

Introduktion till rättsdatalagret

För vidareutnyttjare av rättsinformation

Detta dokument förklarar vad rättsdatalagret är, vilken information som lagras där, och hur man kan komma åt och vidareutnyttja denna.

Om du har några frågor med anledning av detta dokument eller angående rättsinformationsprojektet i stort, kontakta gärna Heléne Lundgren på helene.lundgren@dom.se.

Den senaste versionen av detta dokument finns tillgänglig på <http://dev.lagrummet.se/dokumentation/introduktion/intro-vidareutnyttjare.pdf>

Ändringar

Datum	Ändring
2014-01-30	Uppdaterat kontaktperson
2010-11-19	Initial version

Innehåll

- 1. Den korta versionen
- 2. Ordlista / begreppsdefinitioner
- 3. Bakgrund
- 4. Åtkomst till rättsdatalogret
- 5. Tjänstegränssnitt

Rättsdatalogret är i skrivande stund (oktober 2011) visserligen i skarp drift, men ganska tomt eftersom endast enstaka informationsproducenter levererar information till systemet. För att kunna testa och visa systemet har vi simulerat ca 40 000 dokument av olika typer (lagar, förordningar, rättsfallsreferat och olika typer av förarbeten) med grundläggande metadata och hänvisningar, och importerat detta i en demoversion av systemet.

I detta dokument används för närvarande adresser till detta demosystem. Sådana adresser kan genomgående översättas till motsvarande adresser till i det skarpa systemet, vilket kommer vara lämpligt när detta system fylls med information. Detta görs genom att byta ut hostnamnet `rinfo.demo.lagrummet.se` mot `rinfo.lagrummet.se`, och `service.demo.lagrummet.se` mot `service.lagrummet.se`.

Exempel:

```
http://rinfo.demo.lagrummet.se/publ/sfs/1999:175
```

|
V

```
http://rinfo.lagrummet.se/publ/sfs/1999:175>
```

Exempel:

```
http://service.demo.lagrummet.se/-/publ?q=eko*
```

|
V

```
http://service.lagrummet.se/-/publ?q=eko*
```

1. Den korta versionen

För att lista alla dokument som innehåller ord som börjar på "ekonom":

```
$ curl -H "Accept: application/json" \  
http://service.demo.lagrummet.se/-/publ?q=ekonom*
```

För att lista alla dokument utgivna av Finansinspektionen:

```
$ curl -H "Accept: application/json" \  
http://service.demo.lagrummet.se/-/publ?publisher.iri=*/finansinspektionen
```

För att hämta all tillgänglig information om ett visst dokument, inklusive information om utgivare och andra dokument som hänvisar till detta (adressen kan läsas ut från träfflistan i något av de ovanstående exemplen) :

```
$ curl -H "Accept: application/json" \  
http://service.demo.lagrummet.se/publ/fffs/2009:1/data
```

Hämta själva dokumentet i ett angivet format (adress tillgängliga format kan utläsas från informationen i förra exemplet), så som det har levererats till rättsdatalogret från den publicerande myndigheten:

```
$ curl -H "Accept: application/pdf" \  
http://rinfo.demo.lagrummet.se/publ/fffs/2009:1
```

Hämta information **om** dokumentet, så som det har levererats:

```
$ curl -H "Accept: application/rdf+xml" \  
http://rinfo.demo.lagrummet.se/publ/fffs/2009:1
```

Hämta ett Atom-flöde över de senaste dokumenten som publicerats i systemet:

```
$ curl -H "Accept: application/rdf+xml" \  
http://rinfo.demo.lagrummet.se/feed/current/
```

Hämta en komplett lista över dokument som tidigare publicerats i systemet genom att följa alla <link rel="prev-archive">-länkar rekursivt (i enlighet med [RFC 5005](#)).

Bläddra bland alla dokument på

<http://service.demo.lagrummet.se/view/browse/publ> eller sök med fritext- eller fältsökning på <http://service.demo.lagrummet.se/ui/>.

2. Ordlista / begreppsdefinitioner

Informationspost

Ett sammanhållande begrepp för ett enskilt dokument, dess identifierare och metadata om dokumentet. Informationsposten innehåller även information om när själva posten skapades eller uppdaterades.

Informationsproducent

En myndighet som enligt rättsinformationsförordningen (1999:175) har ett informationsansvar, dvs en skyldighet att se till att viss information tillgängliggörs och sprids elektroniskt.

