

# DZHK Central Data Management Process Description and Data Protection Concept

---

DEUTSCHES ZENTRUM FÜR HERZ-KREISLAUF-FORSCHUNG E.V.  
(GERMAN CENTRE FOR CARDIOVASCULAR RESEARCH)

Version 1.2, 24 March 2014

Published by:

University Medicine Greifswald  
Institute for Community Medicine

represented by the Managing Director  
Prof. Dr. med. Wolfgang Hoffmann, MPH

Ellernholzstr. 1-2  
D-17487 Greifswald

Tel: +49 (0) 3834 86 7750  
Fax: +49 (0) 3834 86 7752  
email: wolfgang.hoffmann@uni-greifswald.de

University Medical Centre Göttingen  
Department of Medical Informatics

represented by the Director  
Prof. Dr. med. Otto Rienhoff

Robert-Koch-Straße 40  
D-37075 Göttingen

Tel: +49 (0) 551 39 3431  
Fax: +49 (0) 551 39 22493  
email: otto.rienhoff@med.uni-goettingen.de

## Content

---

<b>A</b>	<b>PROJECT-SPECIFIC SECTION.....</b>	<b>5</b>
<b>A1</b>	<b><i>CDM project overview.....</i></b>	<b>6</b>
<b>1</b>	<b>Project structures and partners.....</b>	<b>6</b>
1.1	DZHK.....	6
1.2	Institute for Community Medicine .....	6
1.3	Department of Medical Informatics, University Medical Centre Göttingen .....	7
1.4	Involved institutions.....	9
1.5	Tasks.....	9
<b>A2</b>	<b><i>Specifications for the TTP subproject (Greifswald) .....</i></b>	<b>12</b>
<b>1</b>	<b>Workflows and data flows .....</b>	<b>12</b>
1.1	Requirements and technical approach .....	12
1.2	Data flows within the CDM.....	13
1.3	Workflows .....	15
<b>2</b>	<b>Further protection requirements .....</b>	<b>21</b>
<b>3</b>	<b>Supplementary technical and organisational measures.....</b>	<b>21</b>
3.1	Integrating the secuTrial eCRF system.....	21
3.2	Separating identifiable information .....	21
3.3	Dealing with informed consent forms.....	22
3.4	Personnel measures.....	23
<b>A3</b>	<b><i>Specifications for the Data Handling subproject (Göttingen) .....</i></b>	<b>24</b>
<b>1</b>	<b>Workflows and data flows .....</b>	<b>24</b>
1.1	Data collection .....	25
1.2	Data capture.....	25
1.3	Data storage and management.....	25

1.4	Transfer Office: Data preparation and transfer.....	26
1.5	Involved groups of persons.....	29
2	<b>Contractual basis for cooperation between the TTP and the DH unit .....</b>	<b>29</b>
<b>B</b>	<b>PROJECT-INDEPENDENT SECTION.....</b>	<b>30</b>
<b>B1</b>	<b><i>Trusted Third Party .....</i></b>	<b>31</b>
1	<b>Independent Trusted Third Party (TTP) processes .....</b>	<b>31</b>
1.1	Unique identification .....	31
1.2	Pseudonymisation .....	32
1.3	Consent, authorisation and revocation .....	32
1.4	Working with secondary data .....	34
1.5	Participating in the Use & Access process.....	34
2	<b>Technical systems.....</b>	<b>34</b>
2.1	Master person index (MPI).....	34
2.2	Pseudonymisation service .....	35
2.3	Consent Manager.....	35
2.4	Architecture .....	36
3	<b>Protection requirements.....</b>	<b>37</b>
3.1	Legal basis.....	37
3.2	Storing personal data.....	39
3.3	Determining protection requirements .....	41
4	<b>Technical and organisational measures .....</b>	<b>42</b>
4.1	ICM-VC institutional data protection concept .....	42
4.2	Network protection.....	43
4.3	Audit trail.....	44
4.4	Data transfer .....	44
4.5	Data security .....	44
4.6	Failure protection.....	45
4.7	Spacial separation.....	46
4.8	Personnel measures.....	46
<b>B2</b>	<b><i>Data Handling unit .....</i></b>	<b>47</b>
5	<b>Data Handling unit processes.....</b>	<b>47</b>

