

## Vill du medverka i projektet?

Du kan bidra med teknisk kunskap i produkt-utformning (inkluderar programutveckling, webanvändargränssnitt, layout, testning, mm), administrativ erfarenhet vid behandling av inkommande datafiler liksom att rapportera in egna forskningsresultat.

Hör av Dig till någon av oss!

### KONTAKTER

**Christer Gustavsson** (DIS), [christer.gustavsson@dis.se](mailto:christer.gustavsson@dis.se)

**Anders Ardo** (EIT, Lunds universitet), [anders.ardo@eit.lth.se](mailto:anders.ardo@eit.lth.se)  
[www.eit.lth.se/staff/Anders.Ardo](http://www.eit.lth.se/staff/Anders.Ardo)

### INFORMATION

DisForum

[www.dis.se/sv/projekt/genealogisk-databas.html](http://www.dis.se/sv/projekt/genealogisk-databas.html)

Internetfonden

[www.internetfonden.se/rikstackande-genealogisk-databas](http://www.internetfonden.se/rikstackande-genealogisk-databas)

Testsystem

<http://rgd.eit.lth.se:8085/>



FINANSIELLT STÖD FRÅN

**.se | internetfonden**

# RGD

RIKSTÄCKANDE GENEALOGISK DATABAS

## Sammanslagen släktforskning

Sveriges historiska befolkning med släktrelationer

Web-tjänster

**RIKSTÄCKANDE GENEALOGISK DATABAS (RGD)**  
är ett samarbetsprojekt mellan DIS ([www.dis.se](http://www.dis.se)) och  
EIT vid Lunds universitet ([www.eit.lth.se](http://www.eit.lth.se))

**Övergripande målsättning med en Rikstäckande Genealogisk Databas (RGD) är att skapa en databas över Sveriges historiska befolkning med släktrationer och utan dubletter – baserat på släktforskares samlade forskningsresultat.**

#### FUNKTIONALITET

- Indatavalidering / egenkontroll av GEDCOM-filer (1-3)
  - Namnkontroll, listar formella fel och möjliga felregistreringar av kön.
  - Ortkontroll, listar angivna platser som inte är svensk församling eller land.
  - Dubblettkontroll, listar möjliga dubblettkandidater.
- Matchning av två GEDCOM filer (4, 5)
 

Släktforskare, med viss del av forskningen gemensam, kan analysera likheter/skillnader
- Avvikelser i matchat data (6)
 

Forskare som hittat gemensamma anor kan hitta detaljskillnader hos matchade personer.
- Sammanslagning av matchat släktforskningsdata (7)
 

Två matchade GEDCOM filer kan sammanföras till en gemensam fil utan dubletter.
- Skapa GEDCOM fil av sammanslagen släktforskningsdata (8)
 

Resultatet från sammanslagningen kan laddas in i ett släktforskningsprogram.

#### TEKNIK/IMPLEMENTERING

Machine Learning,  
Information Retrieval,  
Python, MongoDB, Lucene,  
SVM

#### ROADMAP

- Alpha-release
  - slutet oktober 2014
- Beta-release
  - slutet november 2014
- System 0.5
  - slutet januari 2015
- System 1.0
  - maj 2015

