

*Universitair Informatiemanagement*

**Dataopslagbeleid Universiteit Twente**

Inhoudsopgave

Voorwoord .....	2
Inleiding.....	3
1. Afbakening.....	3
2. Algemeen beleid .....	4
3. Instellingssystemen.....	5
4. E-mail en gebruikersdata.....	5
Bijlagen .....	8
A. Analyse .....	8
B. Instellingssystemen.....	11
C. E-mail .....	13
D. Gebruikersdata.....	14
E. Onderzoeksdata.....	17

## **Voorwoord**

Het onderhavige document legt de basis voor een expliciet opslagbeleid van de UT. Het behandelt beleidskeuzes uit mogelijke functionaliteiten van opslag, beveiligingsniveaus en wijzen van bekostiging.

De consultatieversie van het document is tot stand gebracht in een intensieve samenwerking tussen het Universitair Informatie Management en het ICT Servicecentrum. Daarin is de balans gezocht tussen de technische mogelijkheden van aanbod voor opslag en de vraag naar opslag in het licht van noodzaak voor de bedrijfsprocessen en kostenbesparing. De keuzes passen in de ICTS-trajecten naar uitrol UT werkplek en virtualisatie van server- en opslagcapaciteit voor werkplekken en instellingssystemen.

De beleidskeuzes worden ter advisering aan de UCB voorgelegd, alvorens ze ter bekrachtiging het CvB aan te bieden.

## Inleiding

Dit document geeft in hoofdstuk 1 een afbakening van waarop het behandelde opslagbeleid betrekking heeft. De hoofdstukken 2, 3 en 4 definiëren het opslagbeleid en vormen de kern van het document. Ze behandelen achtereenvolgens de algemene keuzes, de specifieke keuzes voor instellingssystemen en die voor e-mail en gebruikersdata. Aan het document is een bijlage toegevoegd voor een verantwoording van en toelichting op de beleidskeuzes.

De hoeveelheid opgeslagen data groeit continu, de afhankelijkheid van de organisatie van digitale gegevens wordt steeds groter. Dit geldt niet alleen voor de ondersteunende processen, maar meer en meer ook voor de primaire processen onderwijs en onderzoek. De hoeveelheid digitale gegevens groeit wereldwijd sneller dan de opslagcapaciteit.<sup>1</sup>

Mede naar aanleiding van de evaluatie van de calamiteit eind 2008, waarbij gedurende langere tijd schijven voor dataopslag van een deel van de medewerkers niet benaderbaar waren, heeft het CvB de universitaire informatiemanager gevraagd om vanuit het perspectief van vraagbeheersing een expliciet universitair opslagbeleid op te stellen.

Doordat de kosten van opslag per volume-eenheid sterk dalen, kan de kostengroei beperkt blijven. Gezien het belang van dataopslag en de omvang van de kosten is het zinvol bewust stil te staan bij de keuzes die ten aanzien van dataopslag worden gemaakt.

## 1. Afbakening

De scope van het dataopslagbeleid zijn alle aspecten betreffende de opslag van digitale gegevens door medewerkers en studenten en van alle gegevens welke door informatiesystemen van de UT gebruikt worden. Het dataopslagbeleid spreekt zich daarmee niet uit over opslag van papieren dossiers, microfilms, analoge (lucht)foto's en films, maar alleen over digitale data.

Het UT dataopslagbeleid heeft zodoende betrekking op de volgende typen data:

- data opgeslagen in databases van instellingssystemen
- e-mail van medewerkers
- gebruikersdata van medewerkers, zoals
  - o office documenten
  - o digitale foto's, geluidsopnamen en video's
  - o digitale luchtfoto's en satellietbeelden
  - o experimentele meetgegevens
  - o overig onderwijsmateriaal
  - o andere wetenschappelijke data
- e-mail van studenten, vooraanmelders en alumni
- gebruikersdata van studenten, vooraanmelders en alumni

Het dataopslagbeleid heeft raakvlakken met het archiefbeleid<sup>2</sup> en het informatiebeveiligingsbeleid<sup>3</sup>. Voor consequenties van Archiefwet, Auteurswet en Wet Bescherming Persoonsgegevens, wordt dan ook verwezen naar het betreffende vastgestelde UT-beleid.

Aan digitale opslag van data zijn meerdere aspecten verbonden, zoals:

- beschikbaarheid, het op kunnen roepen van data op ieder gewenst moment en plaats
- opslagflexibiliteit, het kunnen opslaan van data op ieder gewenst moment en plaats
- verliesvermijding, het voorkomen dat data definitief verloren gaat
- samenwerking, het mogelijk maken dat data gedeeld kan worden ten behoeve van eenvoudige opslag
- doorzoekbaarheid, voorzieningen voor het kunnen vinden van opgeslagen data
- kostenbeheersing, waaronder aanschaf-, beheers- en energiekosten

<sup>1</sup> Zie ook het IDC White paper "The Expanding Digital Universe", maart 2007

<sup>2</sup> Het archiefbeleid is beschreven in de 'eBusiness Case Archivering UT' (Kenmerk 380.424 van 25-09-2007).

<sup>3</sup> Zie <http://www.utwente.nl/sectr/im/security> Informatiebeveiligingsbeleid Universiteit Twente 2010

- beveiliging, zoals beschreven in het informatiebeveiligingsbeleid
- risicoacceptatie, het vaststellen van maximaal aanvaardbare risico's

## 2. Algemeen beleid

Voor alle dataopslag geldt een afgewogen beleid. Periodiek zal op het juiste niveau een afweging gemaakt worden tussen aspecten als capaciteit, performance, kosten, noodzaak van directe beschikbaarheid, risico van uitval, risico dat data verloren gaat en risico van verlies van vertrouwelijkheid. Vanuit de vraag worden eisen aan de opslag qua capaciteit, performance, beschikbaarheid en maximaal aanvaardbare risico's gesteld.

ICTS levert als product dataopslag. De klanten nemen deze af, de gebruikers maken er gebruik van. De voorwaarden zijn uitgewerkt in een dienstverleningsovereenkomst.

De algemene beleidskeuzes met betrekking tot de opslag van data zijn als volgt.

- ✗ Voor de opslag van data zijn een beperkt aantal opslagregimes beschikbaar op basis van keuzes van aanvaarde versterking van de toegang tot de data of verlies van de data. Zie de hierna volgende opsomming.
- ✗ De opslagvoorziening wordt doorberekend. Dit gebeurt op basis van het gereserveerde gebruik (dus niet van het daadwerkelijke gebruik).
- ✗ Jaarlijks worden de tarieven van doorberekening voor de opslagregimes vastgelegd in de generieke dienstverleningsovereenkomst tussen ICTS en UT.
- ✗ De kosten en tarieven dienen marktconform en transparant te zijn.
- ✗ ICTS registreert het gereserveerde en daadwerkelijke gebruik en rapporteert hierover periodiek, waarbij een onderverdeling wordt gemaakt per toepassing, opslagregime, beheerseenheid, subeenheid (vakgroep, leerstoel of afdeling) en gebruiker.
- ✗ ICTS draagt zorg voor voorlichting aan gebruikers over dataopslag, zodat zij in het gebruik de juiste afweging kunnen maken tussen kosten en risicobeheersing.

