

Eindrapportage

Hublab-2. Naar succesvolle implementatie van het Liferay platform in historisch onderzoek.

(Toward successful implementation of the Liferay platform in historical research)



Stefan Dormans
Jan Kok

Virtual Knowledge Studio for the Humanities and Social Sciences

Internationaal Instituut voor Sociale Geschiedenis

Datum: januari 2010



Deze publicatie verschijnt onder de Creative Commons licentie Naamsvermelding 3.0 Nederland.

Inhoudsopgave

1. Inleiding	5
2. Werkpakketten	6
WP1. Gebruikersonderzoek	6
WP2. Verhoging gebruikersvriendelijkheid	9
WP3. Datasharing	12
WP4. Verbetering communicatie	12
WP5. Ondersteuning workflows	13
WP6. Interoperabiliteit	13
WP7. Kennisdisseminatie	14
WP8. Helpdesk	15
WP9. Projectmanagement	15
3. Kennisdisseminatie en –uitwisseling	15
4. Evaluatie	17
Standlijnen en planning	20
5. Financiële verantwoording	21
Bijlagen	23
Bijlage 1: verslag van literatuurstudie	24
Bijlage 2: eindverslag gebruikersonderzoek	37
Bijlage 3: vergelijking Sharepoint en Liferay	52
Bijlage 4: voorstel licentieregeling	58
Bijlage 5: NCESS paper	62

Samenvatting

In het project HubLab2 werd de dagelijkse praktijk van een aantal collaboratories in de sociaal-economische geschiedenis onderzocht. Welke rol speelde online samenwerkingsplatforms in hun onderlinge communicatie, hun data sharing, en hun opslag van bestanden en documenten? Tegelijkertijd vormden hun ervaringen het uitgangspunt van verdere verbeteringen van het platform, in dit geval gebaseerd op het open source pakket Liferay.

De resultaten van het project zijn een uitgebreid gebruikersonderzoek, dat ook is ingezet in wetenschappelijke voordrachten en publicaties, zowel ten behoeve van wetenschapsdynamici als historici. Aan de technische kant zijn veel verbeteringen doorgevoerd, die het platform veel gebruikersvriendelijker hebben gemaakt. Enkele innovaties konden als open source worden aangeboden aan de omvangrijke gemeenschap van Liferay gebruikers. Documenten rond datasharing vinden hun weg in het collectiebeleid van het IISG.

Hoewel het platform eenvoudig, visueel aantrekkelijk, en goed navigeerbaar is, blijft het feitelijk gebruik ervan in de collaboratories beperkt tot het opslaan van teksten en bestanden. Het gehoopte stroomlijnen en documenteren van discussies, rond bijvoorbeeld definities en data standaardisatie, blijft uit. Mogelijk komt dit door de relatieve late oplevering van het platform en de vrij trage werkprocessen binnen de meeste collaboratories. Dit zou betekenen dat het nog te vroeg is voor een gefundeerd eindoordeel. Het is echter ook aannemelijk dat de vrijblijvende organisatiestructuur van de meeste onderzochte collaboratories een intensivering van het gebruik in de weg staat.

Op basis van de inzichten uit het gebruikersonderzoek komen we tot de volgende zes aanbevelingen ten aanzien van het optimale gebruik van collaborative software in de geesteswetenschappen en de sociale wetenschappen:

- Op dit moment is de grootste meerwaarde van *collaborative software* voor de dagelijkse onderzoekspraktijk de mogelijkheid die het biedt om grote bestanden te delen en te bewaren. De *document library* dient dan ook een centrale plaats in het platform te krijgen, stabiel te zijn, en eenvoudig te bedienen en te vinden zijn.
- Voor de communicatie tussen onderzoekers zal *collaborative software* niet snel de bestaande communicatiemiddelen vervangen. In aanvulling op deze software

blijft het dan ook nodig om na te denken over de rol van e-mail en mailinglijsten – alsmede *face-to-face contact* - in het samenwerkingsproces.

