

Cristin - import av institusjonsdata fra lokale systemer

Tore Vatnan, 2004-2010

Endringshistorikk	
2010-09-30	Endret navn fra "Frida" til "Cristin"
2009-05-25	organisasjon/enhet/akronym økt til 12 tegn
2008-09-03	Lagt til fnrErstatter for person

Generelt

Cristin forutsetter at enkelte data om brukerinstitusjonen ligger i databasen før registrering gjøres. Det gjelder:

- Data om enheter (steder)
- Data om personer (= ansatte, samt eksterne stipendiater og emerituser)
- Data om tilknytningen mellom personer og enheter (ansettelser)

Disse dataene må importeres jevnlig slik at de er mest mulig oppdatert i Cristin.

Importformatet er XML hvor strukturen er definert av et XML schema (se under). De enkelte elementene og attributtene er navngitt etter "Java-standard" med små (engelske) bokstaver. Det innebærer at navnene starter med liten forbokstav. I navn som er satt sammen av flere ord skrives ordene i ett - det første ordet med liten forbokstav, og de påfølgende med stor forbokstav, f.eks. "stednavnBokmal".

Importformatet angir de elementene som er naturlig å importere i forhold til den funksjonaliteten som er implementert i Cristin i dag, eller som er planlagt implementert i nær fremtid. Ny funksjonalitet i Cristin vil kunne fordre at importformatet utvides/endres.

Det skal benyttes vest-europeisk tegnssett (ISO Latin 1 / ISO 8859-1). Dette skal angis i første linje i filen som da skal være

```
<?xml version="1.0" encoding="ISO-8859-1"?>
```

Grafikk

Elementer angis som rektangler med navn øverst og evt. datatype nederst til høyre. Elementene inneholder enten ett eller flere underelementer, eller de inneholder tekst.

Stiplede linjer angir ikke-obligatoriske elementer.

Sekvenser angis med . Dette betyr at underelementene må forekomme i angitt rekkefølge.

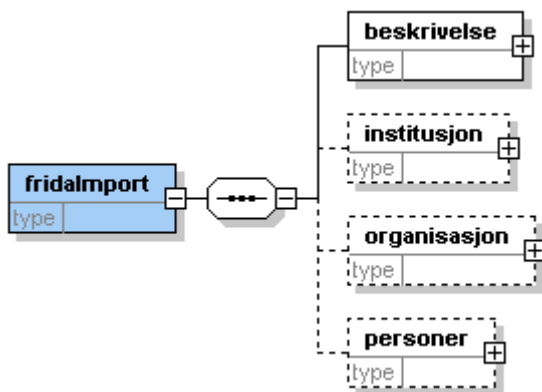
"+" og "-" (pluss og minus) angir om elementet/sekvensen/valget er ekspandert (om underelementene er synlige/ikke-synlige). For oversiktliggheitens skyld er en del beskrivelser av elementer splittet opp på en slik måte at noen av underelementene ikke er ekspandert/beskrevet, men at beskrivelse er gitt et annet sted.

Antallet lovlige forekomster av et element er angitt under dersom det avviker fra 0..1 (angitt med stiptet linje) eller 1.

Supplerende dokumentasjon

- Schema
- Beskrivelse av schema (HTML, Word)
- Eksempel

Hovedelementer



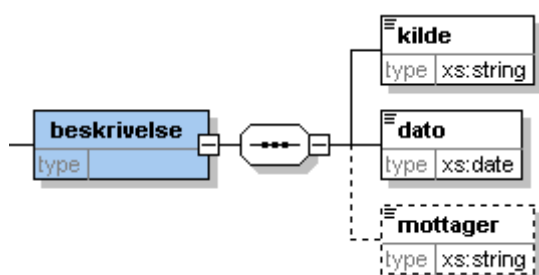
En importfil består av beskrivelse av dokumentet (kilde og dato) og informasjon om institusjon (navn og kontaktadresser), organisasjon (enheter som representerer en hierarkisk struktur), samt personer (med ansettelses).

