

Matching study to registry data: maintaining data privacy in a study on family based colorectal cancer

**Daniel NASSEH ^a, Jutta ENGEL ^{a,b}, Ulrich MANSMANN ^a,
Werner TRETTER ^b and Jürgen STAUSBERG ^a**

a IBE, Ludwig-Maximilians-Universität München, Germany

b Munich Cancer Registry, Germany

02.09.2014 - Istanbul



Study of family based colorectal cancer

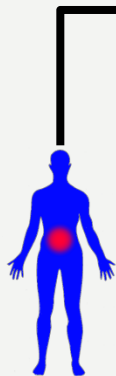
MIE 2014



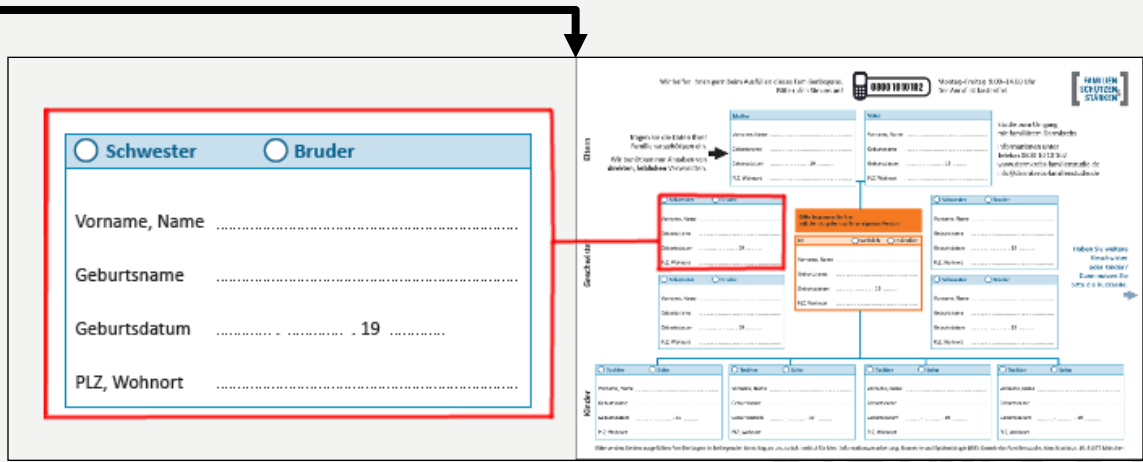
Index-Patient
Newly diagnosed with
colorectal cancer



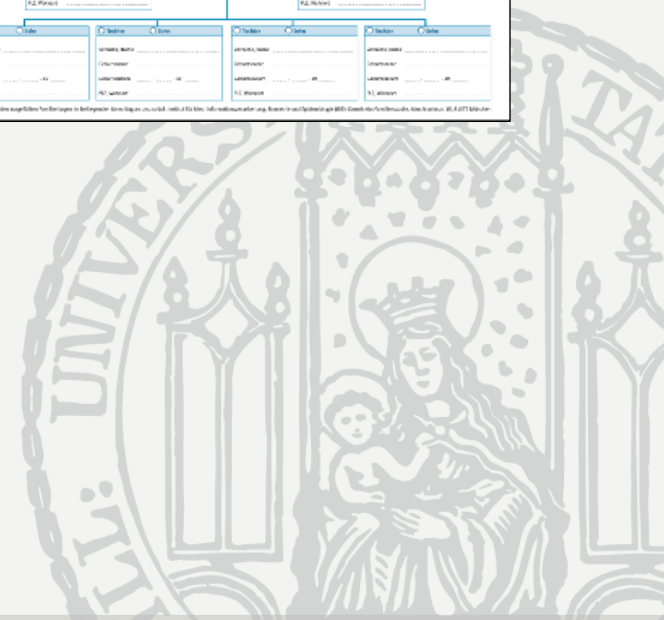
Study of family based colorectal cancer



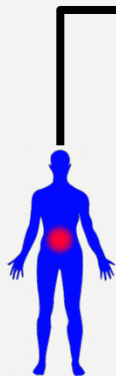
Index-Patient
Newly diagnosed with
colorectal cancer



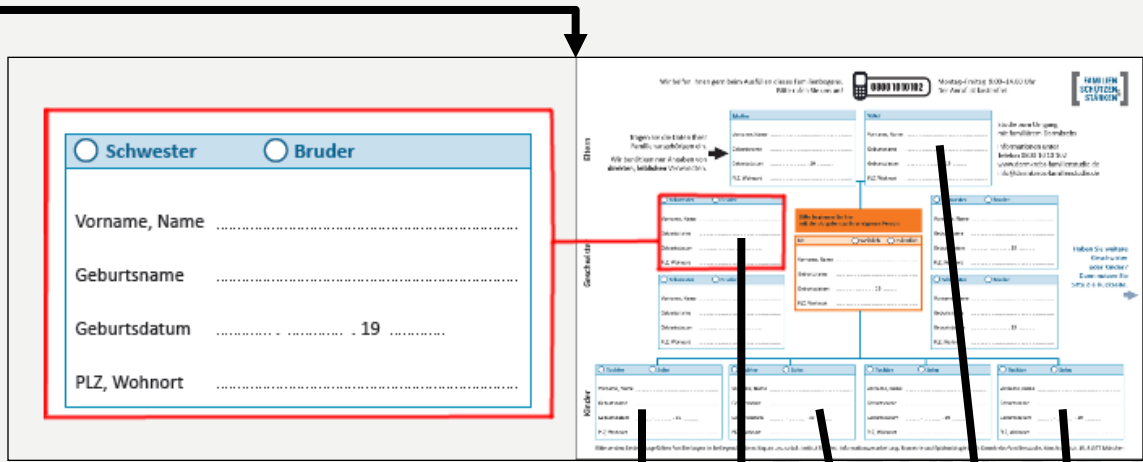
*IDAT: Firstname, Lastname,
Date of Birth, Adress ...*



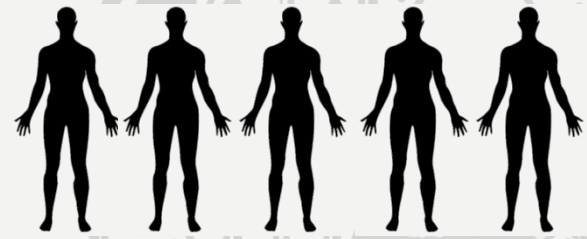
Study of family based colorectal cancer



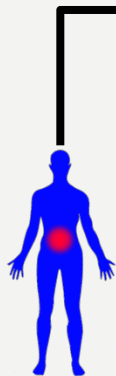
Index-Patient
Newly diagnosed with
colorectal cancer



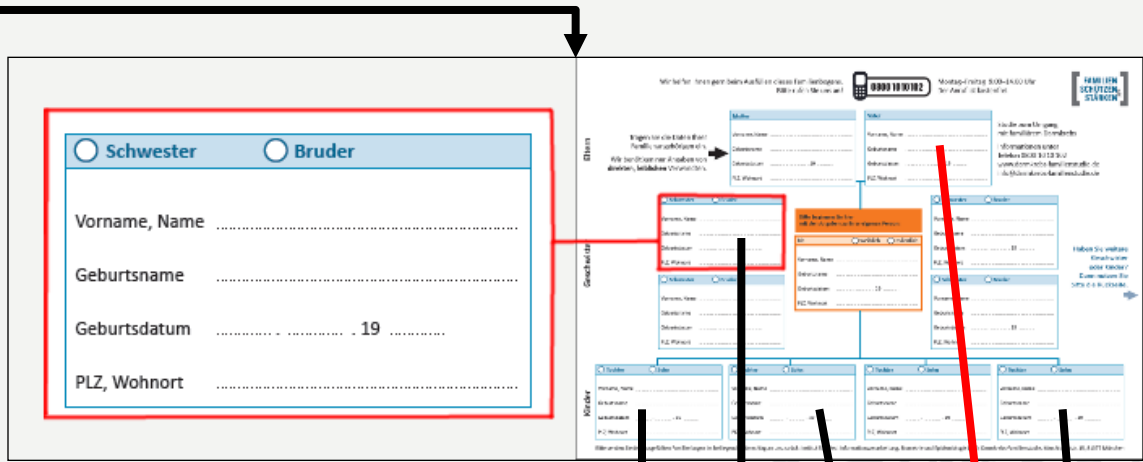
*IDAT: Firstname, Lastname,
Date of Birth, Adress ...*