<b>6</b>	<b>Technical systems</b> .....	<b>48</b>
6.1	secuTrial .....	48
<b>7</b>	<b>Protection requirements</b> .....	<b>50</b>
7.1	Processed data types .....	50
7.2	Applicable legal basis .....	51
<b>8</b>	<b>Technical and organisational measures</b> .....	<b>52</b>
8.1	IT infrastructure used.....	52
8.2	Server virtualisation .....	53
8.3	Process description in accordance with NDSG § 8 .....	53
<b>C</b>	<b>ANNEX</b> .....	<b>54</b>
<b>1</b>	<b>Technical illustrations</b> .....	<b>55</b>
<b>2</b>	<b>List of abbreviations</b> .....	<b>59</b>
<b>3</b>	<b>Glossary</b> .....	<b>59</b>
<b>4</b>	<b>Bibliography</b> .....	<b>60</b>
<b>5</b>	<b>Attachments</b> .....	<b>62</b>

## A Project-specific section

---

# A1 CDM project overview

## 1 Project structures and partners

---

The following section introduces the framework of the project in which the concept of the Central Data Management (CDM) unit is to be implemented within the German Centre for Cardiovascular Research (DZHK). The CDM unit consists of the independent Trusted Third Party (TTP) and the Data Handling (DH) unit. It will also explain the project-specific requirements and the resulting measures.

### 1.1 DZHK

The DZHK is one of the six German Centres for Health Research (DZGs). Their foundation was initiated by the German Federal Ministry of Education and Research (BMBF) and took place between 2009 and 2012. The health centres are organised as "eingetragene Vereine" ("e. V."), which are registered associations under German law. They are registered as part of the Helmholtz- Gemeinschaft Deutscher Forschungszentren e.V. (Helmholtz Association of German Research Centres), making them research networks within Germany's largest non-university scientific organisation. The intention is that the health centres work in close cooperation with one another in order to achieve their goals faster.

The aim of the DZHK is to improve the prevention, diagnosis and treatment of heart and cardiovascular diseases, and to ensure by way of wide-ranging cooperation that research results from this area of expertise find their way more quickly into clinical practice.

To achieve this, 27 partner institutions are working at the following seven locations / in the following location groups within the DZHK: Berlin, Göttingen, Greifswald, Hamburg-Kiel-Lübeck, Heidelberg-Mannheim, Munich and Rhine-Main (Frankfurt a.M., Bad Nauheim, Mainz).

These partner institutions are universities, university hospitals and non-university research institutions (several Max Planck Institutes, the Max Delbrück Centre and one Leibniz Centre). The plan is to cooperate with strong partners from the German healthcare sector.

The work of all the health centres is to be overseen and assessed by high-level advisory bodies staffed by international consultants. The aim is to evaluate not only scientific excellence, but also the strategic direction and the established structures and processes.

As with the other DZGs, the DZHK is 90 % financed by the German federal government. The remaining 10 % of the financing come from the federal states where the respective health centre maintains sites (in the case of the DZHK this includes nine federal states). [1]

### 1.2 Institute for Community Medicine

The focus of the Epidemiology of Health Care and Community Health Section of the Institute for Community Medicine (ICM) at University Medicine Greifswald is analytical epidemiology and risk factor research, epidemiology of and research into health care, health systems and transfer research, disease prevention and the promotion of health, and the development and practical implementation of new models in all areas of medicinal care.

In order to complete the processing of large amounts of data (particularly personal data) that is required, competencies in the field of medical informatics and data management have been significantly expanded in recent years. The Institute for Community Medicine has comprehensive experience in the integration of complex clinical and epidemiological data from multicentre studies and decentralised field surveys in central management units based on central database systems.

### 1.3 Department of Medical Informatics, University Medical Centre Göttingen

The Department of Medical Informatics was founded in 1972 and since 1999 has been dedicated to collaborative medical research. It develops new methodical processes for interdisciplinary multicentre collaboration in cooperation with the researchers from Germany and abroad. The institute has provided and continues to provide methodological and technical support for many individual studies and the competence networks for dementia, congenital heart defects and multiple sclerosis, as well as for several clinical research areas and the German Collaborative Research Centre 1002 (Sonderforschungsbereich 1002 [SFB 1002]). Furthermore, the Department of Medical Informatics has contributed to the establishment of the research body known as Technology, Methods, and Infrastructure for Networked Medical Research (Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V. [TMF]). Additionally, the institute was Germany's main representative for grid-based research projects in medicine.