En importfil bør være komplett i den forstand at alle nåværende enheter og personer med ansettelses, bør inngå i filen. Dette muliggjør validering av filen m.h.t. referanser internt og mellom hovedenhetene.

For enkel validering bør "namespace" angis i hovedelementet som

```
<fridaImport xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation="http://frida.usit.uio.no/import/institusjonsdata/schema/Frida-import-1_0.xsd">
```

Beskrivelse



Beskrivelse består av elementene "kilde" og "dato". Denne informasjonen blir ikke importert i Cristin.

kilde

Angivelse av kilden til dataene. Normalt navnet på institusjonen hvor dataene stammer fra.

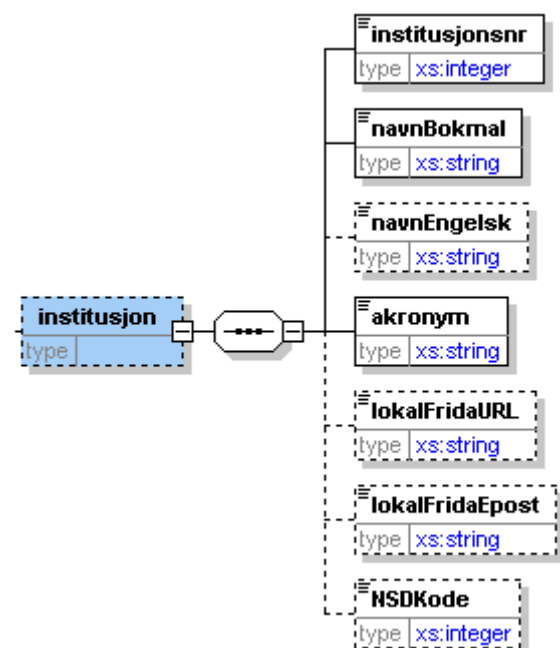
dato

Angivelse av dato for eksport fra lokalt system. Format "åååå-mm-dd".

mottager

Angivelse av hvem som er mottager av dataene. Ikke obligatorisk.

Institusjon



Elementet "institusjon" brukes for å angi navn, akronym, samt URL og e-postadresse for lokal informasjon/brukerstøtte i forbindelse med Cristin.

institusjonsnr

Nummer som representerer institusjonen (NTNU=194, UiB=184, UiO=185, UiTø=186).
Inntil 8 siffer.

navnBokmal

Institusjonens offisielle navn på bokmål. 120 tegn.

navnEngelsk

Institusjonens offisielle navn på engelsk. 120 tegn. Ikke obligatorisk.

akronym

Institusjonens offisielle akronym (NTNU, UiO, ...). 10 tegn.

lokalFridaURL

URL til lokal informasjon om Cristin. 200 tegn. Ikke obligatorisk.