### Opslagregimes

De volgende opslagregimes worden onderscheiden:

1. **Lokaal.** Opslag geschiedt op de harde schijf in een computer, maar ook losse schijven, USB-sticks etc. zijn mogelijk. Deze opslag wordt niet centraal beheerd en kent een hoog risico bij alleen deze enkelvoudige opslag. Lokale opslag is noodzakelijk voor mobiel gebruik wanneer geen verbinding met het internet beschikbaar is. Het is ook nuttig voor vaste werkplekken, omdat bij netwerkstoring doorgewerkt kan worden. Bij een lokale storing kan de data onherstelbaar verloren gaan. Aanschafkosten ca. € 0,06 – 0,50 / GB (in 2010).
2. **Online.** De toegang tot de data is via het netwerk mogelijk. De opslagvoorziening wordt beheerd door ICTS.
  - a. **IJzer.** De data wordt op het netwerk enkelvoudig opgeslagen. Bij een storing kan deze data onherstelbaar verloren gaan. Bruikbaar als secundaire backup-opslag voor lokaal opgeslagen data, welke daardoor met het kopie tegen verlies is beveiligd; of de lokale kopie of de netwerk kopie is beschikbaar. Ook bruikbaar voor data welke geheel opnieuw automatisch te genereren is.
  - b. **Brons.** Opslag is van kwaliteit 'ijzer' met daarnaast iedere nacht een backup. Hierdoor kan normaal gesproken maximaal de netwerkopslag van één dag verloren gaan. Elk gebackupid bestand kent op het netwerk twee kopieën: één (versie) op de opslagvoorziening en één (versie) in de backup. Belangrijke backup's worden op een andere locatie bewaard tegen grotere calamiteiten.
  - c. **Zilver.** 'Brons' plus replicatie (continue tweevoudige opslag op de hardware) ter bescherming tegen hardware storingen. Elk bestand kent drie (versies van) kopieën: twee op de opslagvoorziening, één in de backup. Dit verkleint de kans op dataverlies nog verder.
  - d. **Goud.** Zilver plus snapshots, een kopie techniek die er voor zorgt dat maximaal het werk tot de laatste snapshot verloren kan gaan en bovendien beveiligd is tegen onbedoelde wijzigingen door de gebruiker. Gekozen wordt voor snapshots om de

vier uren. Elk bestand op het netwerk heeft gemiddeld tussen de vier en vijf (versies van) kopieën.

3. **Offline.** Opslag op tape, beheerd door ICTS. Directe toegang tot de data niet mogelijk. Na aanvraag wordt de data online gezet.

Om de verschillende online-opslagregimes te kunnen vergelijken worden de kentallen MTU en MTD gehanteerd. Maximaal Toelaatbare Uitvalduur (MTU) is een maat voor de acceptatie van het risico van verlies van toegang tot de data. Maximaal Toelaatbaar Dataverlies (MTD) is een maat voor de acceptatie van het risico van het verlies van data zelf. Voor de UT worden de volgende keuzes gemaakt. Opgemerkt wordt dat het hier de MTU en MTD van de opslagvoorziening betreft. Voor de gebruiker zullen uiteindelijk ook de beschikbaarheidsrisico's voor netwerk, applicatieservers en de eigen werkplek relevant zijn.

Opslagregime	MTU <sup>4</sup>	MTD	Relatief risico	Kosten (2010)
IJzer	eeuwig	alles	hoog	ca. € 1,- / GB / jr
Brons	twee weken	één dag	normaal	ca. € 2,- / GB / jr
Zilver	twee dagen	één dag	laag	ca. € 3,- / GB / jr
Goud	acht uur	vier uur	zeer laag	ca. € 5,- / GB / jr

### 3. Instellingssystemen

De opgeslagen data welke gebruikt wordt door de instellingssystemen is via het netwerk beschikbaar. Gegevens uit het verleden kunnen nog enige tijd direct online vanuit het systeem beschikbaar en nodig zijn. Door deze zo veel mogelijk en waar mogelijk Offline te archiveren wordt de opslagvoorziening ontzien.

Afhankelijk van hoe de risico's worden ingeschat zal per systeem door de systeemhouder een keuze voor een opslagregime (Brons, Zilver, Goud) gemaakt dienen te worden. Bij onderlinge afhankelijkheid voor opslag van systemen wordt een gezamenlijke keuze gemaakt.

- Bij instellingssystemen is de houder verantwoordelijk voor de keuze van het juiste opslagregime en voor de verantwoording van de kosten van de betreffende dataopslag. Eventuele doorberekening van de storage aan de systeemhouder (i.p.v. bekostiging via CO) wordt meegenomen in het overleg rond budgetsturing.
- De houder maakt in samenspraak met ICTS en UIM de keuze van het opslagregime.
- Het gekozen opslagregime wordt per systeem vastgelegd in de specifieke dienstverleningsovereenkomst tussen ICTS en de houder.

### 4. E-mail en gebruikersdata

#### *Uitgangspunten*

- Zakelijke data, via de werkplek beschikbaar en relevant voor werk en studie, wordt in ieder geval op de lokale harde schijf of andere opslagmedia van de werkplek opgeslagen. Indien het voor het zakelijke gebruik noodzakelijk is wordt de data ook op opslagvoorzieningen online via het netwerk opgeslagen. ICTS voorziet in de faciliteit waarmee de data op de lokale werkplek bij verbinding met het netwerk steeds wordt gesynchroniseerd met de data op de online opslagvoorziening.
- De kosten van het gereserveerde gebruik door medewerkers van de totale online dataopslag voor e-mail en gebruikersdata worden doorberekend aan de beheerseenheden.
- Voor accounts van niet-UT-ers gelden in principe dezelfde regels als voor medewerkers, de eenheid welke het account aanvraagt draagt de kosten.
- Het is de werknemer toegestaan het ICT- en Internetgebruik in beperkte mate in te zetten voor privédoeleinden. Dit dient echter geen significante meerkosten voor de UT te veroorzaken. Privé data wordt daarom alleen lokaal en niet online op het netwerk

<sup>4</sup> De waarden in de kolom MTU dienen nader bezien te worden, of deze keuzes daadwerkelijk de meest juiste en haalbare zijn

opgeslagen. ICTS richt de werkplekken zo in dat dit voor medewerkers gemakkelijk te realiseren is.