Currently almost 40 scientists, documentation officers and administrative staff are working in close cooperation with other institutions on the aforementioned long-term projects and on methodological support for the German Centres for Health Research (DZGs). While the competence networks in medicine represent the decade of methodological breakthrough in collaborative medical research, the infrastructure for the national centres is being established on a new level in coordination with the TMF. This new level bundles the experiences of the last ten years and combines them with the data and requirements of personalised medicine. Since around 2014, the participating associations have been able to rely on the use of a powerful research infrastructure that stands up to international comparison.

### Central Data Management joint project

With its communication dated 26 March 2013, the DZHK Board of Directors commissioned University Medicine Greifswald (represented by the ICM) and the University Medical Centre Göttingen (represented by the Department of Medical Informatics [MI]) to implement the Central Data Management (CDM) unit as a joint project for multicentre studies and registers within the DZHK. This federal approach presents the opportunity to realise spacial and organisational separation of medical and personally identifiable information in accordance with a uniform standard. Medical research data is usually captured as part of multicentre studies. The following studies and registers have been selected for funding as part of an external assessment process, the requirements of which have been included insofar as they were known at the time of the creation of this document:

1. Early Versus Late Left Ventricular Assist Device Implantation (VAD)
2. Systolic Dysfunction to Congestive Heart Failure Cohort Study (TransitionCHF)
3. Translational Registry for Cardiomyopathies (TORCH)

The Central Data Management unit is a joint project between partners that are equal and operate under their own authority. It consists of the Trusted Third Party (TTP) (under the responsibility of the Institute for Community Medicine at University Medicine Greifswald [Prof. Dr. med. Wolfgang Hoffmann]), the Data Handling (DH) unit (under the responsibility of the Department of Medical Informatics at the University Medical Centre Göttingen [Prof. Dr. med. Otto Rienhoff]) and the IT Lab situated at the DZHK Main Office (Prof. Matthias Nauck).

The responsibility for ethical aspects and their harmonisation in the DZHK is realised by the Institute of Epidemiology II at the Helmholtz Zentrum (Prof. H.-E. Wichmann) and the Department of Medical Statistics and Epidemiology (Prof. Frank Kuhn) at the Technical University of Munich as part of an additional ethics project by DZHK Programme Group 7. Furthermore, every study and register independently creates a specific ethics concept and an informed consent (IC) form, which they submit for assessment by the responsible ethics committees and then send to the persons responsible for ethics at the CDM unit after a positive ethics vote. The task of the individual study centres is to recruit participants and capture data afterwards. An IT Office (comprising an IT Lab and the DZHK Main Office's IT Management Office) is to be established at the DZHK Main Office, which is based in Berlin. Among others, this office will assume the function of the DZHK Commissioner for Data Protection and responsibility for coordination with the relevant federal state authorities.

The relationships mentioned are illustrated in the organisation chart below (see Figure 1).

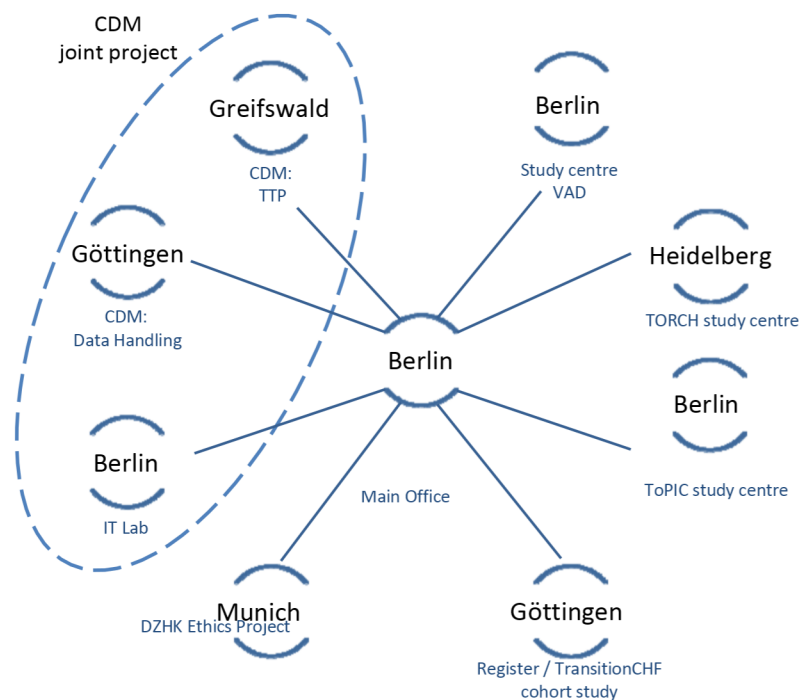


Figure 1: DZHK Central Data Management joint project partners



## 1.4 Involved institutions

The Ethics Project is responsible for preparing the required informed assent and consent forms in terms of content. The CDM joint project uses only the informed consent forms that have been agreed upon and harmonised in advance.