- Gezien het interpretatieve karakter van geesteswetenschappelijk onderzoek vormt *collaborative software* een te smalle basis voor gedistribueerde samenwerkingsprojecten van dit formaat. In de opzet van dergelijke projecten dient dan ook altijd voldoende rekening gehouden te worden – in praktische en financiële zin – met plenaire bijeenkomsten.
- De implementatie en het beheer van *collaborative software* vereist een goede institutionele inbedding, waarbij gezorgd wordt voor stabiele financiële en technische ondersteuning van het pakket. Zelfs een relatief licht *open source* pakket als Liferay vereist veel ICT-ondersteuning.
- Om de bereidheid onder onderzoekers te vergroten om gebruik te maken van *collaborative software* is een overzichtelijke, herkenbare en intuïtieve interface van belang. Uitgangspunt dienen daarbij niet de mogelijkheden van het pakket te zijn, maar de wensen van de onderzoekers.
- Projectleiders dienen een duidelijke visie op de rol van de *collaborative software* in het gehele samenwerkingsproces te hebben en tevens zelf het goede voorbeeld geven. Zonder initiatief en visie van sommigen zal de meerderheid van de gebruikers niet geneigd zijn om actief gebruik te maken van de software.

1. Inleiding

Hoe werken collaboratories in de praktijk? Het doel van dit project was na te gaan of en hoe de bij de voorgaande SURF tender ingerichte omgevingen functioneren als ondersteuning bij collectief onderzoek. In het kader van het SURFshare project Hublab (2007-2008) zijn virtuele platforms ingericht voor een aantal werkgroepen van historici. Het gaat hierbij om internationale samenwerkingsprojecten rond het verzamelen, standaardiseren, analyseren, ter beschikking stellen en archiveren van data. Met deze data kunnen wereldomspannende processen rond arbeid en de verdeling van welvaart worden geanalyseerd. De betreffende data zijn veelzijdig van aard en kunnen zowel kwalitatief (beroepsaanduidingen) als kwantitatief (prijzen, organisatievormen, actievormen) als serieel zijn (lonen) en ook op microniveau liggen (huwelijken, geboorten, migratiebewegingen, beroepen).

Voor elke groep is een platform ingericht met het open source pakket Liferay. Hoe werkt dit in de praktijk? Wie maken precies gebruik van het collaboratory platform en van welke functionaliteiten daarvan? Hoe sluit het platform aan op de onderzoekspraktijk, op het groepsproces en op de verwachtingen van gebruikers? In het kader van dit project is bovendien bekeken hoe de omgeving aantrekkelijk gemaakt kan worden, of de omgeving naar andere werkgroepen kan worden uitgebreid en hoe datasharing kan worden ondersteund.

Hublab-2 is vooral een gebruikersonderzoek van de collaboratory platforms. Welke aangeboden functionaliteiten sluiten goed aan bij de onderzoekspraktijk? Welke functionaliteiten verdienen nadere aanpassing of blijken in de praktijk overbodig? Hoe 'presteert' het pakket Liferay wat dit betreft ten opzichte van SharePoint? In hoeverre werken de mensen buiten het systeem om (bijvoorbeeld via email, mailinglists) en waarom? Ligt dat aan de omgeving of eerder aan de manier van werken (taakverdeling, hiërarchie, mate van vrijwilligheid) binnen deze collaboratories? Onder welke omstandigheden lijken collaboratories in de humaniora optimaal gebruik te maken van de platforms of *virtual research environments*? Om inzicht te krijgen in het samenspel tussen wetenschappelijke groepsdynamiek en virtuele werkomgevingen is een zestal groepen intensief gevolgd en onderling vergeleken.

Tevens zijn een aantal praktische verbeteringen gerealiseerd. Het gaat hierbij om vergroting van de gebruikersvriendelijkheid, het beter ondersteunen van workflows, het optimaliseren van datasharing, de verbetering van de communicatie binnen de groep en bevordering van de interoperabiliteit van het materiaal verzameld binnen de verschillende groepen.