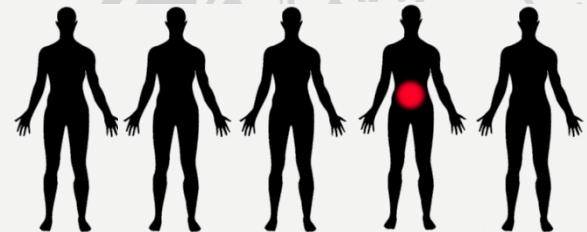
Study of family based colorectal cancer



Index-Patient
Newly diagnosed with
colorectal cancer

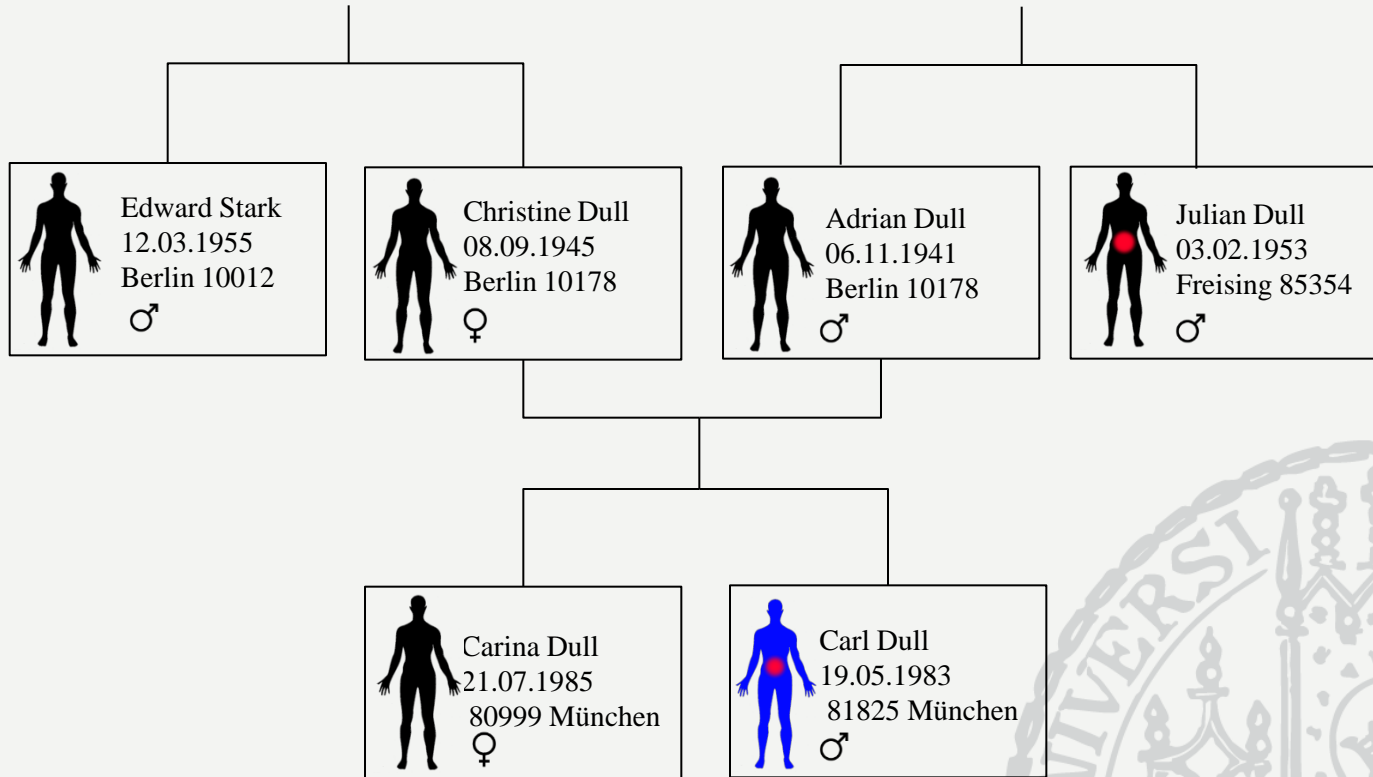


*IDAT: Firstname, Lastname,
Date of Birth, Adress ...*

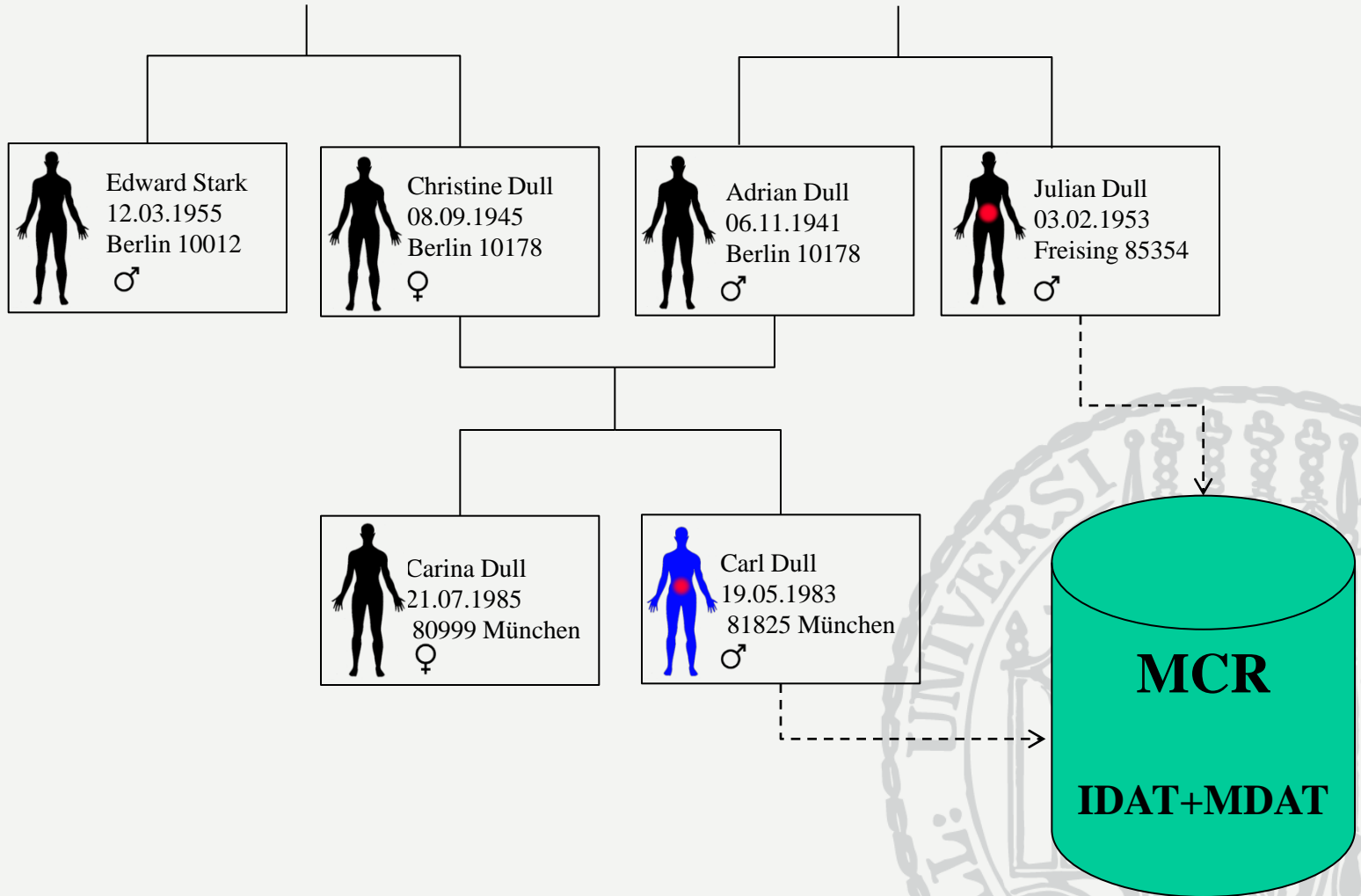


The task: Identifying relatives who had a history of cancer in the past. If we find one, we also need the corresponding medical data.

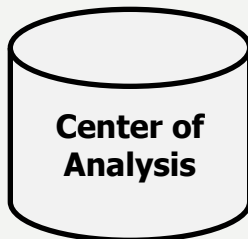
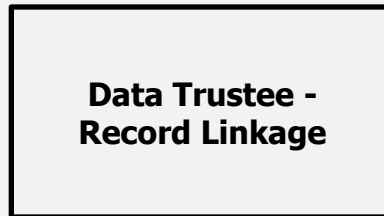
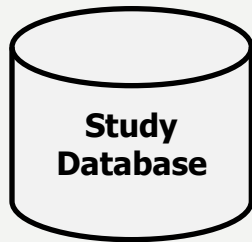
Identification: Matching study data to registry data (IDAT)



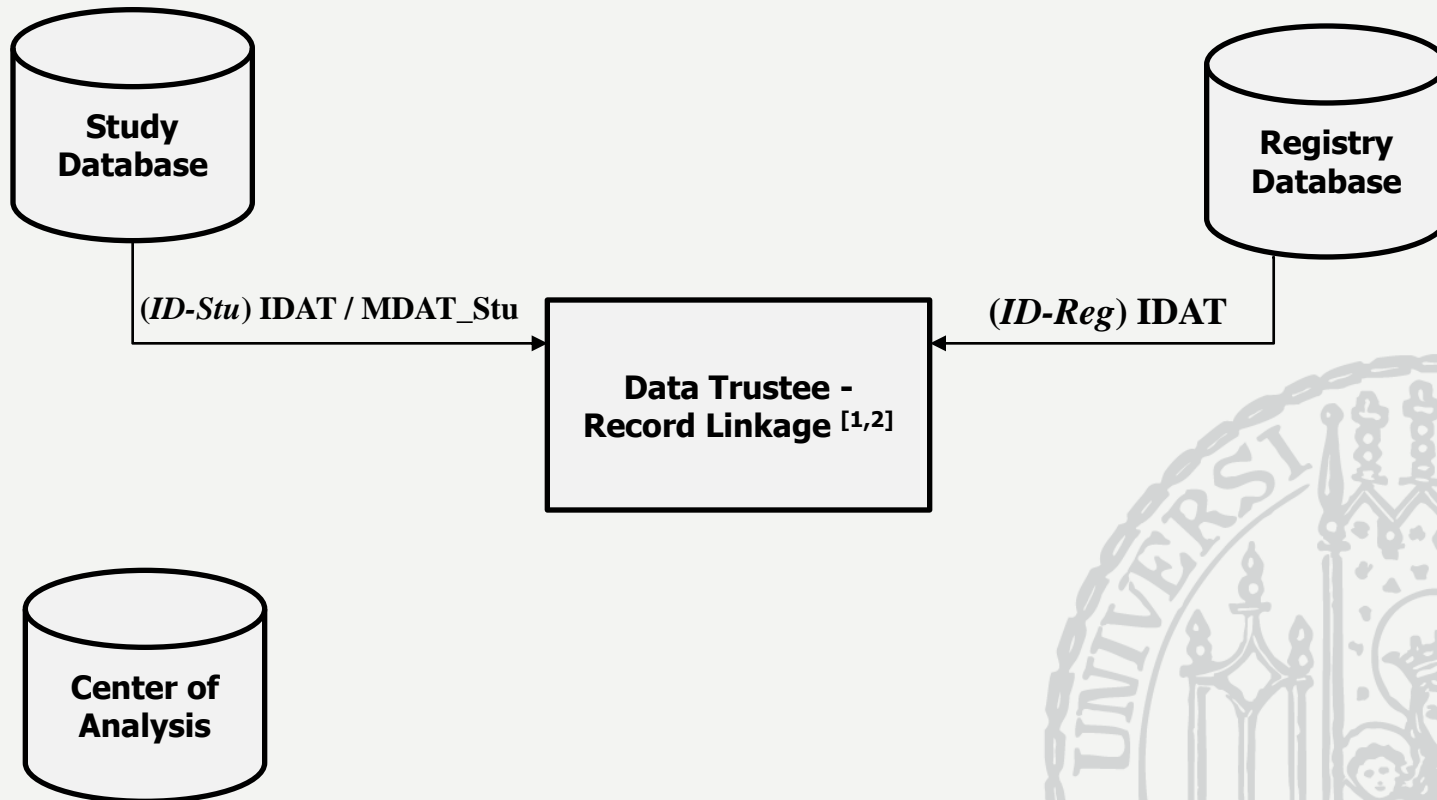
Identification: Matching study data to registry data (IDAT)