The services provided by, responsibilities of and relationship between the partners involved are regulated by separate cooperation agreements. These are currently still being coordinated between the partners and will ensure that all partners can operate equally, independently and under their own authority. Plans include both an internal cooperation agreement for the CDM joint project and a cooperation agreement between the CDM unit and the Ethics Project. Furthermore, cooperation agreements between the individual projects (registers/studies), the CDM unit and the Ethics Project are in the process of coordination.

## 1.5 Tasks

The technical and organisational measures necessary to successfully implement the Central Data Management unit can essentially be reduced to the three following aspects. These are illustrated additionally in Figure 2.

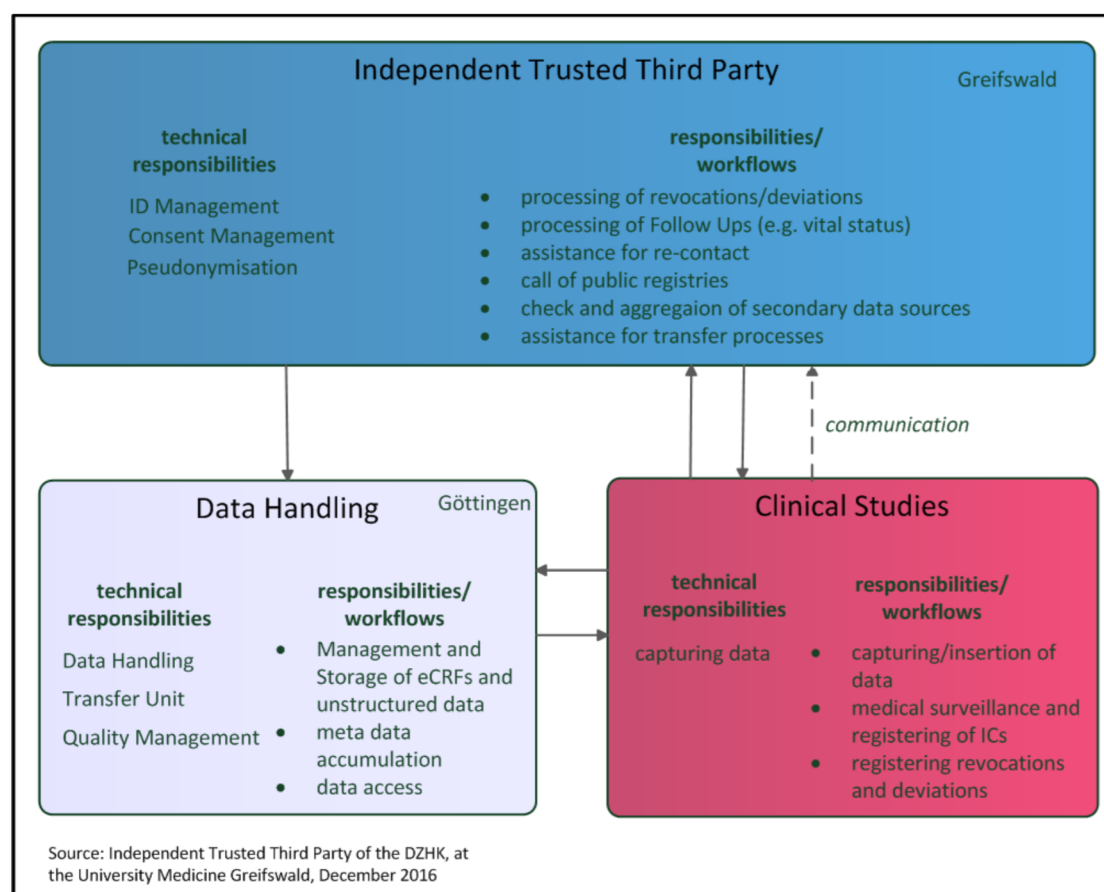


Figure 2: Responsibilities within the Central Data Management unit

### *Trusted Third Party (TTP)*

The Trusted Third Party assumes responsibility for the management and storage of the personally identifiable information (IDAT) as part of the transfer of functions. The transfer of data to the TTP is coordinated by the participating studies, registers and cohorts (SRC) with the respective German federal state commissioner for data protection. Data processing within the TTP must be based on informed consent without exception.