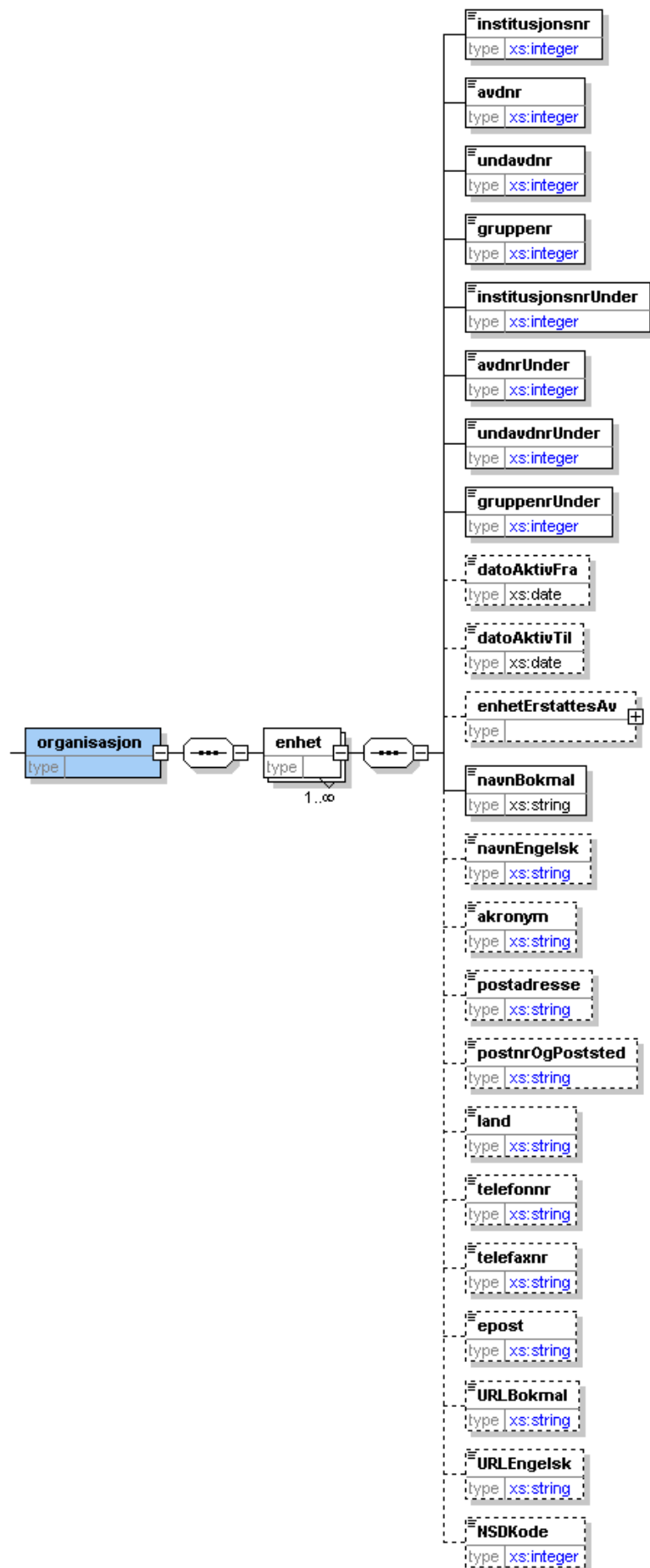
lokalFridaEpost

E-postadresse til lokal brukerstøtte. 80 tegn. Ikke obligatorisk

NSDKode

Institusjonens NSD-kode

Organisasjon/enheter



Elementet "organisasjon" består av én eller flere forekomster av elementet "enhet" som representerer data om institusjon, fakultet, institutt eller gruppe. Hver enhet identifiseres med et tuppel bestående av institusjonsnr, avdnr (avdelingsnummer - fakultet), undavdnr (underavdelingsnummer - institutt) og gruppenr. Dette utgjør nøkkelen for enheten i databasen, og skal således være entydig. I tillegg til nøkkelen skal hver enhet også ha en fremmednøkkel som peker til "forelder" i institusjonshierarkiet. Toppnivået på institusjonen har en "foreldrepeker" til seg selv.

Selv om nøkkelnavnene indikerer at en kan representere et hierarki med fire nivåer, gir bruken av fremmednøkler mulighet til å representere et hierarki av langt flere nivåer. Det anbefales imidlertid at en begrenser seg til maksimalt fire nivåer. Det er heller ingen begrensning når det gjelder "kobling på tvers". F.eks. kan en avdeling/fakultet med et gitt avdelingsnummer i teorien godt ha underliggende enheter med helt andre avdelingsnumre.