I. Institutional and organisational division of participating parties.



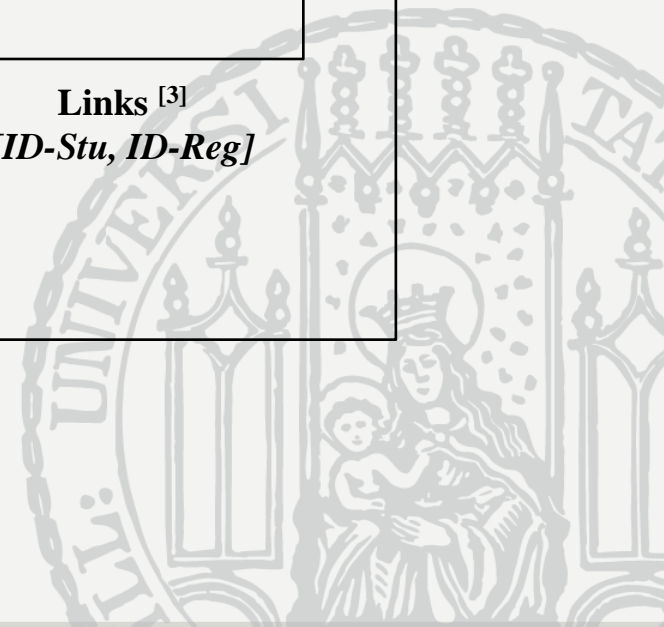
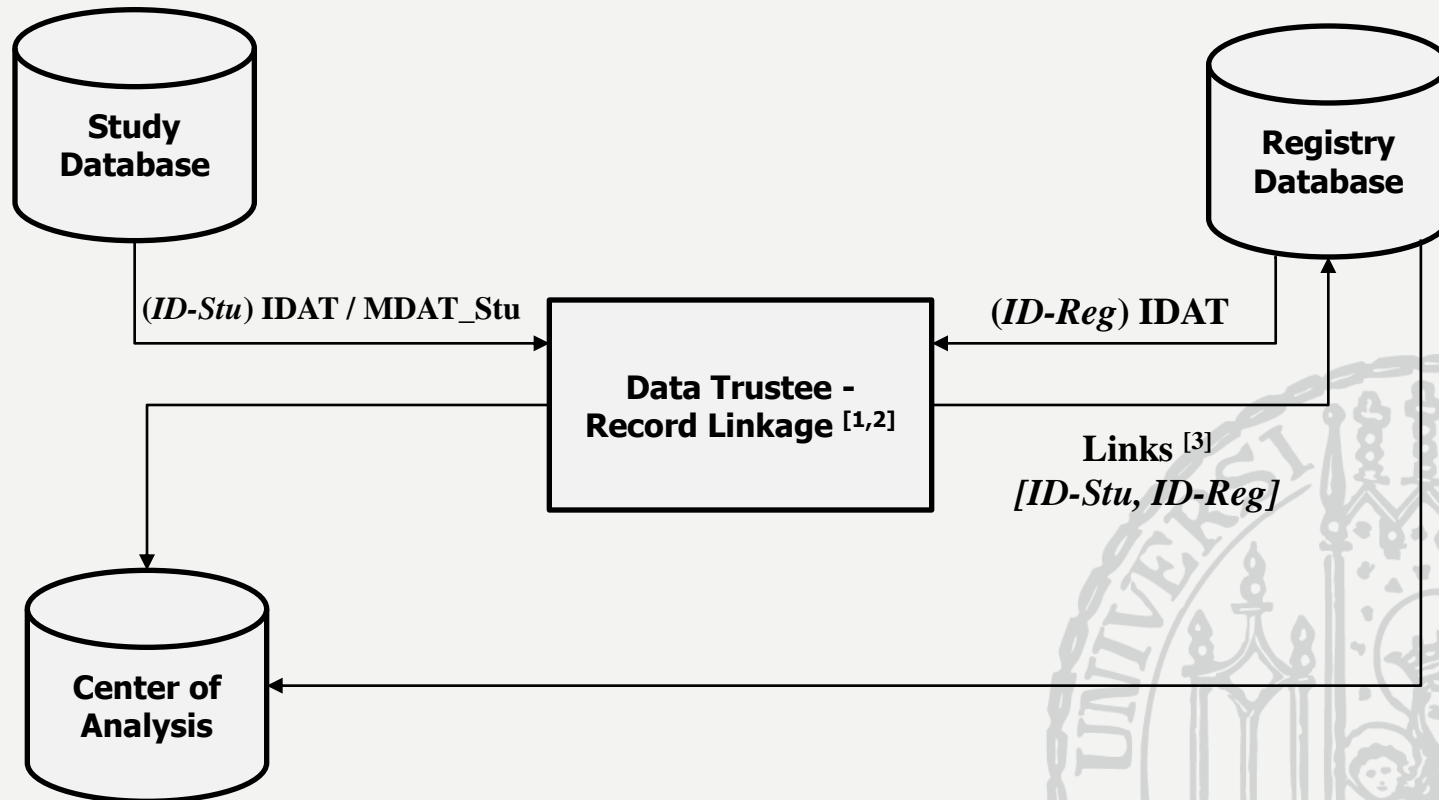
I. Institutional and organisational division of participating parties.



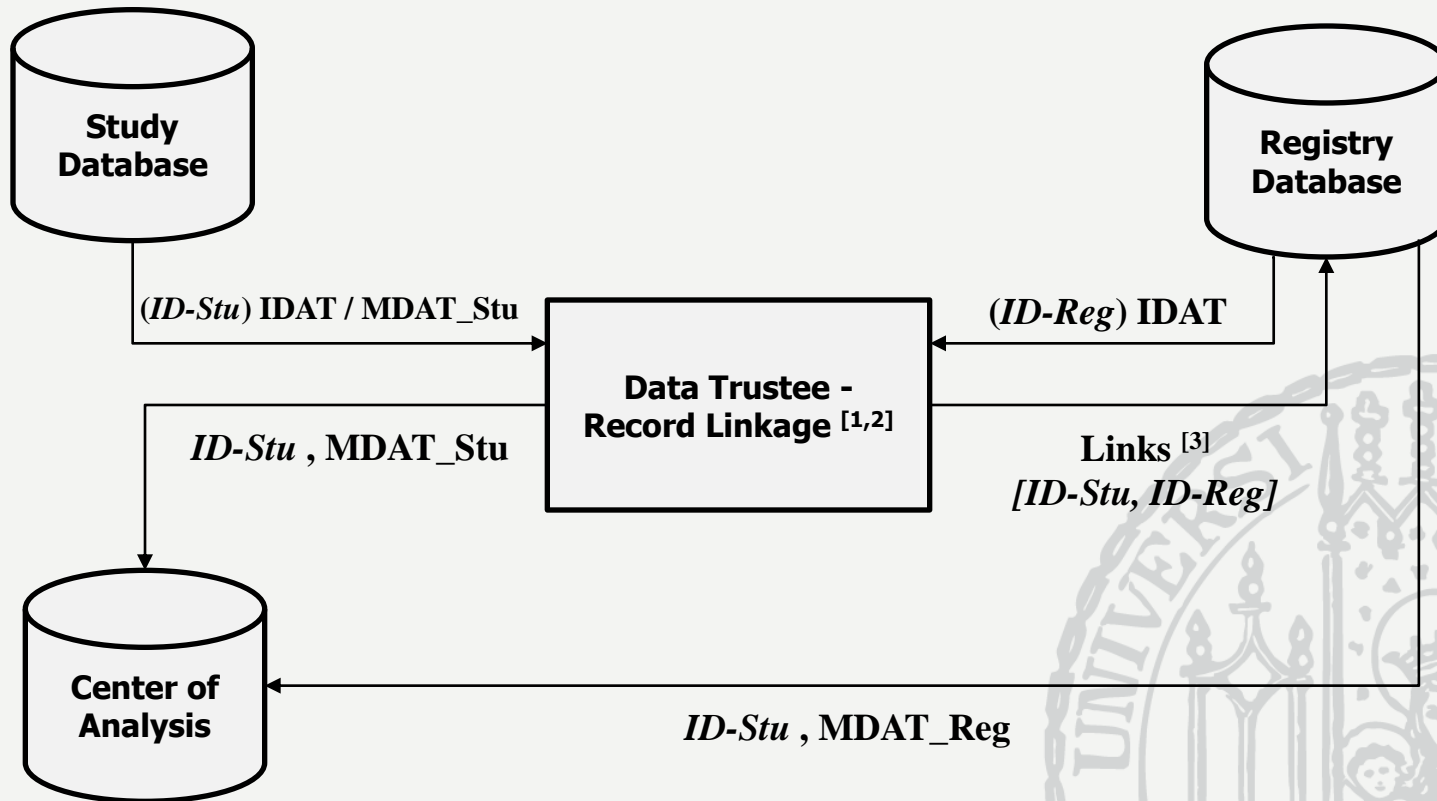
[1] Meyer M. Kontrollnummern und Record Linkage. Das Manual der epidemiologischen Krebsregistrierung. Hentschel S, Katalinie A, editor. Zuckschwerdt. 2011;57-68.

[2] Christen P. Data Matching – Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection. Heidelberg: Springer; 2012.

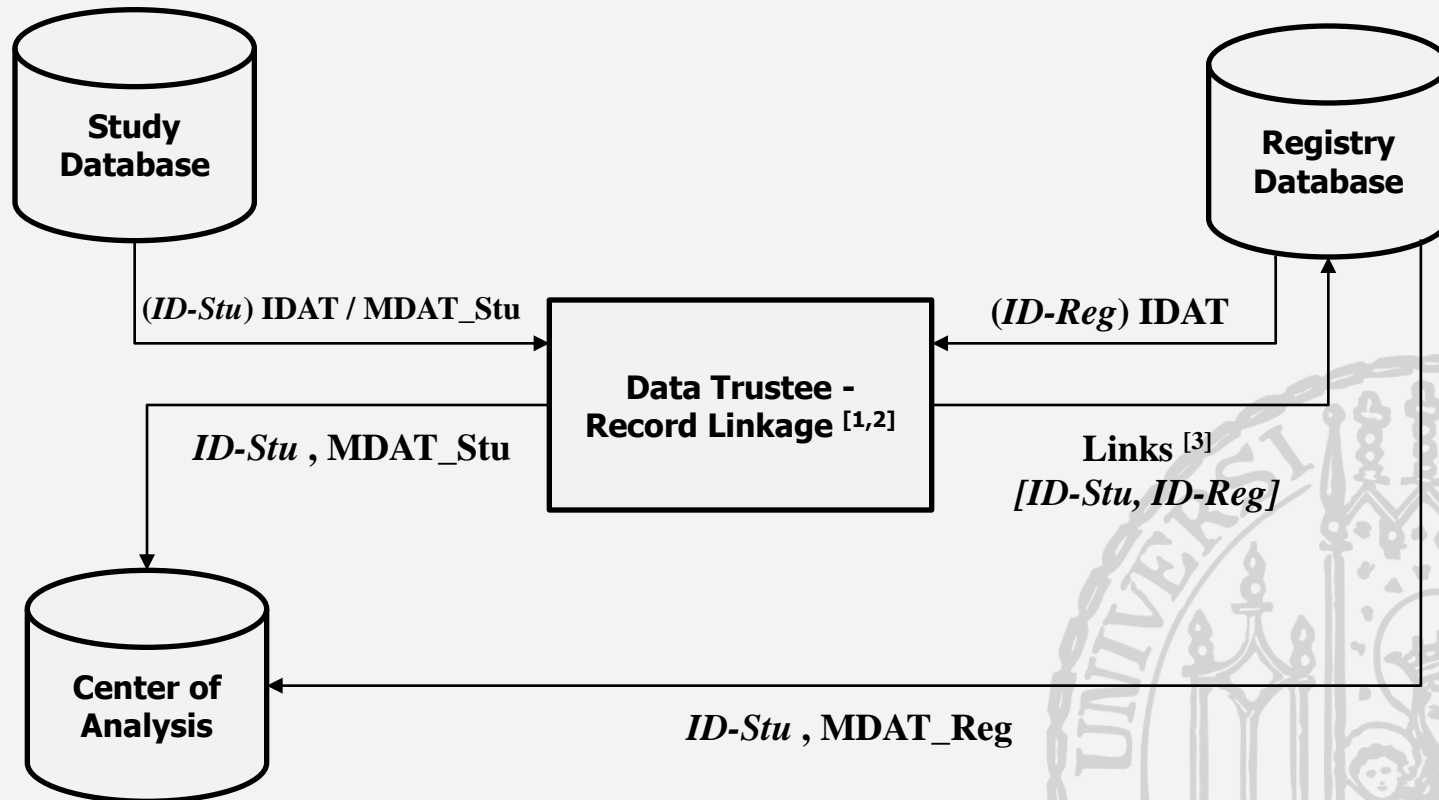
I. Institutional and organisational division of participating parties.