Three essential technical functionalities are necessary for the management of the personally identifiable information: a master person index (MPI), consent management (CM) system and a generalised pseudonymisation service (PSN). The trusted third party receives the personally identifiable information and creates a suitable number of pseudonyms, which are then used and stored with the medical data at the Data Handling unit (see Section B1).

### *Data Handling (DH) unit*

The data handling for all SRCs conducted within the DZHK takes place in Göttingen. This means that a study database is established and operated, and that the data elements for the individual SRCs are modelled there. To achieve this, the personnel in Göttingen are in close coordination with the chief investigators. This coordination is partly advisory (i.e. assistance is provided for uniform, sustainable selection of items for documentation) and partly practical (implementation). Trained documentation officers and documentation officers with an additional master's degree in Medical Informatics are available for these advisory and implementation services. Furthermore, a metadata directory is to be built in the DH in order to fulfil two additional tasks for the long-term comparability of DZHK studies: The first task is to allow a specific item (e.g. diastolic blood pressure) to be captured equally (e.g. sitting after five minutes' rest), stored (unadjusted two- to three-digit integer values) and provided with the same name within the database for every study conducted (e.g. bld\_prsr\_dia). This allows long-term comparability of the DZHK studies conducted. The second task is to allow specific information to be stored within a metadata directory, e.g. information on the circumstances (environmental metadata) of an examination. The same will apply to other data types later on, such as images or information on collected biomaterials.

Other (sometimes complex) data types will become relevant for the DZHK and CDM unit both with the three approved DZHK studies and with new projects in the coming years. This could include electrophysiological data, image data (in coordination with the corresponding DZHK working groups), follow-up data generated from a mobile source (possibly by the patients themselves) or the aforementioned examination metadata. During the course of the project applied for and together with the IT Management Office at the Main Office and TTP employees, DH unit employees create concepts to connect additional databases such as the central DZHK biobank or a DZHK image database and then realise these concepts. These databases can be operated in Göttingen, at other DZHK sites, or with external service providers. However, the CDM unit will ensure uniform architecture for the DZHK in each case. This will be achieved by strategy coordination, implementation planning (and, if necessary, realisation) as well as by cooperation on the specification of processes for establishment and operation.

## *Data exchange and interfaces*

Some of the applications require data exchange between the TTP and the DH unit, or between the TTP and the studies. This can take place either automatically or manually using either electronic means or paper documents. Common interfaces and processes that meet the requirements of data protection law both with regard to electronic exchange and (partially) manual workflows are defined not only by the spacial separation of the data, but also by the different technological systems. The following chapters explain assumptions, conditions and applications for the TTP during interaction with users (DH unit or studies), identify interfaces and define these in greater detail.

## A2 Specifications for the TTP subproject (Greifswald)

### 1 Workflows and data flows

---

The following figures describe the status of the planning and consultation at the time the document was created.<sup>1</sup>

#### 1.1 Requirements and technical approach

In addition to the aspects specific to the TTP that are described here, detailed coordination of all processes with regard to a common interface is necessary for both the DZHK and the Data Handling subproject. This coordination is to be achieved with a focus on technical characteristics and processes (workflows and use cases). It is also necessary to define and describe the internal processes of the TTP.

Cooperation between the TTP and the University Medical Centre Göttingen's Data Handling subproject means that the following conditions are necessary for the implementation of the TTP:

Pseudonymised medical data must be captured in the study centres using electronic case report forms (eCRFs). A basic dataset must be used for all studies, registers and cohorts. Each study must also have a study-specific dataset. IDAT must be captured in order to create the required pseudonyms. However, this is not to be stored in the eCRF.

In the medium- to long-term, unstructured data (e.g. image data, ECG data or free text questionnaire data) will be captured. This (MDAT) data may also be processed by the DH unit in Göttingen or by other institutions. These institutions must provide the same required technical infrastructure. The required transfer office (for the transfer of the collected data for research purposes) should be made available by the DH unit. This data protection concept must be expanded if further data types and/or institutions are added.