Rättsdatalogret

Ett system för att samla in och tillgängliggöra all den rättsinformation som enligt rättsinformationsförordningen ska finnas tillgänglig.

Rättsinformation

Författningar, vägledande domstols- och myndighetsavgöranden, förarbeten och andra rättskällor. Begreppet används här i samma betydelse som i rättsinformationsförordningen.

Vidareutnyttjare

En part (myndighet, företag eller privatperson) som vill vidareutnyttja den information finns i rättsdatalagret för att exempelvis bygga en informationstjänst på nätet.

3. Bakgrund

Den grundläggande offentliga rättsinformationen som idag finns tillgänglig via Internet görs tillgänglig via ett stort antal källor. De centrala tillhandahålls av Regeringskansliet (SFS i tryckt och konsoliderad form samt förarbeten m.m.) och Domstolsverket (Vägledande domstolsavgöranden), men även alla länsstyrelser och de myndigheter som har föreskriftsrätt tillhandahåller sina författningssamlingar i elektronisk form.

De ca 100 tillgängliga källorna görs inte tillgängliga på något enhetligt sätt, utan varje informationsproducent har frihet att ordna navigation, sökning, formatval och övriga detaljer i gränssnittet efter eget huvud. Det saknas även i regel någon form av maskinpassat gränssnitt (API) för att komma åt informationen.

Detta orsakar mycket arbete för den som vill vidareutnyttja informationen. I enlighet med rättsinformationssystemets syfte, att tillgodose enskilda och den offentliga förvaltningen tillgång till grundläggande rättsinformation, är det eftersträvansvärt att detta görs enklare.

Rättsinformationsprojektets syfte är att åstadkomma detta genom att samordna informationsproducenternas publicering så att en gemensamt teknisk standard används med väldefinierade metadatermer, och en gemensam lagringsplats från vilken rättsinformationen kan sökas och hämtas.

4. Åtkomst till rättsdatalagret

En grundläggande tanke i rättsinformationsprojektet är att central insamling av rättsinformation till ett rättsdatalager måste ske för att göra det möjligt att erbjuda ett enhetligt gränssnitt för att bläddra, läsa och söka i rättsinformationen. Detta är dock inte tänkt att medföra en ändring i informationsansvaret, som även i fortsättningen ska vara decentraliserat.

En annan tanke är att tekniska format för informationen måste standardiseras i mer detalj än vad som gjorts hittills. För att rättsdatalagret ska kunna hämta informationen med automatik så krävs att varje myndighet publicerar en *postförteckning*, dvs en lista över vilka dokument som finns tillgängliga. Denna lista ska följa det standardiserade Atom-formatet.

Gränssnittet mellan informationsproducenter och rättsdatalagret är detsamma som gränssnittet mellan rättsdatalagret och vidareutnyttjare. Men istället för att läsa från ca 100 olika datakällor kan en vidareutnyttjare, genom att gå mot det centrala rättsdatalagret, istället läsa från endast en källa.

4.2 Atom-flöden

Startpunkten för att hämta data från rättsdatalagret är dess Atom-flöde som återfinns på <http://rinfo.lagrummet.se/feed/current>. En vidareutnyttjare kan använda de mekanismer som Atom-formatet specificerar för att få uppdateringar om nyttillkomna och ändrade informationsposter i systemet.

Eftersom antalet informationsposter i rättsdatalagret är många fler (tiotusentals) än vad som är praktiskt att lista i en "fysisk" Atom-fil använder rättsdatalagret en mekanism med arkivfiler för äldre material. Huvudflödet innehåller en länk till den nyaste arkivfilen, som i sin tur innehåller en länk till den näst nyaste, och så vidare. Genom att traversera denna kedja kan man hämta hela det "logiska" flödet.

Atom-flödet innehåller i sig inte så mycket faktisk information. Den nås istället genom att följa de länkar som finns i flödet. Varje informationspost uttrycks i Atom-flödet som ett <entry>-element. Inuti detta finns sedan ytterligare tre viktiga element:

- <id>: Anger den maskinläsbara identifieraren, i form av en URI, för dokumentet i fråga. Rättsinformationsprojektet har tagit fram en uppsättning principer för hur dessa identifierare ska konstrueras för olika typer av dokument.
- <link>: Hänvisar till en fil med information om dokumentet, uttryckt i RDF-format. Se nedan för information om vad som ingår i detta.
- <content>: Hänvisar till en fil som innehåller en representation av dokumentet, vanligen i form av en PDF-fil.