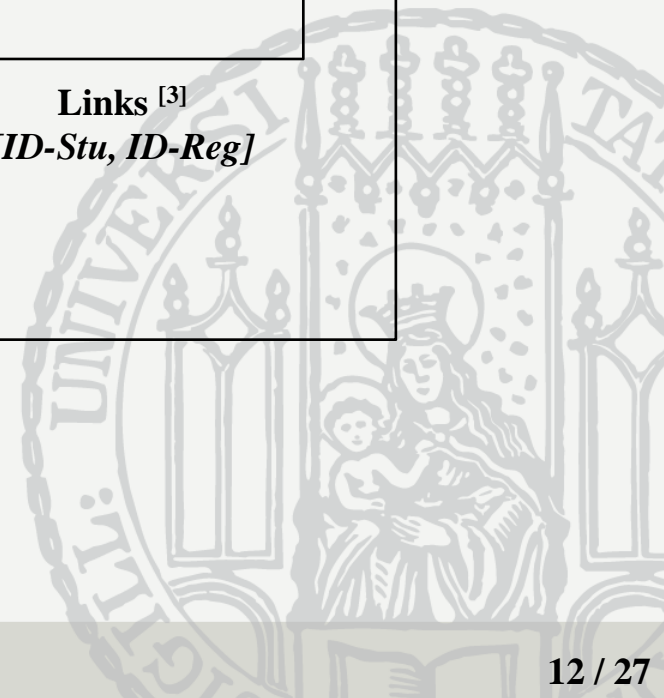
I. Institutional and organisational division of participating parties.



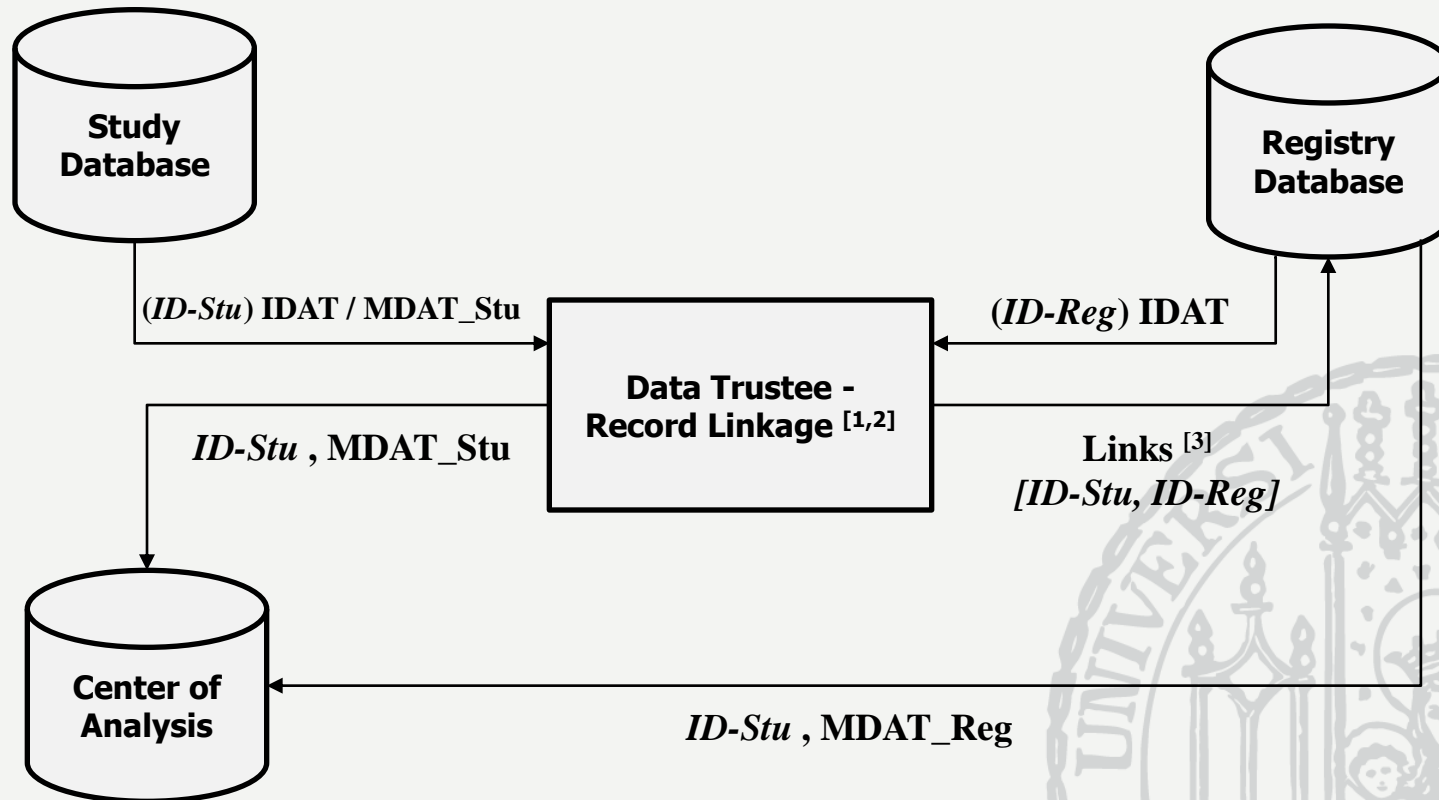
I. Institutional and organisational division of participating parties.



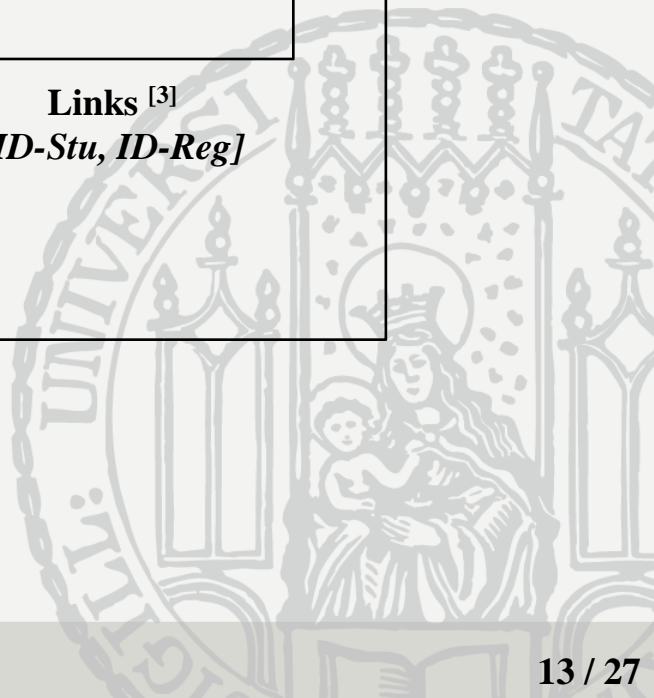
MDAT_Stu <-> MDAT_Reg



II. No transfer of identifying data in human readable form



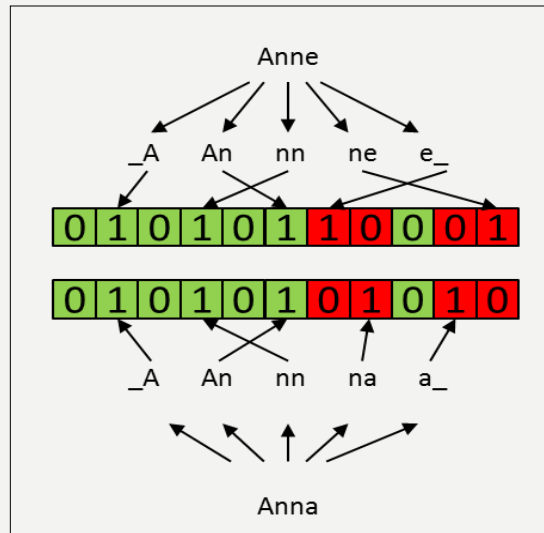
MDAT_Stu <-> MDAT_Reg



Variant 1: Apply Hashfunctions (e.g. SHA-2 [1]) on IDAT

Cleartext	Hashvalue
Meier	05c2d2b4cad1a3f5bf547b484ac6f4a70893e944d5bd6fe0f28db40453bf3f3c
Meyer	876fdfa1d1152c1d024386a1f66e7725f292ef83404fc4d3be79c1b51cc81c45

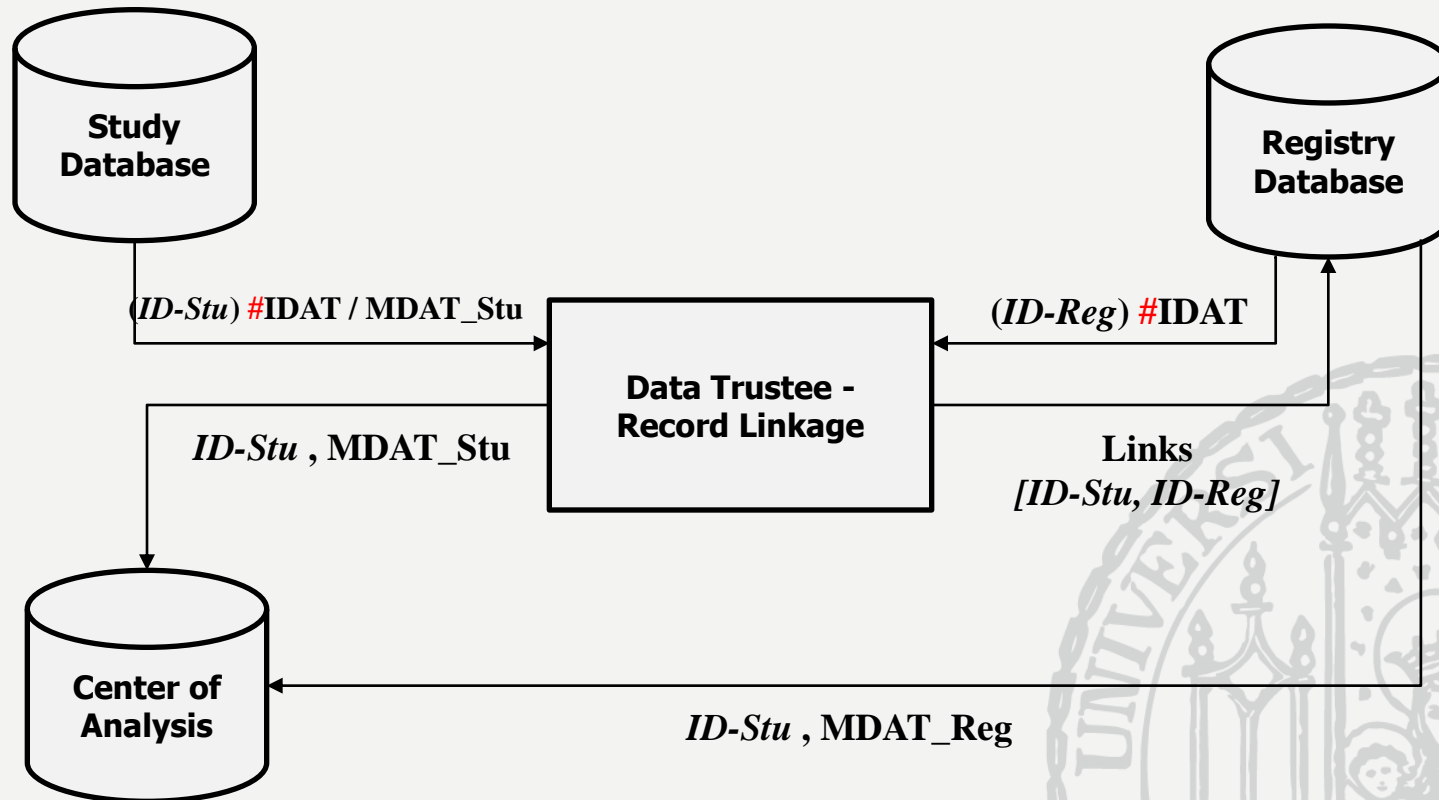
Variant 2: Transform IDAT into Bloomfilters [2]



