.org/TR/vocab-dqv/> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix xsd: <https://www.w3.org/TR/xmlschema11-2/> .
@prefix xxx: <https://ikke.eksisterer.enda/pre_def_kvalitetsmaal/> ;
  skos:note "antar at det blir etablert en oversikt over pre-definerte
  kvalitets(del)dimensjoner og kvalitetsmål som kan refereres med en URI"@nb .

#definisjon av kvalitetsmål
:numberOfMissingItemsMetric
  a dqv:Metric ;
  skos:prefLabel "number of missing items metric"@en ;
  skos:prefLabel "antall manglende elementer"@nb ;
  skos:definition "number of items that are missing in the dataset"@en ;
  dqv:expectedDataType xsd:integer ;
  dqv:inDimension xxx:undercoverage # antar at "undercoverage" er definert .

#eksempel på angivelse av måleresultat "8" (mangler åtte elementer)

:measurementNumberOfMissingItems
  a dqv:QualityMeasurement ;
  dqv:isMeasurementOf :numberOfMissingItemsMetric ;
  dqv:value "8"^^xsd:integer .
```

## Kvalitetsmål Andel manglende elementer (rate of missing items)

### Definisjon:

- Forholdet mellom antall elementer som mangler og antall elementer som skulle være med i datasettet.

### Uttrykt i DQV:

```
@prefix dqv: <https://www.w3.org/TR/vocab-dqv/> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix xsd: <https://www.w3.org/TR/xmlschema11-2/> .
@prefix xxx: <https://ikke.eksisterer.enda/pre_def_kvalitetsmaal/> ;
  skos:note "antar at det blir etablert en oversikt over pre-definerte
  kvalitets(del)dimensjoner og kvalitetsmål som kan refereres med en URI"@nb .
```

```
#definisjon av kvalitetsmål
:rateOfMissingItemsMetric
  a dqv:Metric ;
  skos:prefLabel "rate of missing items metric"@en ;
  skos:prefLabel "andel manglende elementer"@nb ;
  skos:definition "ratio between the number of items that are missing and the number
of the items that should be present"@en ;
  dqv:expectedDataType xsd:decimal ;
  dqv:inDimension xxx:undercoverage # antar at "undercoverage" er definert .
```

#eksempel på angivelse av måleresultat "0.08" (mangler 8%)

```
:measurementRateOfMissingItems
  a dqv:QualityMeasurement ;
  dqv:isMeasurementOf :rateOfMissingItemsMetric ;
  dqv:value "0.08"^^xsd:decimal .
```

## Kvalitetsdeldimensjon Overdekning (overcoverage/commission)

Overdekning kan defineres på helt tilsvarende måte (men "motsatt vis") som for underdekning i avsnittet foran, dvs. om elementer som ikke skulle vært i datasettet. Arbeidsgruppen velger derfor ikke å bruke tid på å eksemplifisere denne deldimensjonen.

**Merknad:** med noe redaksjonelt avvik tilsvarer dette ISO 19157:2013(E) Annex D.2.1.

### Definisjon:

- I hvilken grad datasettet inneholder overflødige elementer

### Merknad til definisjonen:

- Eksklusive [dubletter](#) (som er en egen kvalitetsdeldimensjon)

### Uttrykt i DQV:

```
@prefix dqv: <https://www.w3.org/TR/vocab-dqv/> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix xxx: <https://ikke.eksisterer.enda/pre_def_kvalitetsmaal/> ;
skos:note "antar at det blir etablert en oversikt over pre-definerte
kvalitets(del)dimensjoner og kvalitetsmål som kan refereres med en URI"@nb .

:overcoverage
  a dqv:Dimension ;
  skos:prefLabel "overcoverage"@en ;
  skos:prefLabel "overdekning"@nb ;
  skos:altLabel "commission"@en ;
  skos:definition "the degree to which a particular dataset contains excess items"@en
;
  skos:note "exclusive duplicate"@en ;
```

skos:broader xxx:coverage # antar at «coverage» er definert .

## Kvalitetsdeldimensjon Selektivitet (selectivity)

### Definisjon:

- I hvilken grad datasettet dekker den statistiske populasjonen.

### Uttrykt i DQV:

```
@prefix dqv: <https://www.w3.org/TR/vocab-dqv/> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix xxx: <https://ikke.eksisterer.enda/pre_def_kvalitetsmaal/> ;
skos:note «antar at det blir etablert en oversikt over pre-definerte
kvalitets(del)dimensjoner og kvalitetsmål som kan refereres med en URI»@nb .