4.3 RDF-data

För varje dokument ska finnas en uppsättning grundläggande metadata, vilket innefattar egenskaper som dokumentets titel, beteckning, utgivningsdatum, men också hur dokumentet relaterar till andra dokument. Enligt den begreppsmodell som rättsinformationsprojektet tagit fram finns ett antal olika dokumenttyper, och beroende på dokumenttyp skiljer sig den obligatoriska och frivilliga metadatan åt.

Mer information om vilken RDF-data som ska finnas för de olika dokumenttyperna finns i de olika handböcker som rättsinformationsprojektet tagit fram.

Om man har ett dokument identifierare i form av en URI kan man få ut RDF-datat för dokumentet genom att använda HTTP-protokollets mekanism för *content negotiation*.

```
$ curl -H "Accept: application/rdf+xml" http://rinfo.lagrummet.se/publ/rf/rk/2009:4
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
<rdf:Description rdf:about="http://rinfo.lagrummet.se/publ/rf/rk/2009:4">
  <creator xmlns="http://purl.org/dc/terms/"
    rdf:resource="http://rinfo.lagrummet.se/org/kammarratten-i-stockholm"/>
  <identifier xmlns="http://purl.org/dc/terms/">RK 2009:4</identifier>
  ...
```

Det finns även, som beskrivs ovan, länkar i Atom-flödet som leder direkt till RDF-beskrivningar av dokument, som inte kräver content negotiation.

4.4 Dokument

På liknande sätt som man kommer åt RDF-datat genom att ange att man vill ha just *application/rdf+xml* så kan man komma åt exv en PDF-representation (om sådan finns):

```
$ curl -H "Accept: application/pdf"
http://rinfo.lagrummet.se/publ/rf/rk/2009:4 > rk-2009-4.pdf
```

Även här kan man använda Atom-flödet för att ta reda på vilka olika representationer av resursen som finns tillgänglig.

5. Tjänstegränssnitt

Vid sidan av det grundläggande datalagret finns även en uppsättning **tjänster** som vidareutnyttjare kan använda för att enklare hitta och sammanställa rättsinformationen. Tjänstelagret innehåller ingen information som inte redan finns i datalagret, men informationen presenteras på ett mer användbart sätt. Tanken med tjänstelagret är göra det möjligt för utvecklare att konstruera tjänster på informationen i rättsdatalagret utan att nödvändigtvis behöva hämta alla tiotusentals dokument på förhand.

Tjänstegränssnittet kan nås på <http://service.lagrummet.se/> (Obs! Denna address kan komma att ändras till en address under <http://rinfo.lagrummet.se/> i framtiden.)

5.1 Åtkomst till information om dokument (och andra resurser)

Tjänstelagret kan ge en informationsvy för varje resurs som bygger både på den metadata som kan fås från datalagret för den resursen men också metadata för andra resurser som hänvisar till denna. För en lag kan exempel visas rättsfall som hänvisat till lagen i fråga.

Adressen till denna informationsvy skapas från dokument-URI:n på följande sätt:

`http://rinfo.lagrummet.se/publ/sfs/1999:175`

|
V

`http://service.lagrummet.se/publ/sfs/1999:175/data`

Även här kan man använda content negotiation för att få ut informationen i olika format. Det är även möjligt att istället utöka adressen med en dokumenttypsuffix för att få ett visst givet format (exempelvis `http://service.lagrummet.se/publ/sfs/1999:175/data.n3`). Samma information (i form av en utökad RDF-graf) returneras i samtliga fall, men i olika format.

MIME-typ	Suffix	Kommentar
<code>application/json</code>	<code>.json</code>	Denna vyn använder sig av <u>JSON-LD</u> för att representera RDF i JSON-format.
<code>application/rdf+xml</code>	<code>.rdf</code>	RDF/XML, den äldsta formen av serialiserad RDF-data. Kan vara lämpligt för att vidarebehandla informationen med RDF-baserade verktyg.
<code>text/turtle</code>	<code>.ttl</code>	<u>Turtle</u> , ett RDF-format som är enklare för människor att läsa.