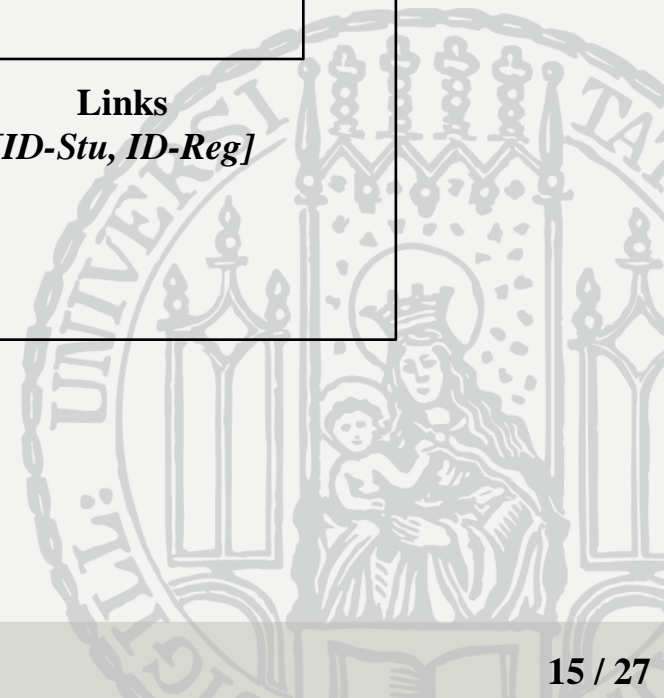
[1] Gilbert H, Handschuh H. Security Analysis of SHA-256 and Sisters. Selected Areas in Cryptography. 2003; 175–193

[2] Schnell R, Bachteler T, Reiher J. Privacy-preserving record linkage using Bloom filters. BMC Medical Informatics and Decision Making. 2009; 9:41

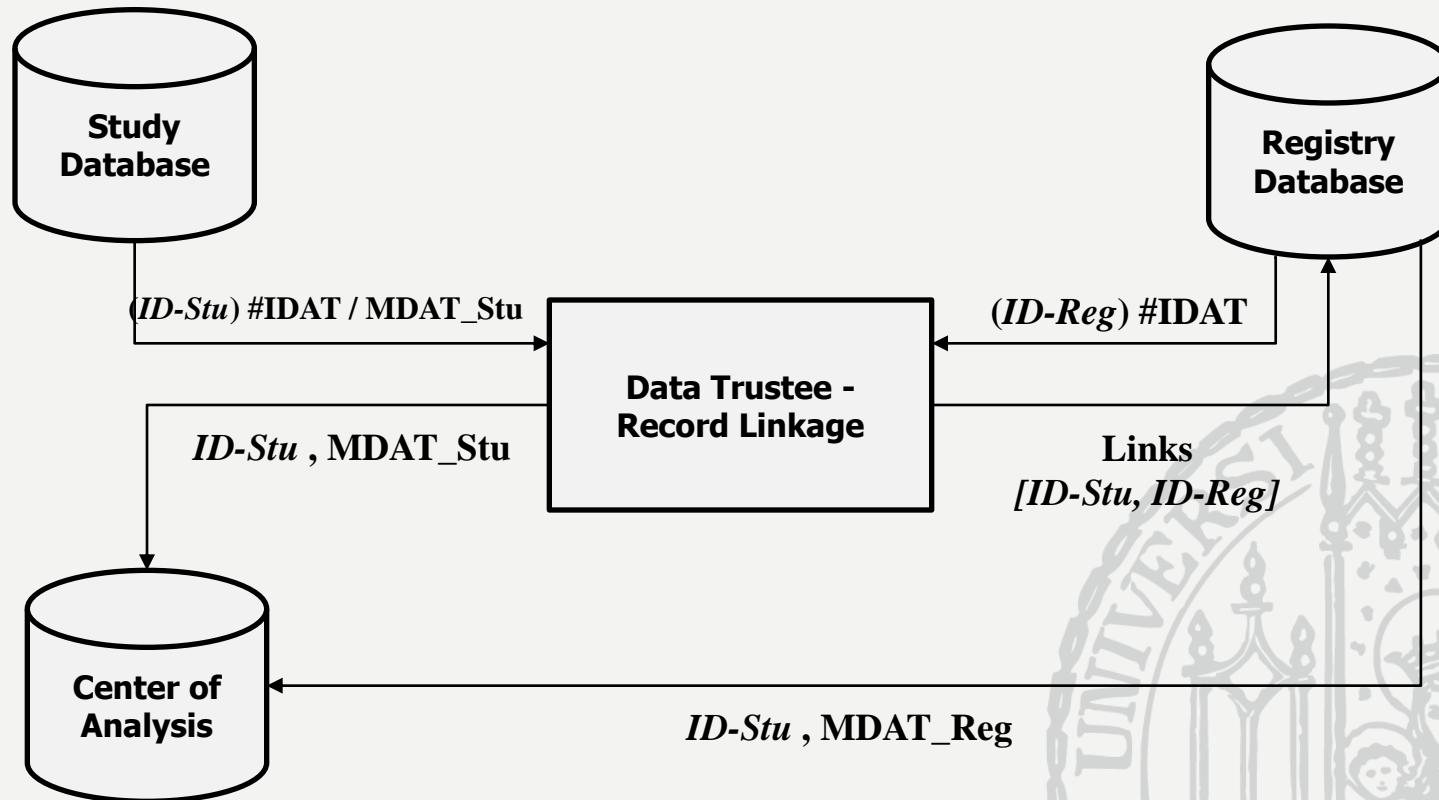
II. No transfer of identifying data in human readable form



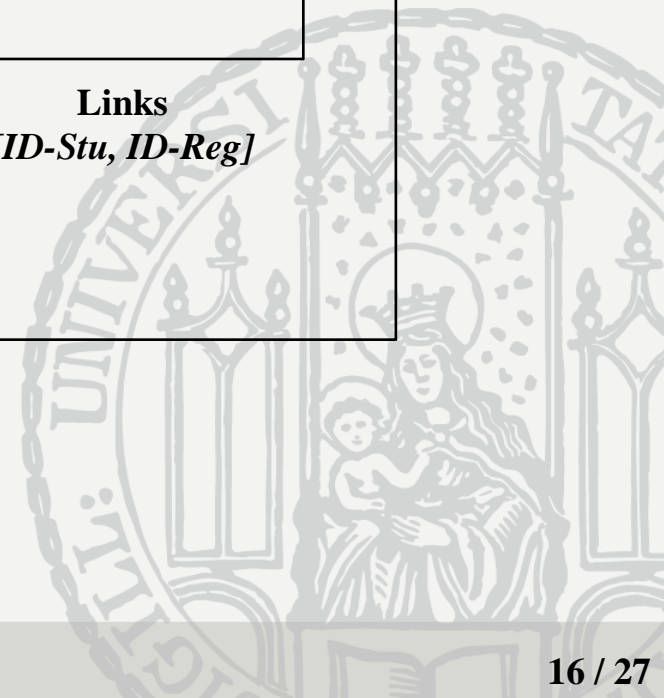
MDAT_Stu <-> MDAT_Reg



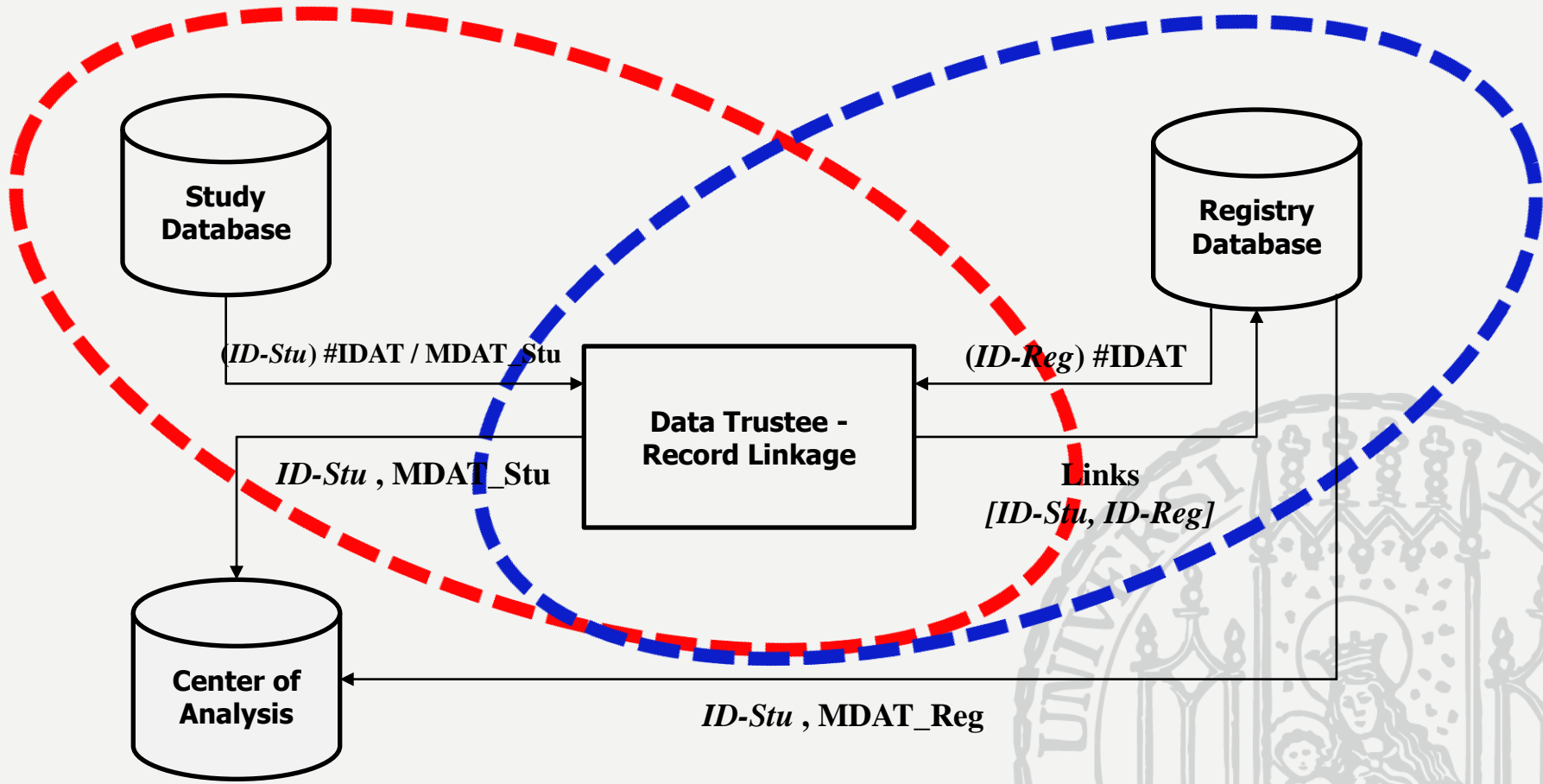
III. Registry-ID and study-ID may only be shared with the data-trustee