### **Inrichting**

- ✘ Het wordt door ICTS de medewerkers gemakkelijk gemaakt om voor de werkplek en het ingestelde opslagregime de reservering van omvang dataopslag voor e-mail en gebruikersdata aan te passen, bijvoorbeeld via een webapplicatie.
- ✘ ICTS richt de werkplekken zo in dat belangrijke zakelijke data zowel lokaal als online op het netwerk beschikbaar zijn. Afhankelijk van het ingestelde opslagregime is in de beveiliging van deze data voorzien.
- ✘ Het terughalen van een beveiligde kopie van data wordt gemakkelijk gemaakt, bijvoorbeeld met een self-service.
- ✘ Voor het opslagregime voor e-mail kiest de UT voor één regime voor alle medewerkers, namelijk Zilver. Voor de e-mail voorziening stapt ICTS met de invoering van de UT werkplek ICTS over op Exchange 2010. Het eigen opslagmechanisme van Exchange is een gegeven en lijkt het meest op Zilver. Naast opslag van de mailberichten op de lokale schijf van de werkplek bestaan er in de opslag- en backupvoorziening op het netwerk drie kopieën.
- ✘ Voor het opslagregime voor persoonlijke en groepsgebruikersdata kiest de UT in de generieke dienstverlening voor één regime voor alle medewerkers, namelijk Brons. De data staan daarmee opgeslagen op de lokale schijf van de werkplek, op de virtuele schijf op het netwerk en in de backup. Waar noodzakelijk door specifieke omstandigheden zal ICTS als maatwerk een ander regime aanbieden.
- ✘ Er wordt door ICTS een mogelijkheid gecreëerd om te archiveren (wetenschappelijke) data goedkoop offline op te slaan, zodat deze niet verloren gaat.
- ✘ Bij een ernstige calamiteit heeft het weer operationeel krijgen van de e-mail voorziening hoge prioriteit in verband met de communicatie met medewerkers.

### **Quota**

- ✘ De quota van gereserveerde dataopslag voor e-mail en gebruikersdata gezamenlijk dienen eenvoudig en individueel aan te passen te zijn. Initieel geldt een default omvang. Bij invoering van dit beleid wordt uitgegaan van het huidige gebruik. (Veel gebruikers hanteren hun mailbox als persoonlijk archief, waardoor deze veel groter is dan hun persoonlijke dataopslag. Vanuit kosten oogpunt is hier voortaan geen enkel bezwaar meer tegen.)
- ✘ De quota voor e-mail en gebruikersdata van studenten, alumni en vooraanmelders worden jaarlijks vastgelegd in de dienstverleningsovereenkomst. De kosten ervan worden centraal gedragen.
- ✘ Het quotum voor gezamenlijke data opslag voor groepen medewerkers wordt met de eenheden afgesproken, inclusief de wijze van doorberekening ervan.

### **Voorlichting**

- ✘ Het gebruik van Outlook archief bestanden (PST-files) op netwerkschijven wordt vanwege het excessieve opslaggebruik bij het backuppen door ICTS met voorlichting actief ontmoedigd. Het verdient aanbeveling de mail dan niet te archiveren, maar gewoon in de mailbox te bewaren. Bestaande PST-files kunnen in de mailbox geïmporteerd worden.
- ✘ De voorlichting door ICTS, over hoe bestaande mogelijkheden van gezamenlijke dataopslag optimaal zowel als efficiënt kunnen worden toegepast, wordt beschreven op de website en in de dienstverleningsovereenkomst.

### **Toekomst**

- ✘ UIM doet onderzoek naar de mogelijkheden van de invoering van een samenwerkingsomgeving (denk aan Surfgroepen of opvolger hiervan). Hierdoor kan onder meer meervoudige opslag van identieke bestanden bij samenwerking in groepen/commissies worden tegen gegaan.
- ✘ UIM en ICTS doen gezamenlijk onderzoek naar de business case van outsourcing van de e-mail en dataopslag-voorziening, vooralsnog van alleen studenten, alumni en

vooraanmelders. Hierbij worden noodzakelijke beveiligingstechnische en juridische aspecten meegenomen.

- ✂ UIM en B&A doen gezamenlijk onderzoek naar het wegnemen van belemmeringen die noodzakelijk langdurig veilig opslaan en efficiënt ontsluiten van (wetenschappelijke) datasets in de weg staan. Daarbij wordt aangesloten bij bestaande 3TU initiatieven.

## Bijlagen

### A. Analyse

Dataopslagbeleid dient gebaseerd te zijn op de analyse van de waarden van de opgeslagen data. Om ongecontroleerde oplopende kosten en ervaren risico's te vermijden dient er beleid voor groeibeheersing en risicobeheersing geformuleerd te worden. Dit leidt onder andere tot een heldere vraag waaraan ICTS met technische voorzieningen kan voldoen.

#### **Waarden**

Data wordt opgeslagen omdat de gebruiker of anderen deze later weer willen kunnen oproepen. Dataopslag heeft bepaalde waarden, die zich laten vertalen in:

- De noodzaak voor de goede werking van instellingssystemen
- Het op kunnen roepen van bepaalde documenten of gegevens
- Het kunnen werken met bepaalde databestanden
- Het veilig kunnen opslaan van vertrouwelijke gegevens of documenten
- Documenten of bestanden kunnen delen met anderen
- Het kunnen verwijderen van niet meer relevante documenten, oude concepten of tussenresultaten
- Op ieder moment deze handelingen kunnen verrichten
- Vanaf iedere locatie deze handelingen kunnen verrichten
- Kunnen voldoen aan wet- en regelgeving

#### **Toepassing**

Dataopslag wordt gebruikt voor:

- Instellingssystemen, zie verder bijlage B.
- E-mail, zie verder bijlage C.
- Individuele gebruikersdata, zie verder bijlage D.
  - o zakelijk
  - o privé
- Gezamenlijke data, zie verder bijlage D.
  - o voor onderwijs
  - o voor onderzoek, zie verder bijlage E.
  - o voor projecten
  - o per eenheid

#### **Behoefte**

De dataopslag dient zo goed mogelijk aan te sluiten bij de daadwerkelijke behoefte. Zowel kwalitatief als kwantitatief dient het juiste evenwicht gevonden te worden. Bij deze afweging spelen aspecten als capaciteit, performance, kosten, noodzaak van directe beschikbaarheid, risico van uitval, risico dat data verloren gaat en risico van verlies van vertrouwelijkheid een rol. Om een juiste afweging van belangen te garanderen is een vorm van kostentoerekening noodzakelijk. Verschillende omstandigheden leiden tot andere keuzes ten aanzien van het gewenste opslagregime.

#### **Groeibeheersing**

De ruimte die nodig is om data op te slaan groeit sterk. Doordat de kosten van opslag per volume-eenheid dalen, blijft de kostengroei nog beperkt. Gezien de toch hoge kosten is groeibeheersing van belang. Dit is alleen mogelijk als inzicht bestaat in de omvang van de groei en de reden daarvan. De eigenaar van de data zal gestimuleerd worden zo nodig een ander opslagregime te kiezen voor (delen van) de data. Dit moet met de nodige voorzichtigheid gebeuren gezien het risico van keuze voor onveilige opslag van waardevolle data dan wel de mogelijk onnodige verhoging van de kosten.

In het kader van groeibeheersing worden een aantal opslagregimes gehanteerd, die uitnodigen tot terughoudendheid in gebruik van de UT-opslag tot een noodzakelijk minimum.



Voor data welke lokaal op de werkplek wordt gebruikt zal ook de lokale harde schijf als primaire opslag gelden. Als extra beveiliging tegen verlies nodig is kan ook secundair online een replicatie/backup van deze data beschikbaar worden gesteld.

Voor volumineuze data welke gedurende langere tijd niet wordt gebruikt is veilige offline opslag aan de orde. De UT kiest ervoor om, alleen wanneer daartoe een noodzaak bestaat, zakelijke data ook op een opslagvoorziening via het netwerk online op te slaan. Verder wordt ervoor gekozen om meerdere online opslagregimes aan te bieden, zodat een afweging tussen kosten en risicobeheersing mogelijk is.