:selectivity
  a dqv:Dimension ;
  skos:prefLabel «selectivity»@en ;
  skos:prefLabel «selektivitet»@nb ;
  skos:definition «the degree to which a particular dataset represents the
statistical population»@en ;
  skos:broader xxx:coverage # antar at «coverage» er definert .
```

## Kvalitetsdeldimensjon Dublett (redundancy/duplicate)

### Definisjon:

- I hvilken grad datasettet inneholder flere enn én forekomst av samme opplysning.

### Uttrykt i DQV:

```
@prefix dqv: <https://www.w3.org/TR/vocab-dqv/> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix xxx: <https://ikke.eksisterer.enda/pre_def_kvalitetsmaal/> ;
skos:note «antar at det blir etablert en oversikt over pre-definerte
kvalitets(del)dimensjoner og kvalitetsmål som kan refereres med en URI»@nb .

:redundancy
  a dqv:Dimension ;
  skos:prefLabel «redundancy»@en ;
  skos:altLabel «duplicate»@en ;
  skos:prefLabel «dublett»@nb ;
  skos:definition «the degree to which a particular dataset contains more than one
instance of the same information»@en ;
  skos:broader xxx:coverage # antar at «coverage» er definert .
```

# Kvalitetsdimensjon Aktualitet (currentness/timeliness)

## Definisjon:

- Graden av “ferskhets” av datasettet, for en spesifikk brukskontekst.
  - Engelsk original: The degree to which data has attributes that are of the right age in a specific context of use.
    - Kilde: [W3C DQV](#).

## Uttrykt i DQV:

```
@prefix dqv: <https://www.w3.org/TR/vocab-dqv/> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .

:currentness
  a dqv:Dimension ;
  skos:prefLabel "currentness"@en ;
  skos:altLabel "timeliness"@en ;
  skos:prefLabel "aktualitet"@nb ;
  skos:definition "the degree to which data has attributes that are of the right age
in a specific context of use"@en .
```

# Kvalitetsdimensjon Samsvar (compliance/conformity)

## Definisjon:

- I hvilken grad datasettet er i samsvar med standarder, konvensjoner, regler eller lignende som regulerer datakvalitet, for en spesifikk brukskontekst.
  - Engelsk original: The degree to which data has attributes that adhere to standards, conventions or regulations in force and similar rules relating to data quality in a specific context of use.
    - Kilde: [W3C DQV](#).

## Uttrykt i DQV:

```
@prefix dqv: <https://www.w3.org/TR/vocab-dqv/> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .

:compliance
  a dqv:Dimension ;
  skos:prefLabel "compliance"@en ;
  skos:altLabel "conformity"@en ;
  skos:prefLabel "samsvar"@nb ;
  skos:definition "the degree to which data has attributes that adhere to standards,
conventions or regulations in force and similar rules relating to data quality in a
```

specific context of use@en .

## Kvalitetsbeskrivelse I samsvar med (conforms to)

### Definisjon:

- Datasettet er i samsvar med gitt standard, spesifikasjon, regel og lignende

### Uttrykt i DQV:

```
@prefix dcat: <https://www.w3.org/ns/dcat#> .
@prefix dqv: <https://www.w3.org/TR/vocab-dqv/> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix dcterms: <http://dublincore.org/documents/dcmi-terms/> .
@prefix foaf: <http://xmlns.com/foaf/spec/> .
@prefix xsd: <https://www.w3.org/TR/xmlschema11-2/> .
@prefix xxx: <https://ikke.eksisterer.enda/pre_def_kvalitetsmaal/> ;
skos:note "antar at det blir etablert en oversikt over pre-definerte
kvalitets(dimensjoner og kvalitetsmål som kan refereres med en URI"@nb .