MDAT_Stu <-> MDAT_Reg

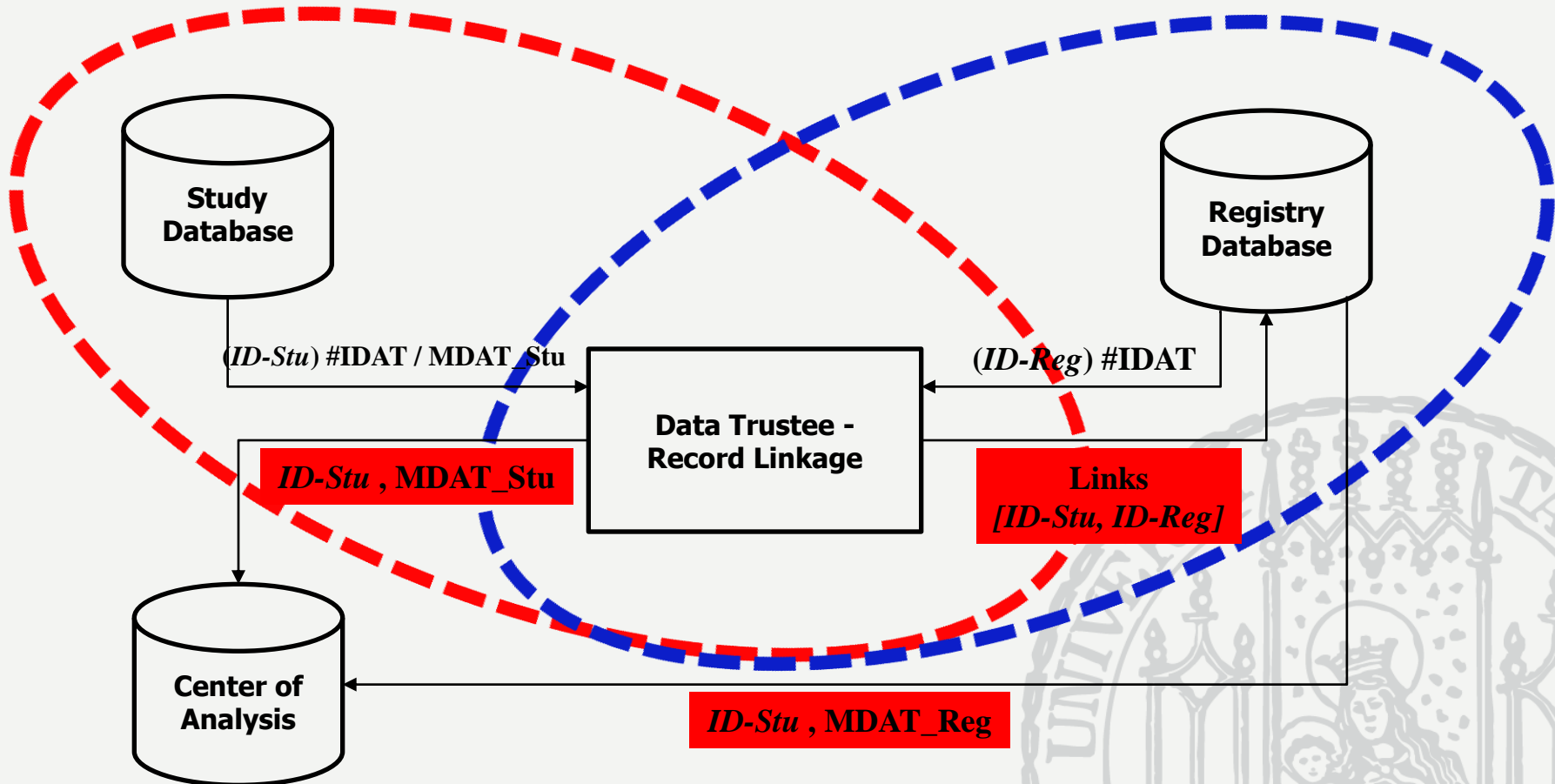


III. Registry-ID and study-ID may only be shared with the data-trustee



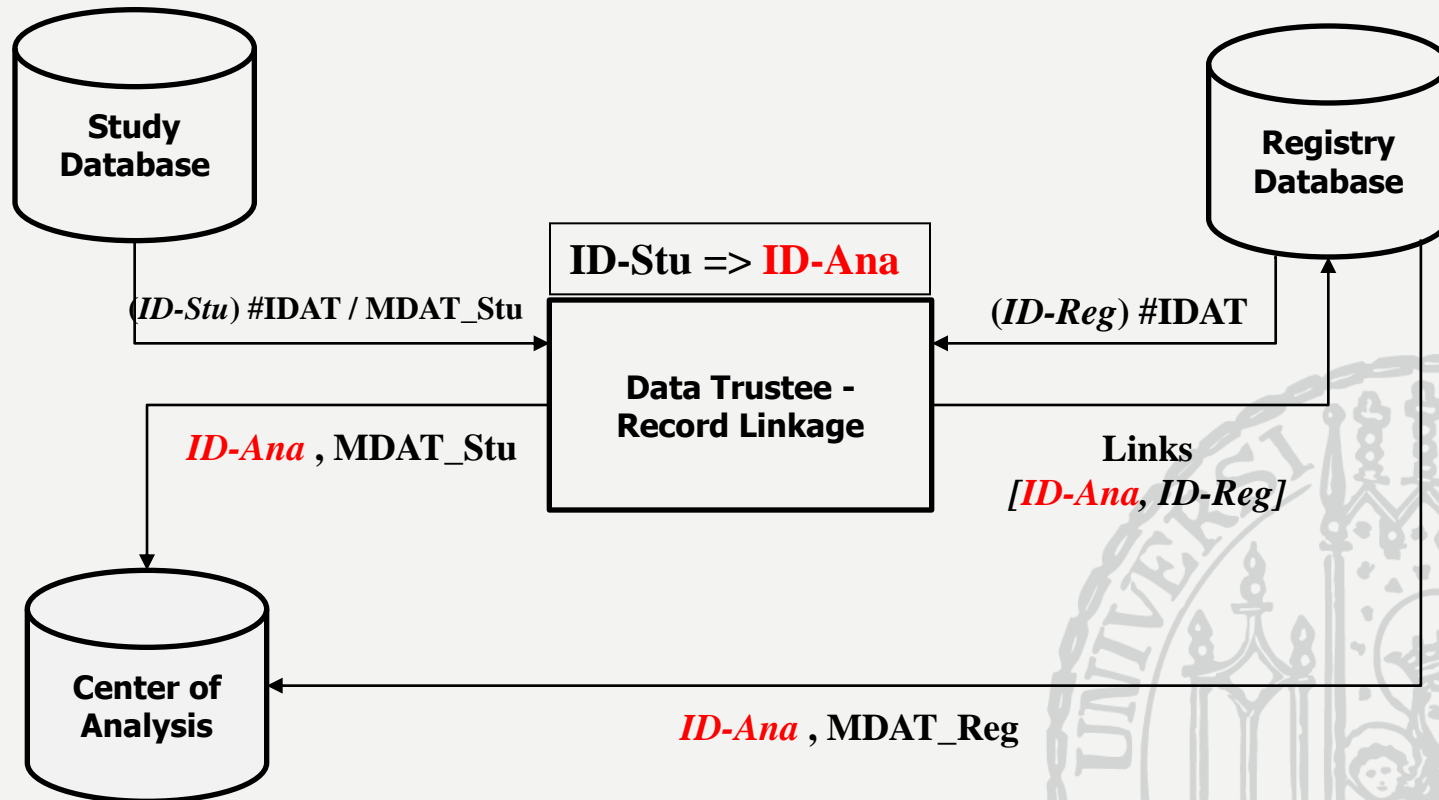
MDAT_Stu <-> MDAT_Reg

III. Registry-ID and study-ID may only be shared with the data-trustee

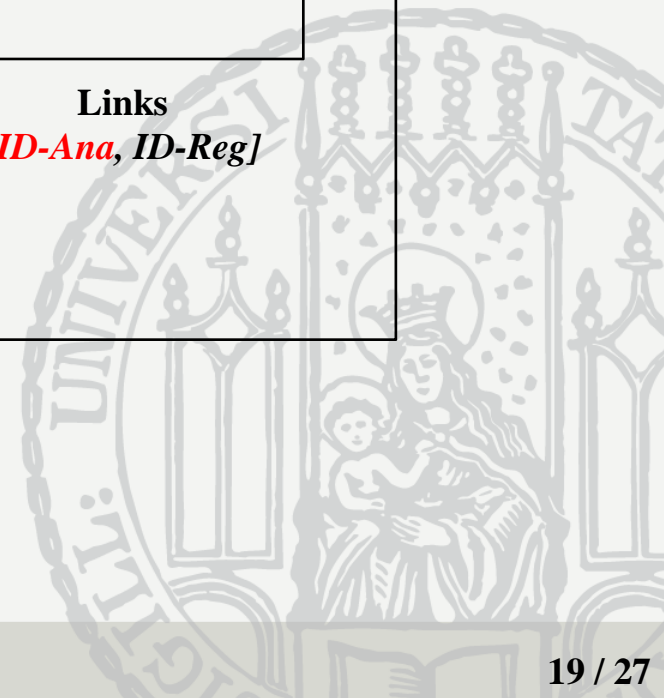


MDAT_Stu <-> MDAT_Reg

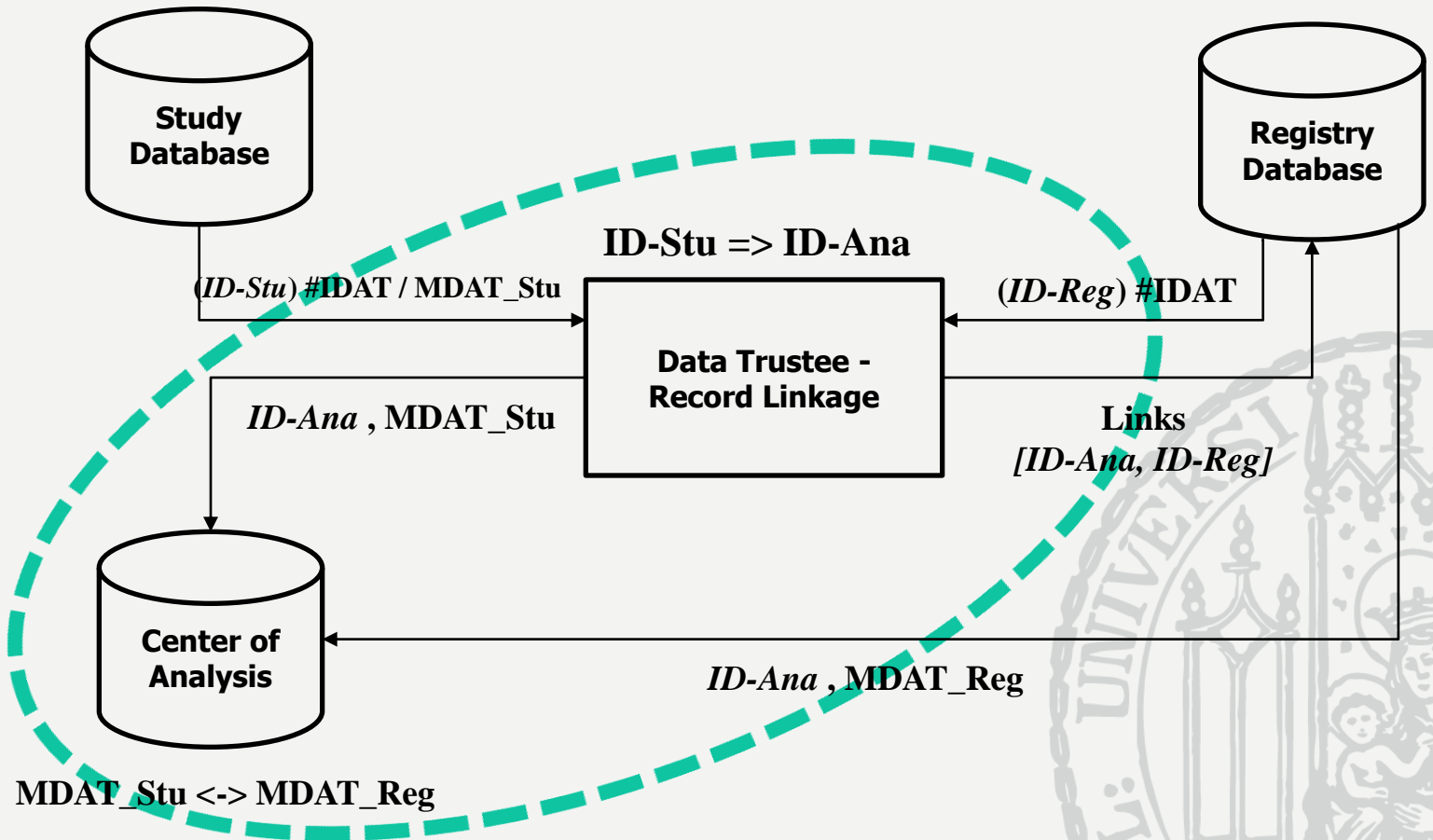
III. Registry-ID and study-ID may only be shared with the data-trustee