Obligatoriske elementer for enhet er nøkkel og fremmednøkkel, samt offisielt navn på enheten på bokmål. Det anbefales imidlertid sterkt å også ta med "kontaktdata" (adresse, telefonnr, telefaxnr og e-postadresse).

institusjonsnr

Tall som representerer institusjonen (NTNU=194, UiB=184, UiO=185, UiTø=186). Inntil 8 siffer.

avdnr

To-sifret tall som normalt angir nivået rett under toppnivået for institusjonen, typisk fakultet.

undavdnr

To-sifret tall som normalt angir nivået angir rett under avdeling/fakultet, typisk institutt.

gruppenr

To-sifret tall som normalt angir nivået rett under underavling/institutt.

institusjonsnrUnder

Tre-sifret tall som viser til overliggende enhets institusjonsnummer.

avdnrUnder

To-sifret tall som viser til overliggende enhets avdelingsnummer.

undavdnrUnder

To-sifret tall som viser til overliggende enhets underavdelingsnummer.

gruppenrUnder

To-sifret tall som viser til overliggende enhets gruppenummer.

datoAktivFra

Datoen enheten ble opprettet. Ikke obligatorisk.

datoAktivFra

Datoen enheten ble nedlagt. Ikke obligatorisk.

enhetErstattesAv

Ikke obligatorisk. Se under.

navnBokmal

Enhetens offisielle navn på bokmål. Maksimalt 512 tegn.

navnEngelsk

Enhetens offisielle navn på engelsk. Maksimalt 512 tegn. Ikke obligatorisk.

akronym

Enhetens offisielle akronym. Maksimalt 12 tegn. Ikke obligatorisk.

postadresse

Enhetens postadresse (f.eks. angivelse av postboks). Ikke obligatorisk, men sterkt anbefalt. Maksimalt 100 tegn.

postnrOgPoststed

Angivelse av postnummer og -sted. Ikke obligatorisk, men sterkt anbefalt. Maksimalt 100 tegn.

land

Benyttes primært dersom enheten er lokalisert utenfor Norge. Maksimalt 100 tegn. Ikke obligatorisk.

telefonnr

Enhetens offisiell telefonnummer. Maksimalt 20 tegn. Ikke obligatorisk, men sterkt anbefalt.

telefaxnr

Enhetens offisiell telefaxnummer. Maksimalt 20 tegn. Ikke obligatorisk, men sterkt anbefalt.

epost

Enhetens offisiell e-postadresse. Maksimalt 80 tegn. Ikke obligatorisk, men sterkt anbefalt.

URLBokmal

URL til enhet (bokmål). Ikke obligatorisk.

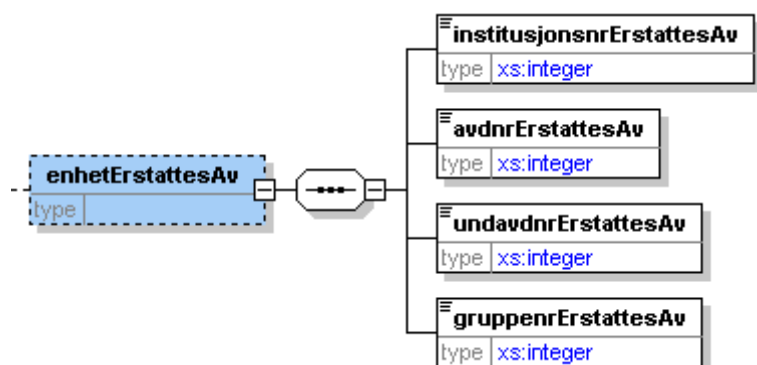
URLEngelsk

URL til enhet (engelsk). Ikke obligatorisk.