Het Hublab-2 project liep van 1 januari 2009 tot 31 december 2009. De projectbegroting van het IISG (in samenwerking met de VKS) was € 71.350, waarvan € 37.459 uit het SURFshare programma.

2. Werkpakketten

WP1. Gebruikersonderzoek

De nadruk in het project lag op het volgen van de gebruikersgroepen in hun onderzoeks- en samenwerkingspraktijk, waarmee we een scherper zicht wilden krijgen op de rol die het Liferay platform daarin kan spelen. Uitgaande van onze eerdere ervaringen is het gebruik van virtuele ontmoetingsruimten sterk bepaald door (1) de urgentie van de gedeelde onderzoeksproblematiek, (2) de aanwezigheid en het gebruik van alternatieve communicatiemiddelen, (3) de gebruikersvriendelijkheid van het platform en (4) de sociale verhoudingen binnen de onderzoeksgroep. Om deze hypothesen te toetsen willen we in de eerste plaats sterk verschillende groepen vergelijken en willen we in de tweede plaats een aantal momentopnames maken, zodat we kunnen zien of technische verbeteringen effect hebben.

Doel: In dit grootste werkpakket zijn gedurende de hele looptijd van het project nauwgezet de ervaringen van de gebruikers met de collaboratory-omgeving Liferay in kaart gebracht, geanalyseerd en gecommuniceerd naar enerzijds de software ontwikkelaars, anderzijds de SURFshare gemeenschap. Einddoel is een eindverslag waarin concrete aanbevelingen worden gedaan voor de optimalisatie van virtuele platforms voor geesteswetenschappelijk onderzoek.

Aanpak:

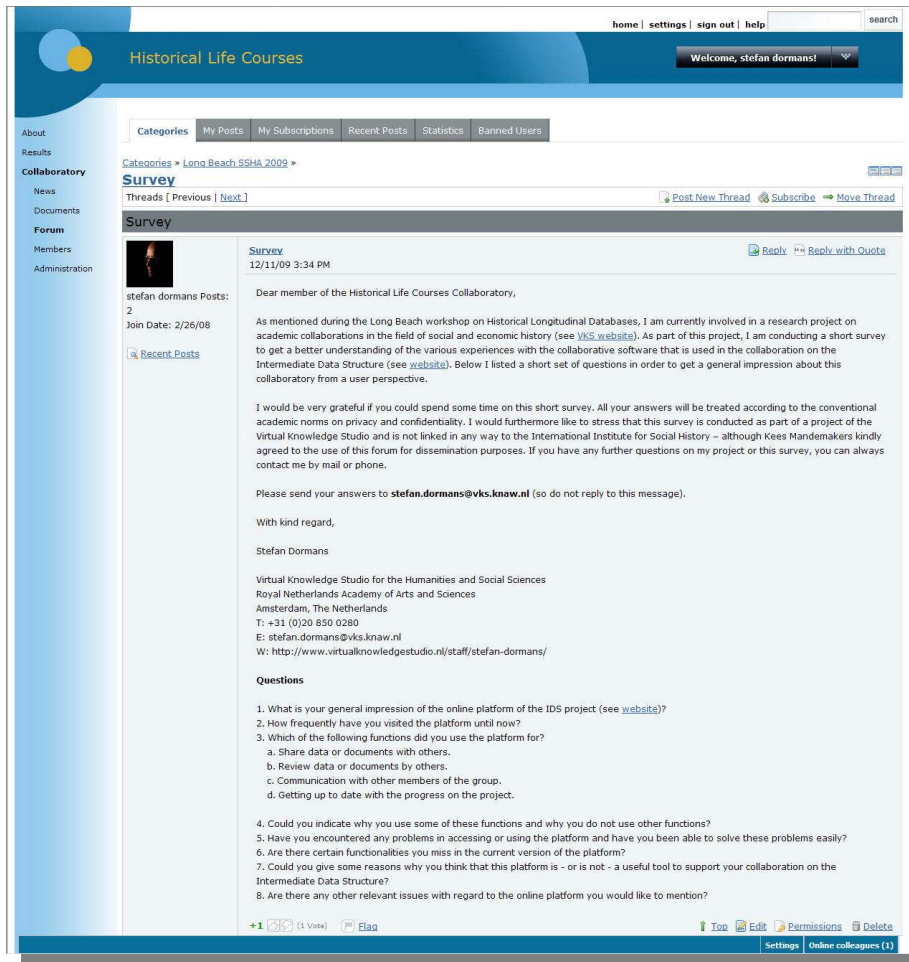
Literatuurstudie: In de periode januari-maart is een literatuurstudie uitgevoerd (zie bijlage 1). Deze studie diende enerzijds om het gebruikersonderzoek theoretisch te funderen en anderzijds verkende de studie de specificiteiten van samenwerking met behulp van virtuele platforms in de geesteswetenschappen. In de natuur- en levenswetenschappen is er een langere traditie in dit soort samenwerking en deze projecten zijn uitgebreid bestudeerd binnen diverse wetenschapsgebieden. In de literatuurstudie wordt een vertaalslag gemaakt van dit debat naar de praktijken in de geesteswetenschappen.