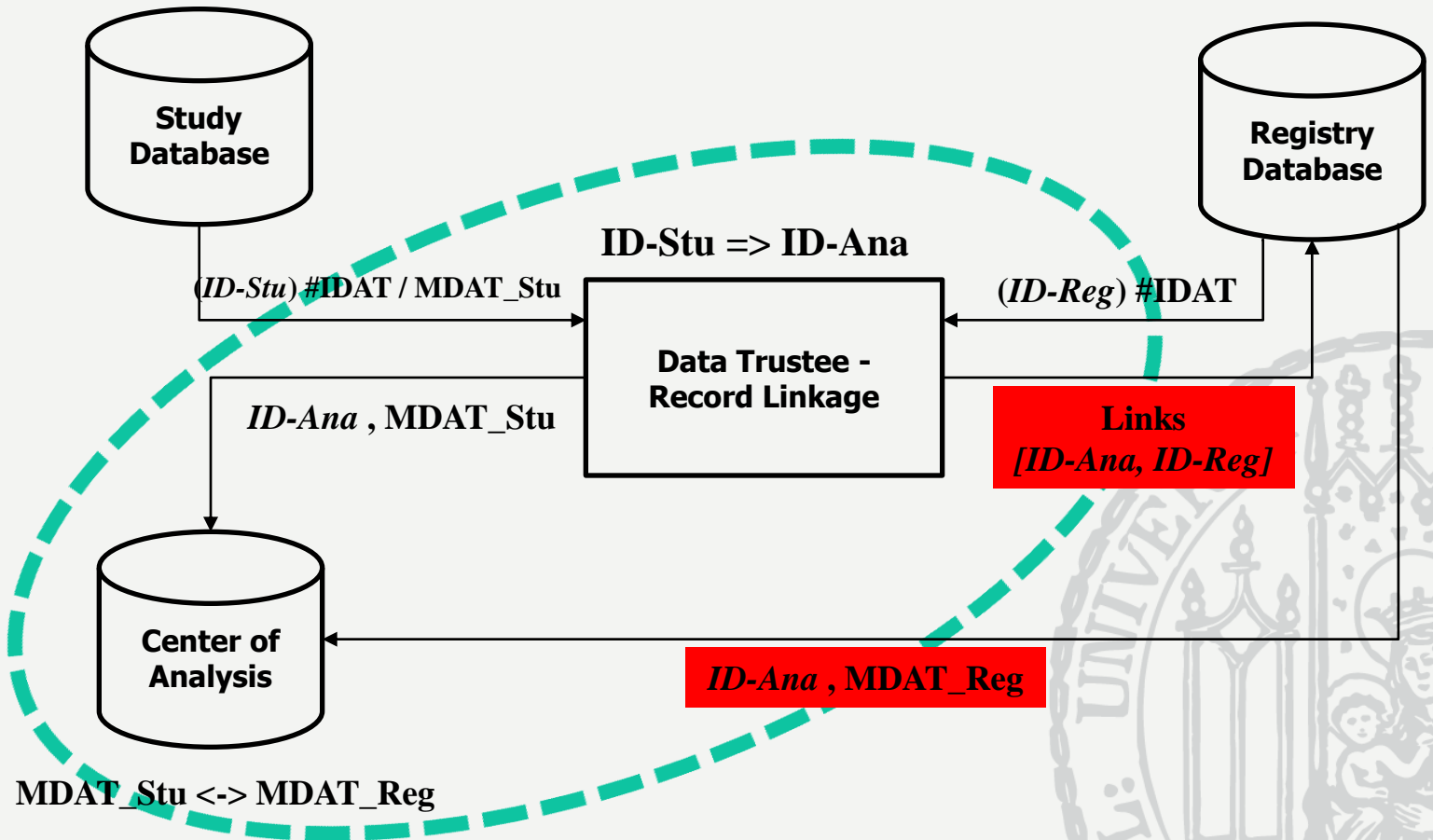
MDAT_Stu <-> MDAT_Reg



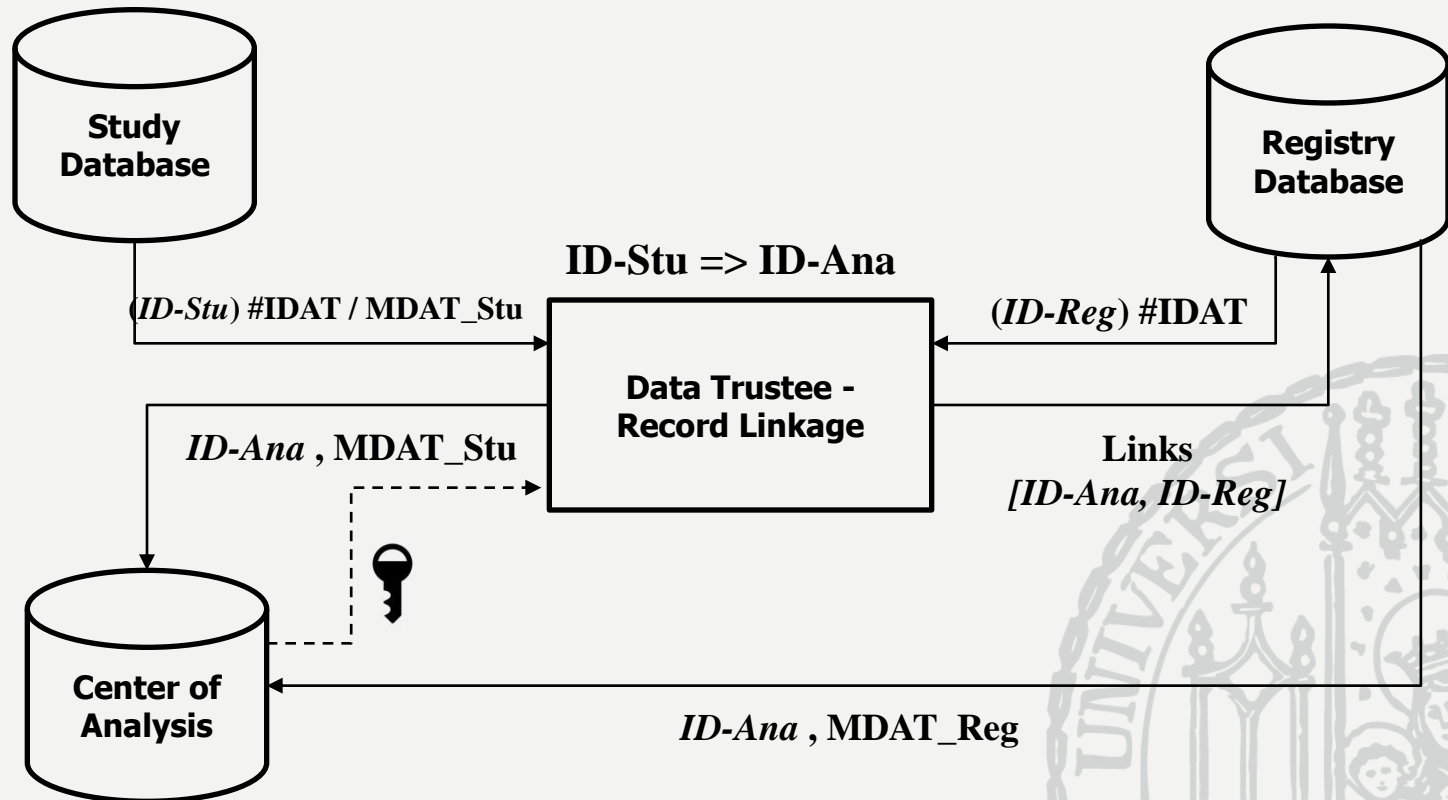
IV. Analysis-ID may only be shared between the Data Trustee and the Center of Analysis



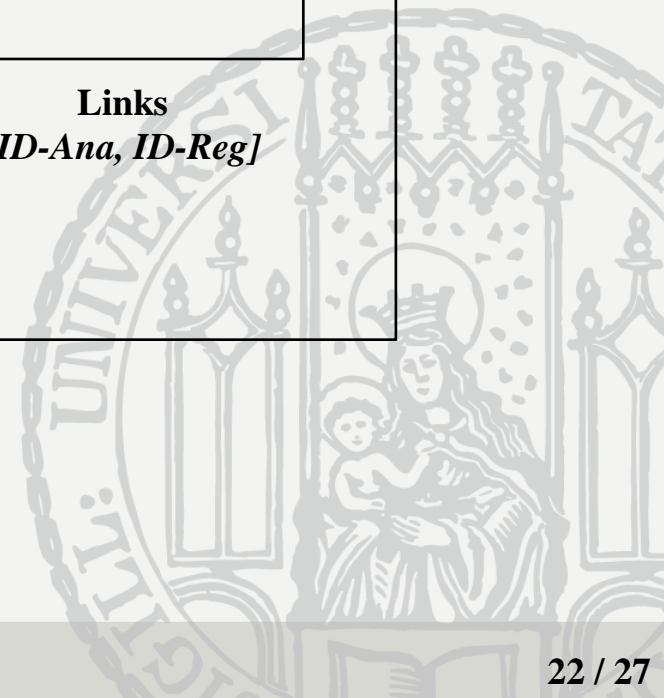
IV. Analysis-ID may only be shared between the Data Trustee and the Center of Analysis