The eCRFs are developed with the help of the web-based tool secuTrial, which was developed by iAS Berlin GmbH. The sole contract partner of iAS GmbH is the University Medical Centre Göttingen. The cooperation agreement of the CDM joint project stipulates that the TTP has a right to be involved in the definition and testing of the interface. The relevant studies are responsible for the form and content of the electronic report forms.

---

<sup>1</sup> The purpose of this status is to begin data capture. The plan is to update the content of the versions (e.g. by adding additional processes) as soon as the defining aspects have been coordinated between the partners involved. The proposed extensions are support during the transfer of data and materials (transfer office process) and the integration of modalities other than eCRFs (e.g. the biobank and imaging procedures).

secuTrial has a central roles and rights system. As required, it is therefore only possible to allow a selection of persons or groups of persons (doctors, studies) to amend identifiable information. Defined, contractually agreed interfaces allow for the necessary interaction between the Data Handling unit, Trusted Third Party and secuTrial.

There are also plans to manage biosamples using a laboratory information and management system (LIMS). The current plan is for the DZHK Main Office's IT Management Office to prepare an invitation to tender, the specifics of which are currently being developed.

Personal data, consents and authorisations, as well as all the pseudonyms and allocations created at a single point in time, are saved at the TTP. The TTP manages both this data and other metadata on studies, projects and locations. The data is encrypted and stored in accordance with the TTP's data protection concept and the relevant security regulations with regard to access and backup (see Section B1, Chapter 4).

## 1.2 Data flows within the CDM

Figure 3 (below) illustrates the planned data flow between the individual subprojects.<sup>2</sup> The technical details for the individual processes can be found in the annex (see Figures 16-19).

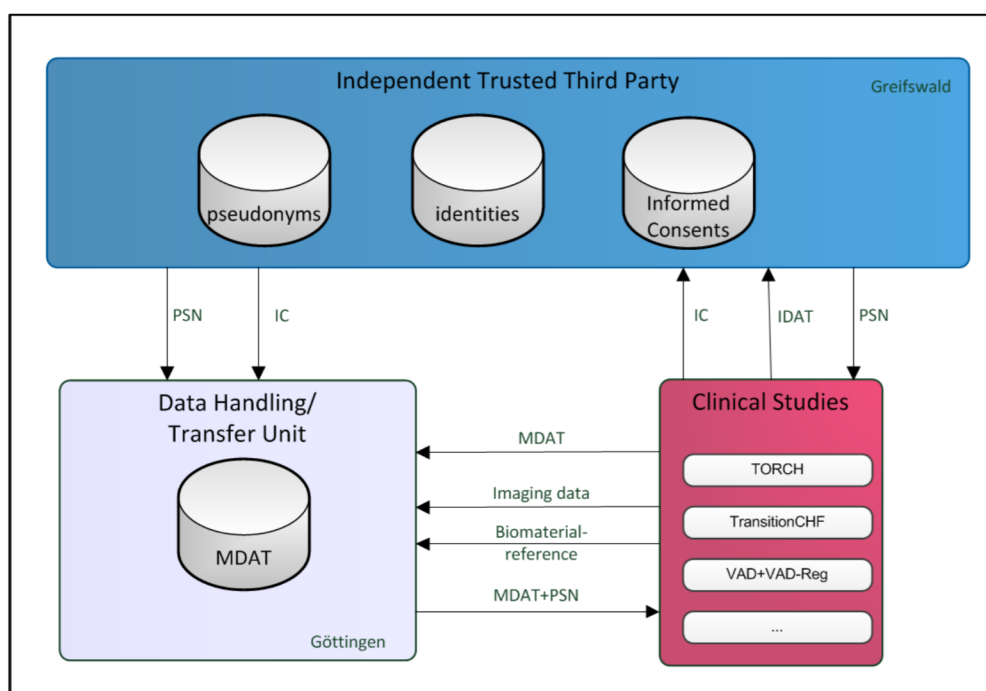


Figure 3: Data flows within the Central Data Management unit

Upon inclusion of a person, the first step is to send the informed consent information and identifiable information to the TTP. The unique PSN generated for this person by the TTP is then transferred to the study centre and (together with the informed consent information) to the Data Handling unit.

<sup>2</sup> No attempt will be made to illustrate the authentication mechanisms and provision of the eCRF using secuTrial at this point.