Bij instellingssystemen is de houder verantwoordelijk voor de keuze van het juiste opslagregime en voor de kosten van de dataopslag.

Bij e-mail wordt voor de hele universiteit het opslagregime bepaald en worden de kosten naar gebruik doorberekend aan de eenheden. Voor de opslag van gebruikersdata wordt een beperkt aantal mogelijke opslagregimes aangeboden en worden de kosten naar gebruik doorberekend aan de eenheden.

Het is de werknemer toegestaan het ICT- en Internetgebruik in beperkte mate in te zetten voor privédoeleinden. Dit dient echter geen meerkosten voor de UT te veroorzaken.

### **Communicerende vaten**

De grootte van bijlagen die meegestuurd worden met e-mail groeit sterk. Het huidige quotum van bijvoorbeeld 500 MB per medewerker op de mailbox kan een averechts effect hebben op de kosten. Ten eerste besteden medewerkers tijd aan het afwegen wat wel of niet weg te gooien. Als hier veel tijd mee gepaard gaan kunnen de daarmee samenhangende kosten zelfs een veelvoud zijn van de opslagkosten.

Sommige medewerkers slaan documenten op binnen een folderstructuur, anderen geven de voorkeur aan het bewaren van de documenten bij de e-mails en deze te ordenen binnen Outlook. Het dient geen redelijk doel om via quota het ene gedrag te bevorderen ten koste van het andere. Het gereserveerde opslagquotum zal dus betrekking hebben op de gezamenlijke opslag van e-mail en gebruikersdata.

Voor archivering van mailberichten kunnen bij gebruik van Outlook zogenaamde PST-files aangemaakt worden waarin alle berichten in één geheel bestand staan. Bij elke wijziging ontstaat er een volledig nieuw bestand. Als dit bestand op de netwerkschijf wordt opgeslagen wordt hiervan iedere dag weer een complete nieuwe backup gemaakt. Er gaat zodoende veel opslagcapaciteit en dus kosten mee gemoeid.

### **Cloud Computing**

Het uitbesteden of samen met anderen organiseren van delen van de ICT-voorzieningen wordt steeds populairder. Hierbij kan gedacht worden aan het gezamenlijk met andere universiteiten beheren van een instellingssysteem, het uitbesteden van de e-mail voorziening, het opslaan van documenten en het delen daarvan.

Voorgesteld wordt om voor de e-mail voorziening voor studenten, samenwerkingsomgeving en documentopslag voor studenten businesscases op te stellen, waarbij ook rekening wordt gehouden met beveiligingstechnische en juridische aspecten.

### **Volumineuze wetenschappelijke data**

Voor ruwe wetenschappelijke data geldt vaak dat deze groot in omvang zijn, maar na afloop van het project toch niet verloren mogen gaan. Er kan immers nog een vervolgonderzoek komen, een nieuwe afstudeerder of PhD'er welke nieuwe aspecten kan onderzoeken. In deze gevallen is een mogelijkheid van goedkope opslag gewenst. Dit geldt vooral voor grote (verzamelingen van) bestanden zoals experimentele meetgegevens, digitale foto's, geluidsopnamen, video's, luchtfoto's en satellietbeelden. Hergebruik van bestaande data betekent een grotere efficiëntie en lagere kosten.

Nu wordt vaak dit soort data door een medewerker enkelvoudig lokaal opgeslagen met alle risico's van dien. Wanneer de opslag minder ad-hoc gebeurt neemt de kans op hergebruik toe, wat kosten van het opnieuw verkrijgen van (soortgelijke) data voorkomt. In verband

daarmee is het terug kunnen vinden relevant van dergelijke data en documenten en het toevoegen van metadata.

## **Risicobeheersing**

Risico's kunnen verschillende oorzaken hebben:

1. Hardware storing. Er bestaan verschillende kwaliteiten hardware, maar alle hardware gaat een keer kapot. Maatregelen ter beperking van de gevolgen van storingen zijn redundant opslaan (replicatie) en backup.
2. Software bug. Foutvrije software bestaat niet en hoewel deze soort fout weinig optreedt dienen de gevolgen beperkt te worden. Dit kan door belangrijke data op technisch verschillende gegevensdragers op te slaan.
3. Menselijke fout. Goede voorlichting, beveiliging en functiescheiding zijn van belang. Vergissingen zullen echter blijven voorkomen. Versiebeheer en door de gebruiker zelf terug te halen backups kunnen de gevolgen beperken.
4. Externe oorzaak. Belangrijke data op geografisch gescheiden locaties opslaan helpt.

Iedere waarde, zoals hier boven beschreven, kan aangetast worden en deze risico's moeten beheerst worden. Het informatiebeveiligingsbeleid werkt dit in algemene zin uit. Daarbij kunnen de volgende risico's onderscheiden worden:

1. Verlies van beschikbaarheid
2. Verlies van integriteit
3. Verlies van vertrouwelijkheid
4. Onvoldoende controleerbaarheid

Vanuit dataopslagbeleid wordt gefocust op

1. Verlies van beschikbaarheid, de mate van acceptatie van dit risico kan gekwantificeerd worden met de parameter Maximaal Toelaatbare Uitvalduur. De MTU kan variëren van minder dan een dag (kort) tot meer dan een week (lang).
2. Verlies van data, dit is een subset van het risico verlies van integriteit en kan gekwantificeerd worden met de parameter Maximaal Toelaatbaar Dataverlies. De MTD kan variëren van alle werk gedurende een korte periode (weinig) tot een langere periode (veel).

Vertrouwelijkheid speelt een rol bij

- gegevens opgeslagen door instellingssystemen, zie verder het Informatiebeveiligingsbeleid.
- onderzoeksdata en onderzoeksresultaten. Dit stelt randvoorwaarden aan de opslag, dit komt verder aan de orde in bijlage E Onderzoeksdata.

## **Maatregelen**

Er zijn meerdere technische maatregelen mogelijk om het verlies van beschikbaarheid en het verlies van data te beperken. Des te meer maatregelen er binnen een opslagregime worden genomen des te lager de risico's en des te hoger de kosten voor dat regime.

**Replicatie** is het dubbel opslaan van gegevens. Als er zich een storing voordoet kan deze verholpen worden zonder dat de gebruiker daar last van heeft.

**Snapshots** is het met een bepaald interval, bv vier uur, maken van een opname van alle data. Alleen de gewijzigde data neemt hierbij extra ruimte in. De omvang van deze extra opslag hangt samen met de frequentie waarmee snapshots worden gemaakt. Deze techniek beveiligt tegen onbedoelde wijzigingen. Des te meer versies bewaard worden des te hoger de kosten. Wanneer een replicaset onbruikbaar is, doordat er zowel naar origineel als replica verkeerd is weggeschreven, kan men op de snapshot terugvallen.

**Backup** is het met een bepaald interval, bv een dag, maken van een kopie van alle data. Bij een storing kan de bewaarde versie teruggezet worden.

**Offsite backup** is het met een bepaald interval, bv drie weken, maken van een kopie van alle data en deze elders bewaren. Dit is bedoeld als achtervang bij calamiteiten.