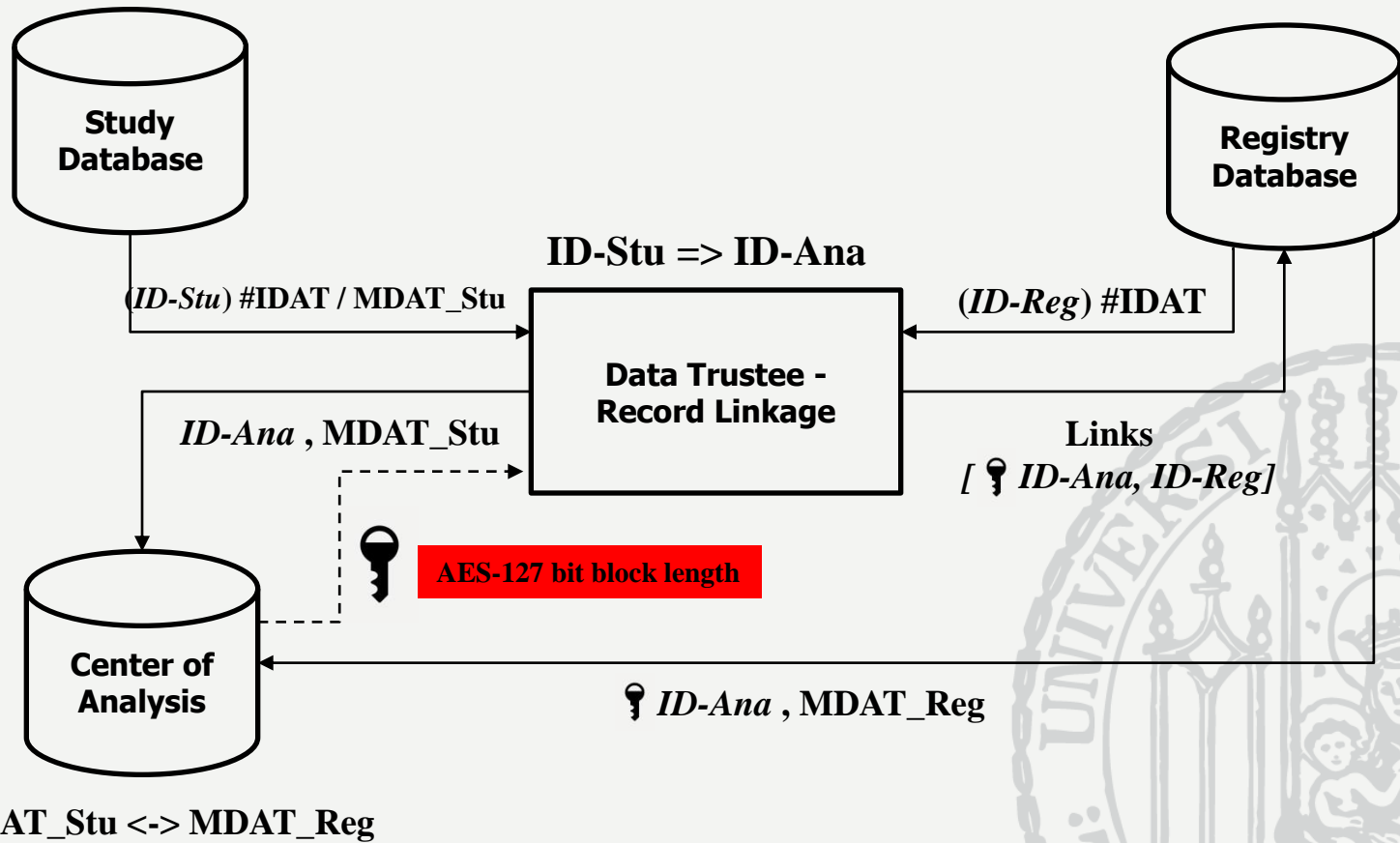
IV. Analysis-ID may only be shared between the Data Trustee and the Center of Analysis



MDAT_Stu <-> MDAT_Reg



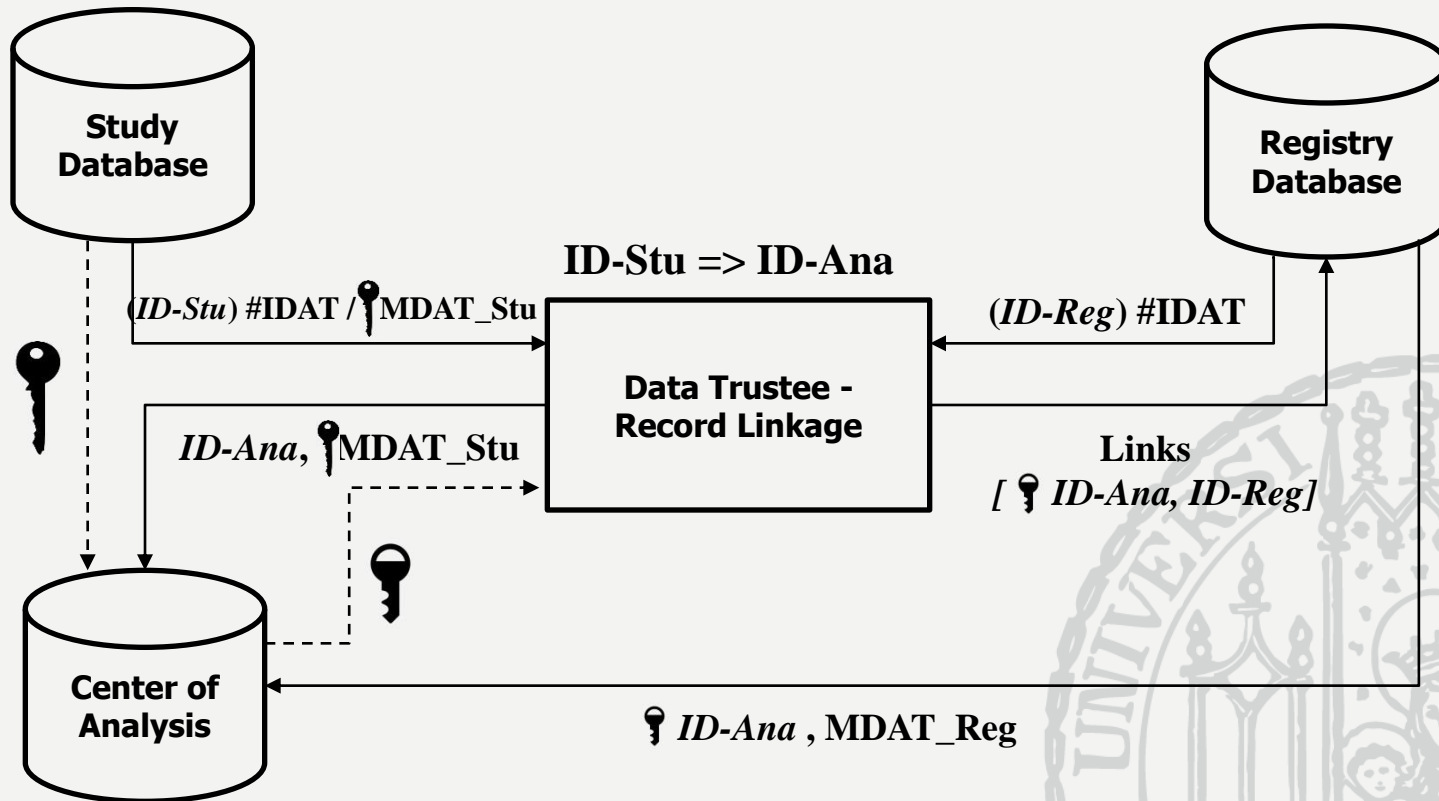
IV. Analysis-ID may only be shared between the Data Trustee and the Center of Analysis



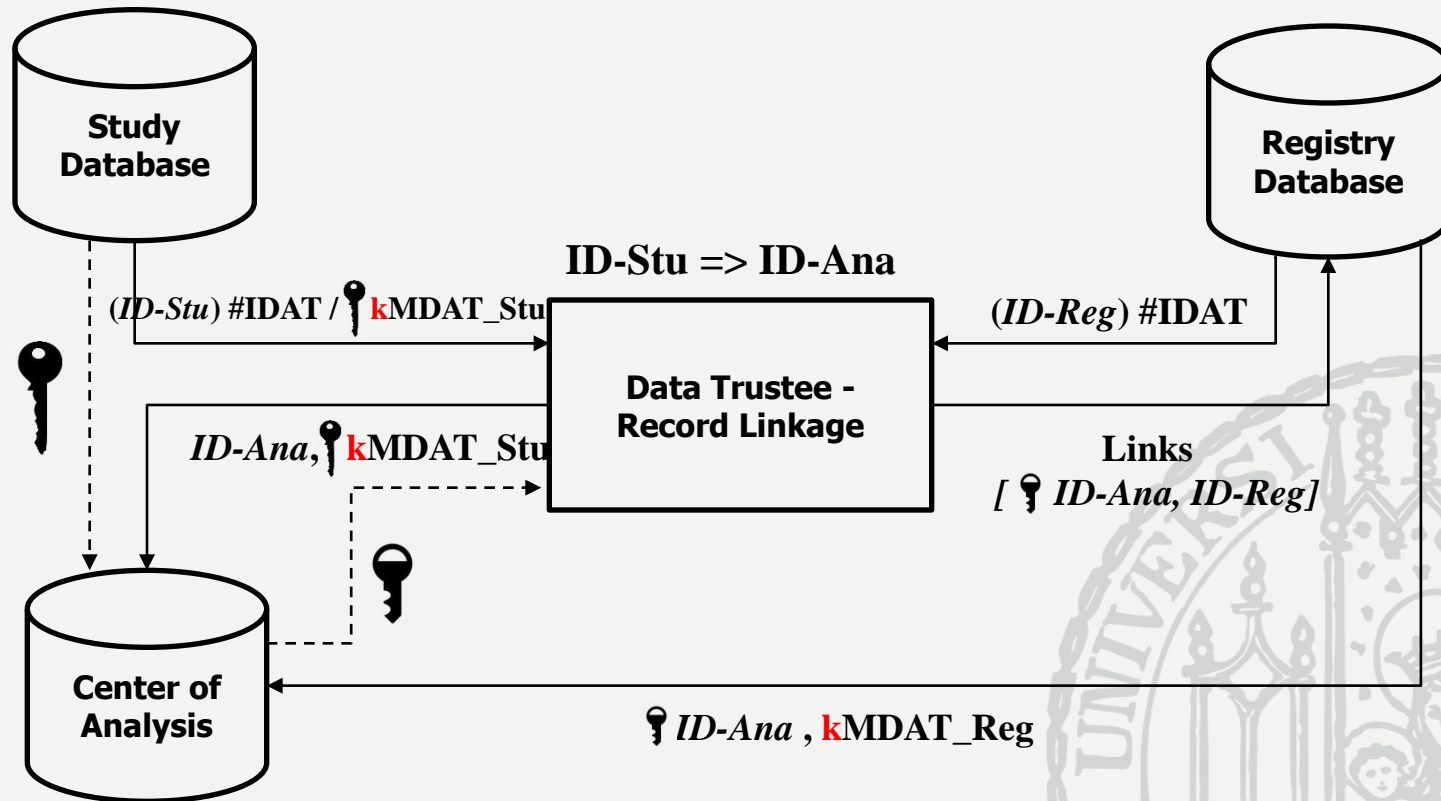
[1] BSI - Technische Richtlinie, Kryptographische Verfahren: Empfehlungen und Schlüssellängen, BSI TR-02102. 2013;

[2] Daemen J, Rijmen V. AES Proposal: Rijndael. 1999;

V. MDAT should only be readable for designated institutions



MDAT_Stu <-> MDAT_Reg



$\text{kMDAT_Stu} \leftrightarrow \text{kMDAT_Reg}$

VII. Additional Layer of assymetrical transport encryption