:myDataset
  a dcat:dataset ;
  dcterms:conformsTo :aQualityStandard .

:aQualityStandard
  a dcterms:Standard ;
  dcterms:title "Standard for ..."@en ;
  dcterms:comment "The standard defines ..."@en ;
  dcterms:issued "2017-08-01"^^xsd:date ;
  foaf:page <https://path.ToThe.Standard/aStandard> ;
  dqv:inDimension xxx:compliance # antar at "compliance" er definert .
```

## Kvalitetsdimensjon Tilgjengelighet (availability)

### Definisjon:

- I hvilken grad datasettet kan nåes av brukere og/eller dataapplikasjoner, for en spesifikk brukskontekst.
  - Engelsk original: The degree to which data has attributes that enable it to be retrieved by authorized users and/or applications in a specific context of use.
    - Kilde: [W3C DQV](#).

### Uttrykt i DQV:

```
@prefix dqv: <https://www.w3.org/TR/vocab-dqv/> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
```

```
:availability
  a dqv:Dimension ;
  skos:prefLabel "availability"@en ;
  skos:prefLabel "tilgjengelighet"@nb ;
  skos:definition "the degree to which data has attributes that enable it to be
retrieved by users and/or applications in a specific context of use"@en .
```

## Kvalitetsdimensjon Relevans (relevancy)

### Definisjon:

- I hvilken grad datasettet inneholder data som dekker behov, for en spesifikk brukskontekst.
  - Engelsk original: Relevancy refers to the provision of information which is in accordance with the task at hand and important to the users' query.
  - Kilde: [W3C DQV](#).

### Uttrykt i DQV:

```
@prefix dqv: <https://www.w3.org/TR/vocab-dqv/> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .

:relevancy
  a dqv:Dimension ;
  skos:prefLabel "relevancy"@en ;
  skos:prefLabel "relevans"@nb ;
  skos:definition "the provision of information which is in accordance with the task
at hand and important to the users"@en .
```

## Ikke-kvantitativ/fritekst beskrivelse Bruksformål (specific usage)

### Definisjon:

- Fritekst beskrivelse av hva datasettet er opprettet/innsamlet for

### Uttrykt i DQV:

```
@prefix dcat: <https://www.w3.org/ns/dcat#> .
@prefix dqv: <https://www.w3.org/TR/vocab-dqv/> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix dc: <http://dublincore.org/documents/dces/> .
@prefix oa: <http://www.w3.org/ns/oa#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix xsd: <https://www.w3.org/TR/xmlschema11-2/> .
@prefix xxx: <https://ikke.eksisterer.enda/pre_def_kvalitetsmaal/> ;
skos:note "antar at det blir etablert en oversikt over pre-definerte
kvalitets(del)dimensjoner og kvalitetsmål som kan refereres med en URI"@nb .
```

```

:myDataset
  a dcat:dataset ;
  dqv:hasQualityAnnotation :usageAnnotation .

:usageAnnotation
  a dqv:QualityAnnotation ;
  skos:prefLabel "usability"@en ;
  skos:prefLabel "egnethet"@nb ;
  skos:definition "hva datasettet er opprettet/innsamlet for"@nb ;
  dqv:inDimension xxx:relevancy # antar at "relevancy" er definert ;
  oa:hasBody :usageDescription ;
  oa:motivatedBy dqv:qualityAssessment .

:usageDescription
  a oa:TextualBody ;
  rdf:value "datasettet er opprettet for ..." ;
  dc:language "nb" ;
  dc:format "text/plain" .

```

## Ikke-kvantitativ/fritekst beskrivelse Egnethet (usability)

### Definisjon:

- Fritekst beskrivelse av hva datasettet er, og ikke er, egnet til

### Uttrykt i DQV:

```

@prefix dcat: <https://www.w3.org/ns/dcat#> .
@prefix dqv: <https://www.w3.org/TR/vocab-dqv/> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix dc: <http://dublincore.org/documents/dces/> .
@prefix oa: <http://www.w3.org/ns/oa#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix xsd: <https://www.w3.org/TR/xmlschema11-2/> .
@prefix xxx: <https://ikke.eksisterer.enda/pre_def_kvalitetsmaal/> ;
skos:note "antar at det blir etablert en oversikt over pre-definerte
kvalitets(del)dimensjoner og kvalitetsmål som kan refereres med en URI"@nb .