**Prioriteit** geeft aan welke data, na een storing, als eerste van backup teruggezet zal worden.

## Opslagregimes

Er kunnen een aantal gewenste opslagregimes onderscheiden worden:

1. **Lokaal.** Directe toegang tot de data mogelijk. Niet centraal beheerd. Geschikt voor (beperkte) privé data en voor zakelijke data in combinatie met online of offline opslag. Noodzakelijk voor mobiel gebruik wanneer geen verbinding met het internet beschikbaar is. Nuttig voor vaste werkplekken, omdat bij netwerkstoring doorgewerkt kan worden. Alleen wanneer het volkomen verloren gaan van de data acceptabel is kan het maken van een kopie achterwege blijven. Aanschafkosten ca. € 0,06 – 0,50 / GB.
2. **Online.** Directe toegang tot de data mogelijk.
  - a. **Goud.** Opslag van de hoogste kwaliteit. Replicatie, snapshots, backup en offsite backup worden toegepast. Na een calamiteit wordt de data met een maximaal verlies van vier uur binnen maximaal acht werkuren teruggezet. Geschikt voor die toepassingen waarbij de MTU korter is dan een dag. Kosten ca. € 5,- / GB / jr.
  - b. **Zilver.** Opslag van hoge kwaliteit. Replicatie, backup en offsite backup worden toegepast. Na een calamiteit wordt de data met een maximaal verlies van 24 uur binnen maximaal twee werkdagen teruggezet. Geschikt voor die toepassingen waarbij de MTU korter is dan een week. Wanneer de MTD weinig is mag geen lagere kwaliteit gekozen worden. Kosten ca. € 3,- / GB / jr.
  - c. **Brons.** Opslag van normale kwaliteit. Backup wordt toegepast. Na een calamiteit wordt de data met een maximaal verlies van 24 uur binnen maximaal tien werkdagen teruggezet. Geschikt voor die toepassingen waarbij de MTU meer is dan een week. In combinatie met lokale of offline opslag geschikt voor alle toepassingen ongeacht de MTU. Kosten ca. € 2,- / GB / jr.
  - d. **IJzer.** Opslag van lage kwaliteit. Enkelvoudige opslag zonder backup. Bij een storing kan de data onherstelbaar verloren gaan. Alleen te gebruiken als secundaire opslag voor lokale data of voor data welke geheel opnieuw automatisch te genereren is. Kosten ca. € 1,- / GB / jr.
3. **Offline.** Directe toegang tot de data niet mogelijk. Na aanvraag wordt de data binnen een dag online gezet. Bij data van zeer grote omvang kan dit langer duren. ICTS heeft nog geen kostenschattting kunnen afgeven.

## B. Instellingssystemen

De data welke gebruikt wordt door de instellingssystemen is online beschikbaar. Gegevens uit het verleden kunnen nog enige tijd direct online vanuit het systeem beschikbaar moeten zijn. Er is op capaciteit te besparen door actief te archiveren waardoor de online gegevens beperkt beschikbaar blijven. Na vaststelling van het beleid zijn per instellingssysteem door de houder de MTU (Maximaal Toelaatbare Uitvalduur) en MTD (Maximaal Toelaatbaar Dataverlies) bepaald.

In het dataopslagbeleid zijn vier opslagregimes geformuleerd, zie onderstaande tabel.

Opslagregime	MTU	MTD	Relatief risico	Kosten (2010)
IJzer	eeuwig	alles	hoog	ca. € 1,- / GB / jr
Brons	twee weken	één dag	normaal	ca. € 2,- / GB / jr
Zilver	twee dagen	één dag	laag	ca. € 3,- / GB / jr
Goud	acht uur	vier uur	zeer laag	ca. € 5,- / GB / jr

Op grond van de gekozen MTU en MTD is een gewenst opslagregime genoteerd. Merk op dat voor websites en portals een beter opslagregime is geselecteerd dan gedefinieerd.

In de tabel op de volgende pagina is de kolom omvang gebaseerd op de som van de huidige omvang van de achterliggende database en de dataopslag van het front-end systeem. Over de dataopslag van de fysieke servers kan nog geen duidelijkheid verkregen worden. Backup vindt nu nog steeds plaats op tape, de kolom omvang geeft de netto enkelvoudige opslag weer.

Functie	Houder	Systeem	MTU	MTD	Opslag-regime	Omvang
Institutional Repository	B&A	ePrints	twee dagen	alles	zilver	geen DB
Lokaal Bibliotheek Systeem LBS	B&A	OCLC	halve dag	halve dag	goud	intern
Federated Search	B&A	Metalib SFX	één dag	één dag	zilver	intern
Post- en Archiefsysteem	B&A	Decos	één dag	halve dag	goud	100 GB
Facilitair Management en Informatie Systeem FMIS	FB	Planon	één dag	één dag	goud	20 GB
Financiële administratie	FEZ	Oracle Financials OFI	één dag	halve dag	goud	1360 GB <sup>5</sup>
Inventarisbeheer	FEZ	IRIS	twee dagen	één dag	zilver	720 GB <sup>6</sup>
Factureren	FEZ	FAMEUS via FIT	twee dagen	één dag	goud	<sup>5</sup>
Declareren	FEZ	RRVD via FIT	twee dagen	één dag	goud	<sup>5</sup>
Contractenregister	FEZ	COBSY	twee weken	één dag	zilver	<sup>6</sup>
Kasboek	FEZ	KAS via FIT	twee dagen	één dag	goud	<sup>5</sup>
Tijdregistratie	FEZ	TAS	twee dagen	één dag	zilver	<sup>6</sup>
Rekenmodellen uurtarieven	FEZ	UUR via FIT	twee dagen	één dag	goud	<sup>5</sup>
Goedkeuringsworkflow	FEZ	GKS	één dag	één dag	zilver	<sup>6</sup>
Interne doorberekening	FEZ	INTD	twee weken	één dag	zilver	<sup>6</sup>
Doorberekening telefoon	FEZ	DOT via FIT	twee weken	één dag	goud	<sup>5</sup>
Europees betaalsysteem	FEZ	APRO	één dag	halve dag	goud	<sup>5</sup>
Research Verantwoording	FEZ/CC	Metis	twee dagen	één dag	zilver	20 GB
Management Informatie	FEZ/CC	MISUT	één week	één dag	brons	120 GB
Personeelsinformatie	HR	Oracle HRM	één dag	halve dag	goud	<sup>5</sup>
Selfservice personeel	HR	Webapplicaties	twee dagen	halve dag	goud	<sup>6</sup>
Inhuur studenten	HR	UTFLEX	twee dagen	halve dag	goud	<sup>6</sup>
Websites en portals	S&C	meerdere	vijf minuten	één dag	platina	extern
CRM	S&C	Orca	één dag	één dag	goud	90 GB
Studenteninformatie	S&O	Osiris	één dag	één dag	goud	745 GB
Leeromgeving	S&O	Blackboard	één dag	één dag	goud	2275 GB
Plagiaatdetectie	S&O	Ephorus	één dag	één dag	goud	geen DB
Promovendi-informatie	S&O	GATS	één dag	één dag	goud	<sup>6</sup>
Rooster	S&O	Syllabus				10 GB
Identity Management Single Sign On Accounts aanmaken	ICTS	RUN, Sun IDM	vijf minuten één dag	één dag	zilver	380 GB
Telefoongids	ICTS	webapplicatie	halve dag	één week	goud	geen DB
Oracle Database	ICTS	Oracle Database				30 GB
Business Objects	ICTS	Business Objects				40 GB

Sommige systemen hebben geen aparte database en bij sommige systemen is de storage intern en al verrekend in de systeemkosten.