Participerende observaties en interviews: Gedurende de hele duur van het project zijn voor het gebruikersonderzoek de diverse collaboratories van het IISG geobserveerd, zowel online als offline. De online observaties betrof het volgen van

mailinglijsten, emailcorrespondenties, en het verkeer op het Liferay platform, alsmede het doornemen en analyseren van alle relevante documentatie die via deze wegen in 2009 gedeeld zijn. De offline observaties vonden plaats tijdens workshops van collaboratories in binnen en buitenland, tijdens vergaderingen over collaboratories en Liferay binnen het IISG, en buiten die formele bijeenkomsten om is gedurende het hele jaar nauw contact geweest met de onderzoekers, de stad en de ICT afdeling van het IISG. Het feit dat de Virtual Knowledge Studio een gebouw deelt met het IISG faciliteerde deze continue informele observaties natuurlijk in sterke mate. Tijdens de gehele onderzoeksperiode zijn tevens diverse etnografische interviews gehouden met gebruikers en ontwikkelaars van het Liferay platform. Aanvankelijk zouden extra diepte-interviews na het afnemen van de enquête afgenomen worden, maar door de vertragingen in WP2 is besloten om alleen enkele korte vraaggesprekken met enkele respondenten te hebben naar aanleiding van de enquête

In het kader van de participerende observaties en interviews is o.a. een tweedaagse workshop bezocht van de “Global Collaboratory on the History of Labour Relations” (13 en 14 maart), alsmede een workshop van de Historical Life Courses Collaboratory in Long Beach, Ca. (11 tot en met 15 november). Tevens werd in Long Beach de jaarlijkse conferentie bezocht van de American Social Science History Association. Diverse andere collaboratories gebruiken deze conferentie om hun resultaten te tonen. Er is o.a. gesproken met de projectleiders van de Historical Violence Database en van Founders and Survivors. Australian Life Courses in Historical context 1803-1920.

Enquête: In december is onder de leden van de twee meest actieve collaboratories, namelijk ‘Historical Life Courses’ en de ‘Global Collaboratory on the History of Labour Relations in the period 1500-2000’, een enquête over de ervaringen met Liferay afgenomen. Beide collaboratories hebben in Liferay respectievelijk 37 en 88 leden, waarbij aangetekend dient te worden dat een deel van die leden niet actief participeert in het onderzoekswerk van de collaboratory. In het geval van de HLC-collaboratory zijn uiteindelijk 25 mensen benaderd en in het geval van de Global Labour Relations-collaboratory is de enquête gestuurd naar eveneens 25 mensen. In totaal hebben 9 respondenten op de enquête gereageerd. In de enquête werden een achttal korte vragen gesteld over de algemene indruk die men van Liferay had, de frequentie waarmee men Liferay bezocht, de functionaliteiten die men al dan niet gebruikte, de gebruiksvriendelijkheid en ondersteuning, alsmede de bruikbaarheid van Liferay voor het samenwerkingsproces. Voor de verspreiding van de enquête is gebruik gemaakt van het Forum en de mailfunctie van Liferay. In het geval van de HLC-collaboratory zijn in dit kader alle leden geabonneerd op een thread in het forum, waardoor ze via de uitnodiging om te reageren op de enquête direct ook geattendeerd werden op de mogelijkheden van het forum.