:myDataset
  a dcat:dataset ;
  dqv:hasQualityAnnotation :usageAnnotation .

:usabilityAnnotation
  a dqv:QualityAnnotation ;
  skos:prefLabel "usability"@en ;
  skos:prefLabel "egnethet"@nb ;
  skos:definition "hva datasettet er, og ikke er, egnet til"@nb ;
  dqv:inDimension xxx:relevancy # antar at "relevancy" er definert ;
  oa:hasBody :usabilityDescription ;
  oa:motivatedBy dqv:qualityAssessment .

```

```
:usabilityDescription
  a oa:TextualBody ;
  rdf:value "datasettet er best egnet for å beregne ..., men kan by på utfordringer
når det også brukes til å analysere ..." ;
  dc:language "nb" ;
  dc:format "text/plain" .
```



## Vedlegg C - URler/navnerom som er brukt

| Prefiks | Navnerom                                                                                              | Beskrivelse                                        |
|---------|-------------------------------------------------------------------------------------------------------|----------------------------------------------------|
| dqv     | <a href="https://www.w3.org/TR/vocab-dqv/">https://www.w3.org/TR/vocab-dqv/</a>                       | Data Quality Vocabulary                            |
| dc      | <a href="http://dublincore.org/documents/dces/">http://dublincore.org/documents/dces/</a>             | Dublin Core Metadata Element Set                   |
| dcterms | <a href="http://dublincore.org/documents/dcmi-terms/">http://dublincore.org/documents/dcmi-terms/</a> | Dublin Core Metadata Terms                         |
| dcat    | <a href="https://www.w3.org/ns/dcat#">https://www.w3.org/ns/dcat#</a>                                 | Data Catalog Vocabulary (DCAT)                     |
| duv     | <a href="https://www.w3.org/TR/vocab-duv/">https://www.w3.org/TR/vocab-duv/</a>                       | Dataset Usage Vocabulary (duv)                     |
| foaf    | <a href="http://xmlns.com/foaf/spec/">http://xmlns.com/foaf/spec/</a>                                 | Friend-of-a-Friend Vocabulary                      |
| oa      | <a href="http://www.w3.org/ns/oa#">http://www.w3.org/ns/oa#</a>                                       | The Web Annotation Data Model                      |
| skos    | <a href="http://www.w3.org/2004/02/skos/core#">http://www.w3.org/2004/02/skos/core#</a>               | Simple Knowledge Organization System               |
| rdf     | <a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#">http://www.w3.org/1999/02/22-rdf-syntax-ns#</a> | Resource Description Framework (RDF)               |
| rdfs    | <a href="http://www.w3.org/2000/01/rdf-schema#">http://www.w3.org/2000/01/rdf-schema#</a>             | RDF Schema                                         |
| xsd     | <a href="https://www.w3.org/TR/xmlschema11-2/">https://www.w3.org/TR/xmlschema11-2/</a>               | XML Schema Datatypes                               |
| xxx     |                                                                                                       | Vokabular som ikke finnes, men som trengs etablert |

# Vedlegg D - Referanser

Informativ del av dette dokumentet.

I dette vedlegget tas med referanser som er aktuelle, men ikke tatt med som normative referanser i dette dokumentet. Referanser er tatt med i en usortert/uprioritert rekkefølge.

- ISO/IEC 25012, Software engineering — Software product Quality Requirements and Evaluation (SQuARE) — Data quality model, <http://iso25000.com/index.php/en/iso-25000-standards/iso-25012>
- ISO 19157:2013, Geographic information — Data quality, <http://www.standard.no/no/Nettbutikk/produktkatalogen/Produktpresentasjon/?ProductID=672821>
- EU, BLUE-ETS <sup>[1]</sup>, <https://www.blue-ets.istat.it/fileadmin/deliverables/Deliverable4.1.pdf>.
- Eurostat, European Statistics Code of Practice (CoP), <http://ec.europa.eu/eurostat/web/quality/european-statistics-code-of-practice>