NSDKode

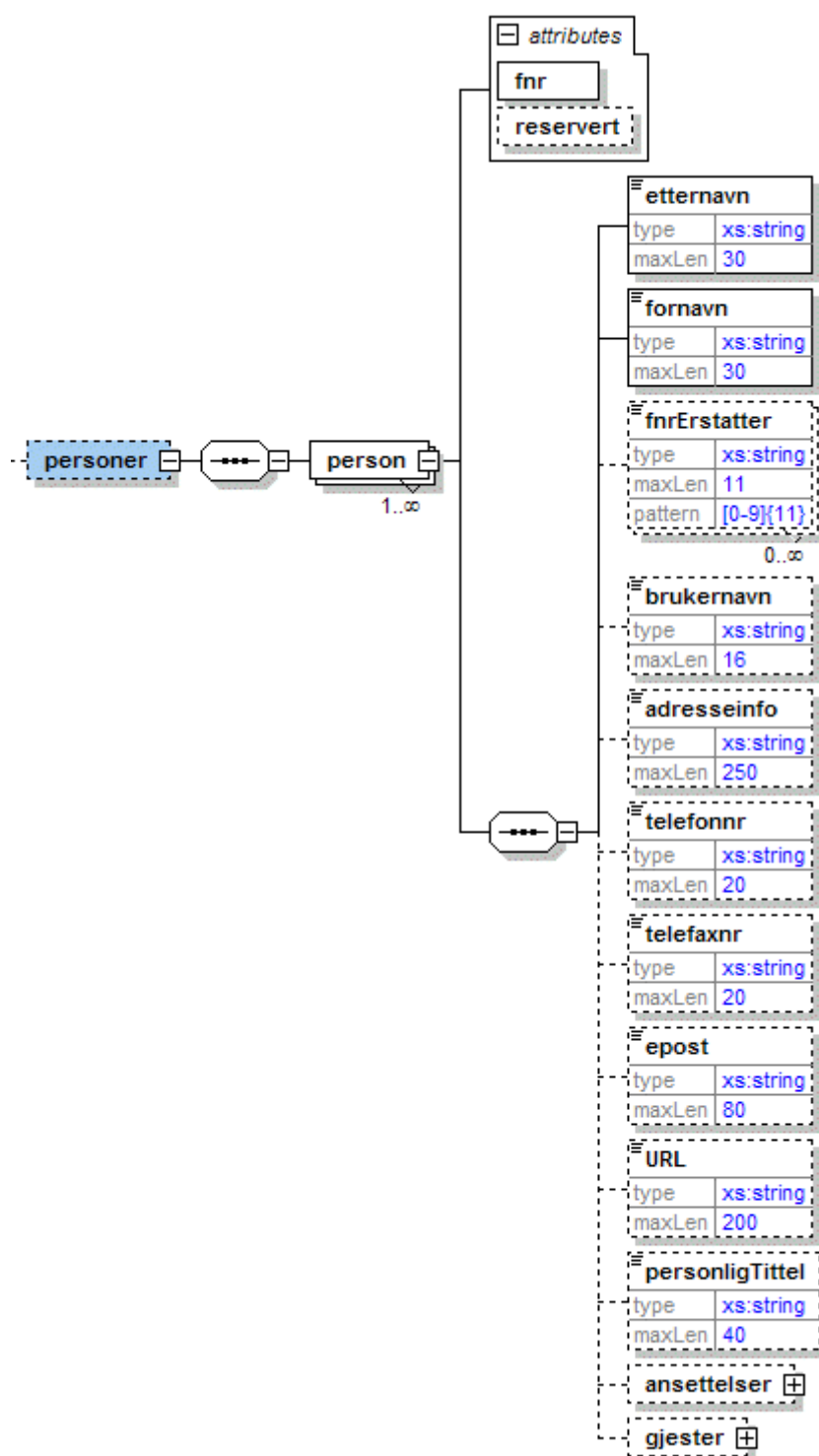
Enhetens NSD-kode

Enhet erstattes av



Når en enhet erstattes av en annen (f.eks. ved at enheten flyttes i organisasjonen e.l.), kan dette angis ved å brukes underelementene i **enhetErstattesAv**. Underelementene utgjør fremmednøkkel til en annen enhet som må være definert tidligere i denne eksporten fra det lokale systemet.

Personer og ansettelseser



Elementet "personer" kan inneholde ett eller flere personelementer. Et personelement inneholder "ansettelser" som skal inneholde ett eller flere ansettelseselementer.

Alle personer som det skal refereres til må importeres til Cristin. Disse vil i utgangspunktet også bli opprettet som brukere med rett til å bl.a. registrere resultater og foreta endringer i egen kompetanseprofil. Fullt fødselsnummer er obligatorisk for disse.