Figuur 1: Post in Liferay-forum (https://collab.iisg.nl/group/hsn/forum/-/message_boards/message/62424)

Kwantitatieve analyse: Om het gebruik van de collaborative software te kunnen volgen is AWSTAT-software geïnstalleerd. Deze software maakt het mogelijk om het bezoek en gebruik van websites in kwantitatieve zin te monitoren. Het verslag van de kwantitatieve analyses is opgenomen in bijlage 2.

Vergelijking met Sharepoint. De voor- en nadelen van het platform Liferay blijken niet alleen uit de ervaringen van de gebruikers, ze komen ook naar voren in de vergelijking met de ervaringen met Sharepoint. Deze vergelijking deden we ten eerste in gesprekken met de leidster en de deelnemers van de collaboratory rond Gilden, die aanvankelijk met Liferay werkte maar is overgegaan op SharePoint. Ten tweede konden we de omgeving en ervaringen van de Virtual Knowledge Studio Collaboratory in onze vergelijking meenemen. De resultaten zijn verwerkt in de vorm van een beknopt deelrapport (Bijlage 3).

Resultaat: Vanwege de verschuivingen in de planning en uitvoering van WP2 is deliverable 2 (de enquête) later uitgevoerd en is deliverable 3 (de interviews) deels gecombineerd met de participerende observatie. Tevens is besloten om D2 en D3

niet als apart deelrapport te publiceren, maar te verwerken in het eindverslag van dit werkpakket (zie bijlage 2).

Deliverable 1: verslag van literatuurstudie: “Collaboratories: from natural sciences to social sciences and humanities” (zie bijlage 1).

Deliverable 2: Enquête gebruikerservaringen (zie bijlage 2).

Deliverable 3: Diepte-interviews gebruikers (zie bijlage 2).

Deliverable 4: Deelrapport met de vergelijking Sharepoint en Liferay (zie bijlage 3).

Deliverable 5: Eindverslag, incl. samenvatting met (beleids)adviezen rond optimaal gebruik van virtuele platforms (zie bijlage 2).

WP2. Verhoging gebruikersvriendelijkheid

Doel: In het eindverslag van HubLab1 werden diverse problemen gesignaleerd wat betreft de gebruikersvriendelijkheid van de Liferay omgeving. De problemen circelden rond navigatie, logische naamgeving, sorteermogelijkheid van bestanden. Voor veel deelnemers bleken dergelijke problemen een grote drempel op te werpen, waardoor het bezoek aan de site al snel gestaakt werd. Vandaar dat het aanpakken van de ‘look and feel’ een grote prioriteit heeft gekregen. Het specifieke doel van dit werkpakket is de drempel voor het gebruik van Liferay voor zowel beheerders van collaboratories als leden omlaag te brengen.

Aanpak: Dit werkpakket draait om de optimalisatie van de omgeving waar het gaat om het snel wegwijs raken van beheerders en gebruikers. In dat kader zijn, in aanvulling op de inzichten uit Hublab1, in de periode januari-maart enkele beheerders en gebruikers geïnterviewd (in totaal vier personen).