[1] BLUE-ETS er et prosjekt under EUs program Framework Programme 7: BLUE - Enterprise and Trade Statistics, hvor en bl.a. så på kvalitetskriterier for administrative datasett brukt til produksjon av offisiell statistikk.

# Vedlegg E - Arbeidsgruppen

Informativ del av dette dokumentet.

## Beskrivelse av arbeidsgruppens oppgave

Arbeidsgruppen fikk i oppgave å utarbeide en spesifikasjon for hvordan kvalitet på datasett skal beskrives i en datakatalog.

## Sammensetning av arbeidsgruppen

Kartverket ved Morten Borrebæk har ledet denne arbeidsgruppen. Følgende etater og personer har ellers vært involvert i arbeidet:

| Virksomhet                        | Person                                                    |
|-----------------------------------|-----------------------------------------------------------|
| Kartverket                        | Morten Borrebæk (leder av arbeidsgruppen)                 |
| Arkivverket                       | Joachim Fugleberg                                         |
| Brønnøysundregistrene             | Espen Slotvik                                             |
| Difi                              | Martin Standley, Jim J. Yang, Øystein Åsnes, Pia Jøsendal |
| Politiet                          | Arne Dybdahl                                              |
| Skattedirektoratet                | Katrine Fredriksen                                        |
| Statens lånekasse                 | Liv Bergliot Simonsen, Gustav Aagesen                     |
| Statens vegvesen                  | Cecilie H. Bratt, Hilde Austlid                           |
| Statistisk sentralbyrå            | Hans Viggo Sæbø                                           |
| <i>Kopimottakere:</i>             |                                                           |
| <i>Direktoratet for e_-helse_</i> | <i>Åsmund Ahlmann Nyre</i>                                |
| <i>Statens pensjonskasse</i>      | <i>Andrea Halvorsen, Margaret McLeod</i>                  |

Mot slutten av arbeidet ble det etablert en mindre arbeidsgruppe som fikk i oppgave å foreslå hvordan det foreslåtte faglige innholdet skal kunne representeres ved hjelp av DQV og implementeres i DCAT-AP-NO. Den mindre arbeidsgruppen har bestått av følgende etater/personer:

| Virksomhet             | Person                                      |
|------------------------|---------------------------------------------|
| Kartverket             | Morten Borrebæk (leder)                     |
| Brønnøysundregistrene  | David Norheim                               |
| Difi                   | Martin Standley, Jim J. Yang, Øystein Åsnes |
| Statistisk sentralbyrå | Hans Viggo Sæbø                             |

# Arbeid i arbeidsgruppen

Arbeidsgruppen startet sitt arbeid fra 25. januar 2017. Har hatt halvdagsmøter ca. annen hver uke.

I perioden frem til sommeren 2017 har arbeidsgruppen hovedsakelig jobbet med: \* Utvekslet erfaringer fra deltagende virksomheter på datakvalitet og måling av datakvalitet Sett spesielt på relevante internasjonale standarder/spesifikasjoner, på området datakvalitet og beskrivelse av datakvalitet, deriblant W3C DQV, ISO/IEC 25012, ISO 19157, BLUE-ETS, CoP. \* Foreslått en standardisert måte å beskrive kvalitet på datasett (se [#heading=h.9auauka2obwd](#)[4 Standardisert metode for å beskrive kvalitet på datasett]) \* Foreslått noen få hoveddimensjoner og deldimensjoner av datasettkvalitet (se [#heading=h.mlfyh2ml9hmk](#)[5 Vedlegg A - Forslag til kvalitetsdimensjoner, kvalitetsdeldimensjoner og kvalitetsmål])

I sommeren 2017, har en mindre arbeidsgruppe jobbet med: \* Eksemplifisering av kvalitetsdimensjoner, kvalitetsdeldimensjoner og kvalitetsmål ved å bruke W3C DQV. \* Forslag til hvordan dette implementeres i DCAT-AP-NO.

Arbeidet ble avsluttet 13. september 2017.