<sup>5</sup> Oracle Financials, Oracle HRM en alle afgeleide systemen maken gebruik van dezelfde database

<sup>6</sup> Andere UT-systemen welke de data niet binnen Oracle Applications opslaan maken gebruik van dezelfde database

## C. E-mail

Studenten moeten aantrekkelijke voorzieningen aangeboden krijgen, ook om de aantrekkingskracht van de UT als studielocatie te bevorderen. Tegelijkertijd kan de UT “gratis” voorzieningen van Google etc. niet beconcurreren. Studenten zullen zelf de afweging maken of en hoe de voorzieningen te combineren. Onderzocht dient te worden of dergelijke voorzieningen via een contract door de UT aangeboden kunnen worden. “If you can’t beat them, join them.”

Medewerkers, studenten en alumni hebben de beschikking over een mailbox. De ruimte welke ter beschikking gesteld wordt via de centrale mailservers is momenteel gemaximeerd op bijvoorbeeld 500 MB per gebruiker, maar kan op aanvraag verhoogd worden. Gebruikers kunnen zelf andere mailarchieven aanmaken en deze op een lokale of netwerkschijf beheren. Een laag quotum heeft een averechts effect op de kosten. Ten eerste besteden medewerkers tijd aan het afwegen wat wel of niet weg te gooien, welke kosten een veelvoud van de opslagkosten kunnen bedragen. Ten tweede worden er voor archivering PST-files aangemaakt en op de netwerkschijf gezet, aangezien deze door Outlook iedere keer geüpdate worden, wordt hiervan iedere dag weer een complete backup van gemaakt. Daarom dient het quotum voor de medewerkersmailbox makkelijk uit te breiden te zijn en dient het gebruik van PST-files ontmoedigd te worden.

Uit analyse van ca. 2 miljoen e-mails (dagelijks: 15000 – 35000) welke door medewerkers begin van dit jaar zijn ontvangen blijkt dat de grootte van het gemiddelde e-mailtje 20 kB bedraagt (dagelijkse variatie: 10 – 30 kB). Echter slechts 10% van de e-mails is groter dan dit gemiddelde (dagelijkse variatie: 3 – 20%). Ongeveer 1% van de e-mails is groter dan 200 kB (dagelijkse variatie: 0,3 – 2 %). De omvang van deze zeer grote e-mails bedraagt ongeveer de helft (53%, dagelijkse variatie: 20 – 73%) van de totale ontvangen mail. Ongeveer 0,15 % van de e-mails (absoluut 5 – 75 per dag) is groter dan 2 MB (dagelijkse variatie: 0,03 – 0,26 %). De omvang van deze enorm grote e-mails bedraagt ongeveer een derde (32%, dagelijkse variatie: 8 – 52%) van de totale ontvangen mail.

Over de tijd groeit de grootte van mailboxen van gebruikers. De noodzaak van groei dient wel beheerst te worden.

Oneigenlijk gebruik treedt op door het ontvangen van grote hoeveelheden van spam. Door voordat de spam wordt afgeleverd bij de gebruiker deze al weg te filteren kan een besparing behaald worden op de benodigde opslag.

Wanneer via e-mail samen wordt gewerkt aan een document dan worden meerdere kopieën van hetzelfde document opgeslagen. Wanneer gebruikers de beschikking hebben over een samenwerkingsomgeving of documentmanagementsysteem dan zal het rondmailen van documenten afnemen.

Via voorlichting zouden gebruikers gestimuleerd kunnen worden grote oude inmiddels niet meer relevante e-mail op te schonen. Het risico hierbij is dat de hoeveelheid werk meer kost dan de kosten van opslag.

### **Risico's**

E-mail wordt door gebruikers gezien als een eerste levensbehoefte, net als water en elektriciteit. Bij calamiteiten blijkt e-mail een onontbeerlijk communicatiemiddel voor de instelling naar de medewerkers. Een langdurige uitval van beschikbaarheid kan verder een negatieve impact op de goede naam van de instelling hebben.

De impact uitvalsduur wordt hoog ingeschat. Voor dataopslag voor e-mail zal het opslag-regime Zilver gehanteerd te worden in verband met de gegeven opzet van Microsoft Exchange 2010.

Voor die e-mail waarbij de gewone maatregelen onvoldoende worden geacht om de vertrouwelijkheid te garanderen, is het aan de gebruiker zelf om aanvullende maatregelen te treffen. Hierbij kan gedacht worden aan encryptie middels private key technologie.

### Vertrek medewerkers en studenten

Bij beëindiging van het dienstverband wordt in principe het e-mail account na een maand gesloten en is de data niet meer beschikbaar. De data wordt na drie maanden vernietigd. Deze periode geeft medewerker en leidinggevende genoeg tijd om zo nodig passende maatregelen te treffen. Wanneer accounts, na bijvoorbeeld pensionering, worden gehandhaafd, blijven de kosten aan de betreffende eenheid toegerekend worden. Studenten accounts worden na drie maanden omgezet in alumni accounts.

### Externe mogelijkheden

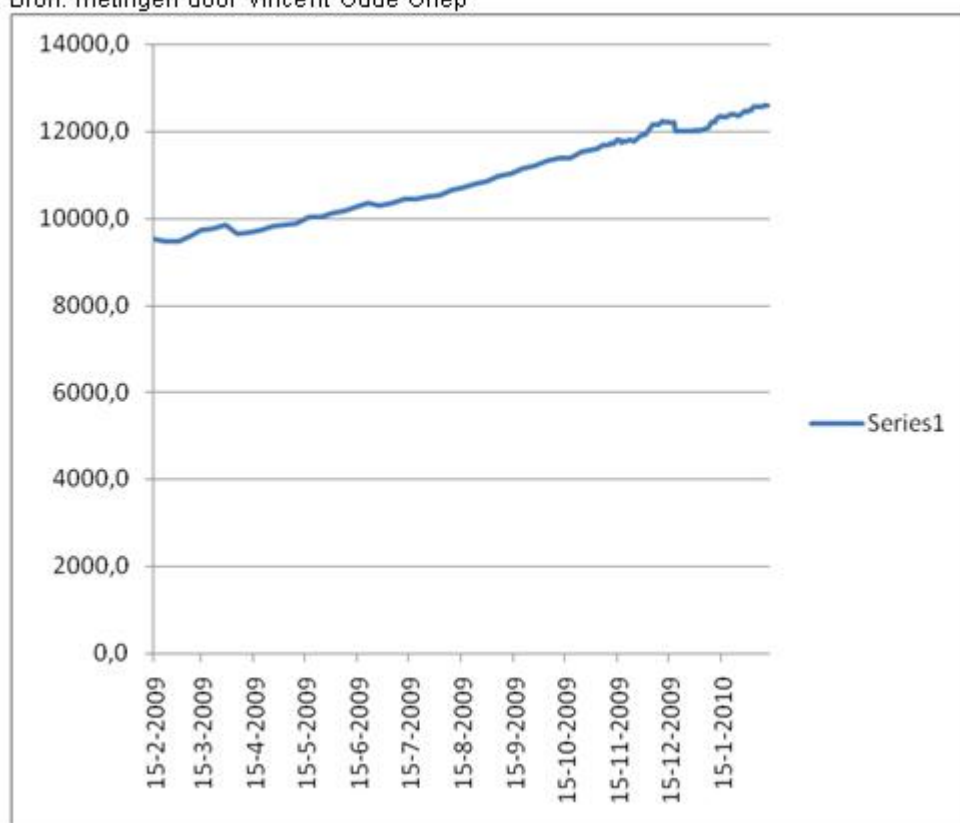
Bedrijven als Google en Microsoft doen onderwijsinstellingen op het eerste gezicht aantrekkelijke aanbiedingen om de mail voor medewerkers en studenten te verzorgen. De beveiligingstechnische en juridische consequenties zijn niet geheel triviaal. Er wordt verder onderzoek gedaan naar business-cases voor de afweging of doorgedaan wordt met het volledig in eigen beheer houden van e-mail of dat voor andere vormen van beheer gekozen zal moeten worden om zo te komen tot right-sourcing.