5.2 Sökning och bläddring

Det går även att ställa frågor mot tjänstelagret och få ut listor på dokument som matchar de parametrar som angetts. Parametrarna anges i URL:en och kan omfatta både dokumentegenskaper och dokumentinnehåll. Den grundadress som frågningen sker mot är

`http://service.lagrummet.se/-/publ`. Svaren returneras som JSON. För varje dokument i träfflistan anges grundläggande information om dokumentet inklusive den adress där den utökade metadatan för dokumentet kan hittas (se avsnitt 5.1).

Observera att bas-URL för nedanstående sökexempel är:

`http://service.lagrummet.se/`

Generella frågor: `q` är den generella frågeparametern. Den söker i samtliga parametrar samt textkroppen. Exempel (returnerar alla dokument som i titel, dokumenttext eller annars innehåller ett ord som börjar med "ekonom"):

```
/-/publ?q=ekonom*
```

Fritextsökning i fält: Vissa dokumentegenskaper kan innehålla fritext som kan fritextsökas i. Dessa är för närvarande `title` (dokumentets titel, exv "Inkomstskattelag (1999:1229)" och `identifier` (dokumentets beteckning, exv "SFS 1999:1229"). Exempel (returnerar alla dokument vars titel innehåller något ord som innehåller texten "djur"):

```
/-/publ?title=*djur*
```

Dokumenttyp: Med parametern `type` (som kan anges flera gånger) kan man avgränsa frågan till en eller flera dokumenttyper. Se nedan under "Tillgängliga parametrar" för hur hela listan av möjliga dokumenttyper kan fås. Exempel (returnerar alla dokument av typen `Lag` eller `Forordning`):

```
/-/publ?type=Lag&type=Forordning
```

Publikationssamling: De flesta dokument i rättssystemet ingår i någon typ av publikationssamling. En `Myndighetsforeskrift` ingår i en `Forfattningssamling`, ett `Rattsfallsreferat` ingår i en `Rattsfallspublikation`, osv. Varje sådant dokument har en egenskap vars värde är URI:n för den publikationssamlingen. Se nedan under "Tillgängliga parametrar" för hur hela listan av möjliga publikationssamlingar kan fås.

Olika dokumenttyper använder olika sorters publikationssamlingar. De som finns definerade är:

- `allmannaRadSerie`
- `forfattningssamling`
- `rattsfallspublikation`
- `utredningsserie`

Finansinspektionens författningssamling (FFFS) har t.ex. URI:n `http://rinfo.lagrummet.se/serie/fs/fffs`, så en sökning på alla dokument i FFFS blir således:

```
/-/publ?forfattningssamling.iri=http://rinfo.lagrummet.se/serie/fs/fffs
```

Likvärdigt men kortare exempel som matchar alla dokument vars
forfattningssamling slutar på /fffs:

```
/-/publ?forfattningssamling.iri=*/fffs
```

Filtrering på förekomst av en viss egenskap: För att inkludera eller
exkludera dokument baserat på förekomsten av en viss egenskap oavsett
värde kan en parameter prefixad med `exists-` anges. Värdet ska då anges som
en boolean, d.v.s. `true` eller `false`. Nedanstående exempel söker på alla
dokument vilka *inte* hänvisar till andra dokument med egenskaperna `andrar`
eller `upphaver`:

```
/-/publ.json?exists-andrar.iri=false&exists-upphaver.iri=false
```

Filtrering på datum: Genom att prefixa datumparametrar med speciella
nyckelord går det att avgränsa sökningen inom specifika tidsangivelser.
Datum ska anges på formen YYYY-MM-DD. Följande prefix är definierade:

`minEx-`

Datumet för den angivna parametern måste vara större än (senare än)
det angivna datumet (`Ex` står för "exclusive").

`min-`

Datumet för den angivna parametern måste vara större än eller lika med
det angivna datumet.

`maxEx-`

Datumet för den angivna parametern måste vara mindre än (tidigare än)
det angivna datumet.

`max-`

Datumet för den angivna parametern måste vara mindre än eller lika
med det angivna datumet.

Det går även att förenkla avgränsningen till inom ett specifikt år genom att
använda prefixet `year-`.

Villkorlig filtrering på datum: Genom att prefixa en filtering med `ifExists-`
kan man exkludera resultat till ett visst datumomfång förutsatt att
datumegenskapen existerar.