Disse personene må ikke blandes sammen med navn på eksterne personer (forfattere o.l.) der referansen kun er navn som ikke (nødvendigvis) identifiserer bestemte personer. Disse eksterne personene trengs ikke å være registrert i Cristin før registrering.

Elementet person har to attributter:

- fnr (fødselsnummer), obligatorisk. Tekst på 11 tegn (sifre).
- reserver, kan ha verdien 'J' eller 'N' (ja/nei), og er ikke obligatorisk. Dersom en angir reserver ('J') vil Cristin ikke vise personens tilhørighet ved institusjonen for eksterne brukere (= ikke innloggede brukere), og vedkommende vil heller ikke synes i kompetansekatalogen. Reservering er kun tenkt benyttet for personer som av meget spesielle grunner skal skjermes.

Følgende persondata inngår:

etternavn

Offisielt etternavn (ikke forfatternavn). Maksimalt 30 bokstaver.

fornavn

Offisielt fornavn (ikke forfatternavn). Maksimalt 30 bokstaver.

fnrErstatter

Tidligere fødselsnummer. 11 sifre.

brukernavn

Brukernavn som benyttes ved autentisering. Maksimalt 16 tegn.

adresseinfo

Adresseopplysninger. Maksimalt 250 tegn. Ikke obligatorisk.

telefonnr

Telefonnummer. Maksimalt 20 tegn. Ikke obligatorisk.

telefaxnr

Personens telefaxnummer.. Maksimalt 20 tegn. Ikke obligatorisk.

epost

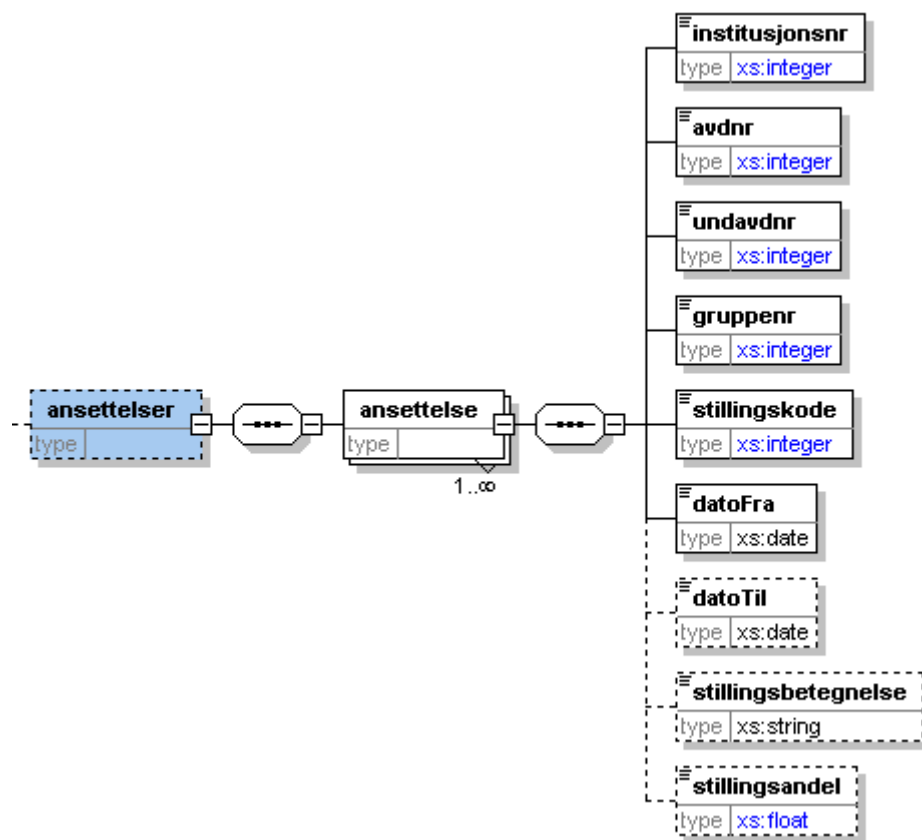
E-postadresse. Maksimalt 80 tegn. Ikke obligatorisk.