## D. Gebruikersdata

Werkplekken beschikken over een lokale dataopslag. Daarnaast wordt er aan medewerkers en studenten een netwerkopslag aangeboden. Verder worden vele mobiele gegevensdragers gebruikt. De behoefte aan dataopslag voor onderzoek loopt sterk uiteen, het hangt sterk samen met het precieze vakgebied en de gekozen werkwijze van individuele onderzoekers.

### Historie dataopslag MB/GW/Diensten en een deel van CTW

Bron: metingen door Vincent Oude Griep



De online dataopslag stijgt momenteel met 30% per jaar.

exclusief functionele emailaccounts

aantal tel.gids 1-6-2009	gids	capaciteit uit DVO 1-1-2010			capaciteit gemeten 1-7-2009			capaciteit afwijking tov DVO 1-7-2009		
		persoonlijke data	groeps data	e-mail	persoonlijke data	groeps data	e-mail	persoonlijke data	groeps data	e-mail

Quota (Gb)			5	5	0,5						
eenheid											
MB	446	401	2230	2230	223	1071	410	170	-1159	-1820	-53
GW	307	295	1535	1535	154	517		125	-1018	-1535	-29
TNW	712	614	3560	3560	356	4610		199	1050	-3560	-157
EWI	712	625	3560	3560	356				-3560	-3560	-356
CTW	414	397	2070	2070	207	5384		400	3314	-2070	193
B&A	48	48	240	240	24	97	9	8	-143	-231	-16
PA&O	66	66	330	330	33	99	42	15	-231	-288	-18
FB	133	114	665	665	66,5	87	70	22	-578	-595	-44,5
ICTS	123	125	615	615	61,5	262	396	27	-353	-219	-34,5
S&O	177	173	885	885	88,5	174	103	31	-711	-782	-57,5
S&C SECR en CVB	53	51	265	265	26,5	152	482	24	-113	217	-2,5
FEZ	65	64	325	325	32,5	127	126	18	-198	-199	-14,5
	56	56	280	280	28	80	51	12	-200	-229	-16
MISC	113		565	565	56,5				-565	-565	-56,5
TSM	37	31	185	185	18,5	32	8	10	-153	-177	-8,5
CTIT	4		20	20	2				-20	-20	-2
SANTA	20		100	100	10				-100	-100	-10
MESA	2		10	10	1				-10	-10	-1
CATU	26		130	130	13				-130	-130	-13
IMPAC	5		25	25	2,5				-25	-25	-2,5
IGS	3		15	15	1,5				-15	-15	-1,5
MD	5		25	25	2,5				-25	-25	-2,5
BMTI	3		15	15	1,5				-15	-15	-1,5
CAT	2		10	10	1				-10	-10	-1
mede- werkers	3532		17660	17660	1766	12692	1697	1061	-4968	-15963	-705
studen- ten	8500		42500	42500	4250						

Uit de tabel blijkt dat het daadwerkelijke gebruik van dataopslag aanzienlijk afwijkt van de afspraken in de bestaande dienstverleningsovereenkomst. De behoeften van de diverse diensten en faculteiten lopen aanzienlijk uiteen. Om deze reden dienen dan ook de afspraken over de gereserveerde capaciteit per eenheid gemaakt te worden.

Er is nog weinig aandacht voor de levenscyclus van documenten. Wanneer een rapport eenmaal is opgeleverd, het lesmateriaal in BlackBoard is geplaatst of een artikel is gepubliceerd dan zijn de vele tussenliggende conceptversies niet meer nodig.

Data wordt regelmatig niet veilig opgeslagen doordat alleen lokale opslag wordt gebruikt. Tegelijkertijd wordt voor grote datasets de huidige centrale opslag als te duur gezien. Daarom dienen er meerdere opslagregimes aangeboden te worden en dienen gebruikers middels voorlichting gestimuleerd te worden de juiste keuzes te maken. Om te voorkomen dat er bij een netwerkstoring niet doorgewerkt kan worden zal alle gebruikersdata ook lokaal worden opgeslagen.

Via voorlichting zouden gebruikers gestimuleerd kunnen worden inmiddels niet meer relevante documenten op te schonen. Het risico hierbij is dat de hoeveelheid werk meer kost dan de kosten van opslag.

Terugvinden van data, documenten wordt door de toename van informatie ook steeds meer een probleem. Wat kan metadata hiervoor betekenen? Wat kunnen zoekmachines hierbij betekenen? Willen we dat steeds meer individuen Google Desktop installeren, met mogelijke gevolgen voor de performance, of zijn er generieke oplossingen mogelijk/wenselijk?

### **Samenwerking en groeibeheersing**

Wanneer samen wordt gewerkt, door medewerkers of studenten, aan een document of project dan worden meerdere kopieën van hetzelfde document opgeslagen. Wanneer gebruikers de beschikking hebben over een samenwerkingsomgeving of document-managementsysteem dan zal het dubbel opslaan van documenten afnemen.

Binnen de diverse gremia worden vele vergaderstukken vooraf verspreid. Door hiervoor een gezamenlijke fileshare, samenwerkingsomgeving of documentmanagementsysteem voor te gebruiken wordt veel dubbele opslag voorkomen. Door de eenduidigheid neemt de gebruiksvriendelijkheid toe en neemt de kans op fouten af.