Nedanstående exempel använder sig av en kombination av mekanismerna
ovan för att söka på *grundförfattningar som var ikraft första januari 2011*:

```
/-/publ.json?max-ikrafttradeddatum=2011-01-01&  
ifExists-minEx-rev.upphaver.ikrafttradeddatum=2011-01-01&
```

```
exists-andrar.iri=false&  
exists-upphaver.iri=false
```

Detta åstadkoms genom att kräva följande av matchande dokument:

- *Ikraftträdandedatumet* inte får vara större än den första januari 2011:
`max-ikrafttradandedatum=2011-01-01`
- *Ikraftträdandedatumet för något som upphäver dokumentet* måste vara större än samma datum, om något sådant existerar:
`ifExists-minEx-rev.upphaver.ikrafttradandedatum=2011-01-01`
- Exkludera dokument som *ändrar* andra dokument:
`exists-andrar.iri=false`
- Exkludera dokument som *upphäver* andra dokument:
`exists-upphaver.iri=false`

Alternativa matchningar: Alla egenskaper som prefixas med `or-` grupperas tillsammans så att enbart en av dem måste matcha. Detta kan användas för att t.ex. söka samtidigt på *antingen* beslutsdatum och utfärdandedatum inom ett visst år:

```
/-/publ?q=&or-year-beslutsdatum=1991&or-year-utfardandedatum=1991
```

Sortering: Många av de parametrar som man kan avgränsa sökningen på kan också användas för att sortera sökresultaten. Parametern `_sort` används, med parametervärdet satt till den parameter man vill sortera på. Sorteringen är vanligen i stigande ordning, men kan fås i fallande ordning om parametervärdet föregås av ett minustecken (-). Exempel (alla dokument i FFFS, sorterat i fallande beslutsdatumordning):

```
/-/publ?forfattningssamling.iri=*/fffs&_sort=-beslutsdatum
```

Paginering: Om inget annat anges returneras endast de första 50 träffarna i en träfflista. Nyckeln `totalResults` i svaret anger hur många träffar som finns totalt för frågan. Genom att ange parametern `_page` (med en siffra som värde, första sidan har sidnummer 0) returneras efterföljande "sidor" med träffar från listan. Genom att ange parametern `_pageSize` med en siffra som värde kan antalet träffar på en sida ändras från 50. Exempel (Listar dokument 21-40 i FFFS):

```
/-/publ?forfattningssamling.iri=*/fffs&_pageSize=20&_page=1
```

Nycklarna `prev` och `next` i resultatet innehåller direktlänkar till föregående/nästa träfflistesida för just den sökningen.

Listor av tillgängliga värden: De "fasta" värden som kan användas för ovanstående parametrar beror på vilken mängd dokument som för tillfället finns inladdat i rättsdatalogret. Från nedanstående vy kan man utläsa vilka parametrar och möjliga värden som finns (notera att vissa parametrar med fritextvärden, som `q`, `title` och `identifier`, inte finns med i nedanstående):

```
/-/publ;stats
```

Det går även att inkludera listor på möjliga värden *inom* ett visst sökresultat (för att stödja s.k. "facetterad" sökning).

```
/-/publ?type=Myndighetsforeskrift&_stats=on
```

5.3 Verktyg för sökning, bläddring och dokumentvisning

Det finns även ett verktyg för att bläddra/söka i rättsdatalogret. Detta kan användas för att utforska API:t och se möjliga värden för olika sökparametrar. För varje träfflista visas även exakt den frågesträng som ger motsvarande resultat.

För varje dokument visas även den utökade informationen som tjänstelagret gör tillgängligt. Verktyget hittas på nedanstående adress:

```
http://service.lagrummet.se/ui/
```

5.4

Liksom för dokument finns det sök- och bläddringsbara index för de organisationer och dokumentserier som ingår i systemet. Dessa nås via:

```
/-/org?_pageSize=1000
```

respektive:

```
/-/serie?_pageSize=1000
```

(`_pageSize` är högt satt här för att visa hur man hämtar hem fullständiga listor. Det är idag inte sannolikt att dessa uppgår till fler än 1000. Om så ändå skulle ske i framtiden kommer förstas resultatlistan att vara försedd med pagineringslänkar så att detta är explicit.)

5.5 Kompletterande resurslistor

En komplett lista med etiketter för resurstyper och egenskaper som ingår i systemet går att nå på:

/var/terms.json