URL

URL til f.eks. personenens web-side ved institusjonen. Maksimalt 200 tegn. Ikke obligatorisk.

personligTittel

Tittel, f.eks. "daglig leder". Maksimalt 40 tegn. Ikke obligatorisk.



Hver person skal ha minst én ansettelse som er aktiv på det tidspunktet eksport til Cristin foretas. Følgende elementer inngår i ansettelseselementet:

institusjonsnr

Nummer som representerer institusjonen (NTNU=194, UiB=184, UiO=185, UiTø=186).
Inntil 8 siffer.

avdnr

To-sifret tall som normalt angir nivået rett under toppnivået for institusjonen, typisk fakultet.

undavdnr

To-sifret tall som normalt angir nivået angir rett under avdeling/fakultet, typisk institutt.

gruppenr

To-sifret tall som normalt angir nivået rett under underavling/institutt.

stillingskode

Stillingskode. Normalt fire siffer. Maksimalt 16 tegn.

datoFra

Angivelse av startdato for denne ansettelsen. Format "åååå-mm-dd". Dersom startdato ikke er kjent/tilgjengelig, settes en fiktiv dato langt tilbake i tid, f.eks. "1900-01-01".

datoTil

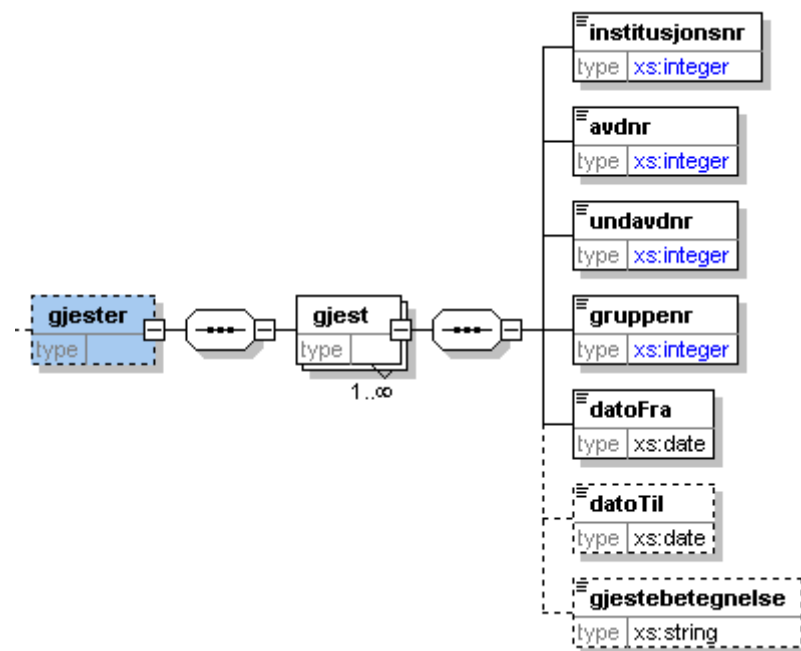
Angivelse av sluttdato for denne ansettelsen. Format "åååå-mm-dd". Dersom sluttdato ikke er kjent/avgjort brukes ikke dette elementet. Alternativt angis fiktiv dato langt frem i tid, f.eks. "3000-01-01". Ikke obligatorisk.

stillingsbetegnelse

Betegnelsen på stillingen i henhold til (statlig) standard. Funksjonelle betegnelser (f.eks. "daglig leder") skal ikke benyttes. Denne betegnelsen vil kun bli tatt inn i Cristin hvis stillingskoden ikke allerede finnes fra før. Maksimalt 40 tegn. Ikke obligatorisk.

stillingsandel

Angivelse av prosentandel av full stilling. Desimaltall mindre eller lik 100.0. Ikke obligatorisk.



gjester brukes for å angi andre tilknytningsforhold enn ansettelser.

institusjonsnr, avdnr, undavdnr, gruppenr, datoFra og datoTil

Som for ansettelse.

gjestebetegnelse

Betegnelse på "gjestetypen".