De UT dient voor dit probleem een oplossing te vinden. Hierbij kan gedacht worden aan:

1. Een eenvoudige manier voor groepen om gemeenschappelijke dataruimte aan te vragen op een netwerkdrive. Speciale aandacht is hierbij nodig voor het weer opschonen van de niet meer actieve groepen.
2. Een eenvoudige manier voor medewerkers om anderen rechten te geven en te beheren op delen van hun dataruimte op een netwerkdrive. Dit is beheersmatig minder complex dan optie 1, aangezien de verantwoordelijkheid bij de gebruiker ligt.
3. Als samenwerkingsomgeving is een Sharepoint omgeving op Surfgroepen.nl nu gratis beschikbaar. Deze mogelijkheden kunnen beter benut worden door de voorlichting hierover te verbeteren. Opgemerkt moet worden dat deze landelijke voorziening waarschijnlijk per 2012 wordt opgeheven.
4. Zelf als UT Sharepoint, of een andere samenwerkingsomgeving, aanbieden. Dit brengt ook beheerskosten met zich mee. Voordat hiervoor gekozen kan worden dient een business-case te worden uitgewerkt.

Het huidige "laissez faire" beleid houdt in dat gebruikers eigen oplossingen gaan zoeken en houdt de bestaande problemen in stand. Om dubbel werk te voorkomen is één standaard oplossing het beste. Het meer gebruiken van Surfgroepen is de goedkoopste oplossing, terwijl deze kwalitatief niet onder doet voor de andere mogelijkheden. Een nieuwe Collaboration omgeving voor de samenwerking tussen instellingen is als opvolger van Surfgroepen in ontwikkeling.

Voor de korte termijn kiest de UT daarom om de bekendheid te vergroten met en de wijze waarop de bestaande mogelijkheden optimaal gebruikt kunnen worden middels voorlichting. UIM volgt de ontwikkelingen bij Surf, zal business-cases ontwikkelen en zal zonedig voorstellen doen om tot right-sourcing te komen.



## Privégebruik

Beperkt gebruik voor privédoeleinden is geregeld in de Gedragscode ICT voor medewerkers.<sup>7</sup> Dit houdt in dat lokale opslag van privédata toegestaan wordt, maar online opslag van deze data niet toegestaan is.

1. Gebruikers zijn zelf verantwoordelijk voor de backup van privé-bestanden, bijvoorbeeld door ze periodiek handmatig weg te schrijven naar een eigen DVD, losse harde schijf of USB-stick.
2. ICTS wordt gevraagd om de omvang te rapporteren van het geregistreerde en daadwerkelijk gebruik van de opslagcapaciteit op het netwerk per medewerker in verband met de doorberekening. Grootverbruikers worden hierbij gesignaleerd.
3. ICTS zal de werkplekken zo in te richten dat folders als "My Music" en "My Pictures" naar folders op de lokale schijf wijzen.
4. Via voorlichting worden medewerkers geïnformeerd over het juiste gebruik van de diverse folders en soorten opslag, ook in verband met de doorberekening van kosten.

## E. Onderzoeksdata

Individuele onderzoekers, kleine en grote onderzoeksgroepen gebruiken dataopslag voor discussies, ideeën, concepten, artikelen, illustraties, ruwe en bewerkte data. Voor de keuze van de dataopslagmethode wordt vaak de weg van de minste weerstand gekozen. Hierdoor worden zowel onnodige risico's gelopen als dat er kansen worden gemist.

Het is voorgekomen dat onderzoeken gestopt moesten worden omdat de benodigde data verloren was gegaan. Daarom is het van groot belang dat gebruikers zich bewust zijn van de kans op dataverlies en zij verstandig gebruik maken van de voorzieningen welke worden geboden. Hiervoor is het nodig dat zij op de hoogte zijn van deze voorzieningen en deze zo worden aangeboden dat zij voldoen voor het beoogde doel en dat individuele gebruikers hier vertrouwen in hebben. Uiteindelijk blijft de gebruiker zelf verantwoordelijk voor de data.

Voor ruwe wetenschappelijke data geldt vaak dat deze groot in omvang zijn, maar na afloop van het project toch niet verloren mogen gaan. Er kan immers nog een vervolgonderzoek komen, een nieuwe afstudeerder of PhD'er welke nieuwe aspecten kan onderzoeken. In deze gevallen is een mogelijkheid van goedkope opslag gewenst. Dit geldt vooral voor grote (verzamelingen van) bestanden zoals digitale foto's, geluidsopnamen, video's, luchtfoto's en satellietbeelden. Hergebruik van bestaande data betekent een grotere efficiëntie en lagere kosten.

Voor de juiste opslag van data is de leider van het betreffende onderzoeksprogramma verantwoordelijk. Zolang het onderzoek loopt is, buiten beschikbaarheid en integriteit, vaak ook de vertrouwelijkheid van belang. Soms omdat de opdrachtgever of dataleverancier eisen hieraan stelt, maar ook om te voorkomen dat resultaten breed bekend worden voordat formeel is gepubliceerd of een promotie heeft plaatsgevonden. Tijdens het onderzoek kan de samenwerking beter gefaciliteerd worden middels "collaboratories", er dient onderzoek gedaan te worden hoe dit beter verwezenlijkt kan worden. De voorlichting door ICTS hoe bestaande mogelijkheden van gezamenlijke dataopslag kunnen worden toegepast wordt geïntensiveerd.

Voor langdurige dataopslag van volumineuze wetenschappelijke datasets wordt door ICTS een mogelijkheid gecreëerd om data goedkoop offline op te slaan, zodat deze in afwachting van een mogelijk vervolgonderzoek niet verloren gaat.

Bij de publikatie van de resultaten van onderzoek wordt steeds vaker het principe van Open Access<sup>8</sup> toegepast, in 2006 heeft de UT de Berlin Declaration<sup>9</sup> ondertekend. Een belangrijk onderdeel van Open Access is Open Data, immers alleen wanneer alle (ruwe) data met het artikel wordt gepubliceerd zijn analyses te herhalen. In 3TU verband is door de bibliotheken

<sup>7</sup> zie <http://www.utwente.nl/im/security/Gedragscode%20ICTmedewerkers/>

<sup>8</sup> Zie [www.openaccess.nl](http://www.openaccess.nl)

<sup>9</sup> Zie <http://oa.mpg.de/openaccess-berlin/signatories.html>

het 3TU datacentrum<sup>10</sup> opgericht om datasets duurzaam op te slaan en te ontsluiten. Voordelen van het gebruik van deze voorziening zijn:

- er geen fysiek verval optreedt. Door gebruik te maken van geavanceerde backup voorzieningen is men ervan verzekerd dat de onderzoeksgegevens altijd in hun oorspronkelijke staat kunnen worden geraadpleegd.
- de datasets gemakkelijk kunnen worden teruggevonden en opgevraagd. Door toekenning van metadata volgens internationale standaarden die onlosmakelijk met de gegevens verbonden blijven.
- de formats waarin de data zijn opgeslagen leesbaar blijven. De programmatuur waarmee de datasets worden gelezen, wordt beschreven en met een koppeling naar de datasets opgeslagen.
- copyrights worden beschermd. De eigenaar bepaalt zelf of en wanneer zijn data voor anderen toegankelijk wordt.

Deze (pilot) voorziening is binnen de UT onvoldoende bekend en verdient een grotere bekendheid. Het meest voor de hand ligt het om aan te sluiten bij de bestaande informatievoorziening rondom Open Access zoals dat op de website en in de Open Access week georganiseerd wordt.

---

<sup>10</sup> Zie <http://datacentrum.3tu.nl>