Een belangrijk aandeel in de verhoging van de gebruikersvriendelijkheid is de installatie van een nieuwe versie 5.2 van de collaborative software Liferay. Deze nieuwe versie biedt niet alleen nieuwe functionaliteiten, maar ook veel meer mogelijkheden om de *look and feel* naar eigen inzicht in te richten. Tevens is het beheer van een community in de nieuwe versie sterk vereenvoudigd. Deze nieuwe mogelijkheden zijn in de periode januari-maart in kaart gebracht en getest in een daarvoor ingerichte testomgeving. Eind augustus is de nieuwe versie naar tevredenheid geïmplementeerd. De volgende stappen zijn in dat kader gezet:

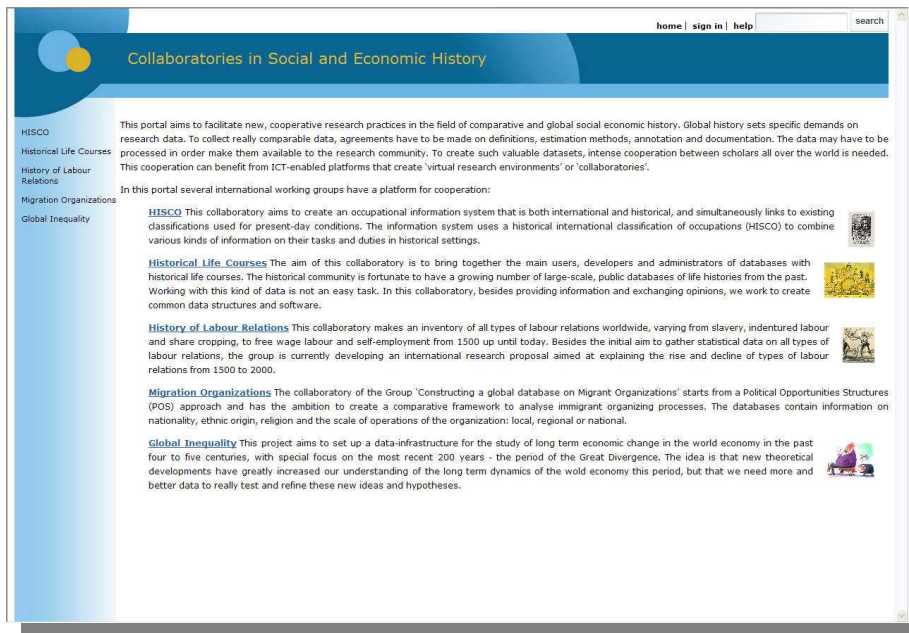
1. Installeren LifeRay trunk (source code) en compile omgeving opzetten: installatie van technische omgeving van nieuwe versie.

2. Ontwikkelen e-mail collaboratory portlet: in de nieuwe versie zat geen geschikte mailfunctie en die is naar aanleiding van de wens hiervoor van de gebruikers ontwikkeld (zie ook WP4).
3. Sorteren bestanden: toevoegen van naam en toevoegdatum in de library documents portlet om bestanden niet alleen alfabetisch te kunnen sorteren.
4. Sorteren forum threads: Nieuwe bijdragen aan threads staan momenteel onderaan. In de nieuwe versie zullen bijdragen anti-chronologisch worden weergegeven: dus, de meest recente bijdrage bovenaan.
5. Handleiding in wiki zetten: voor de handleiding van de 'community owners' en de eindgebruikers is een wiki ingericht.
6. Gebruik meten: installeren en configureren van AWStats plus unsecured webtoegang tot de webstatistiek pagina's zelf.
7. Chat functie aanpassen: de 'online friends'-functie is nu zo ingesteld dat alle leden van een collaboratory als default 'friends' van elkaar zijn. Tevens de term 'friends' vervangen door 'colleagues'.
8. Zichtbaarheid collabs: publieke sites van collabs zijn zichtbaarder gemaakt op het web, met name voor mensen die Google en aanverwante zoekprogramma's gebruiken.
9. Visuele stijl aanpassen: een nieuwe en rustigere stijl implementeren. De nieuwe indeling werkt met een standaard menu- en pagina-indeling. Een expliciet doel was een overzichtelijke lay out en een intuïtieve navigatie.
10. Nieuwe versie testen: De nieuwe portlet en aanpassingen aan bestaande community portlets zijn getest door derden.
11. Migreren collabs: in de periode 26-26 augustus zijn de verschillende laboratoreries gemigreerd naar de nieuwe stijl en versie (zie onderstaande screenshots).

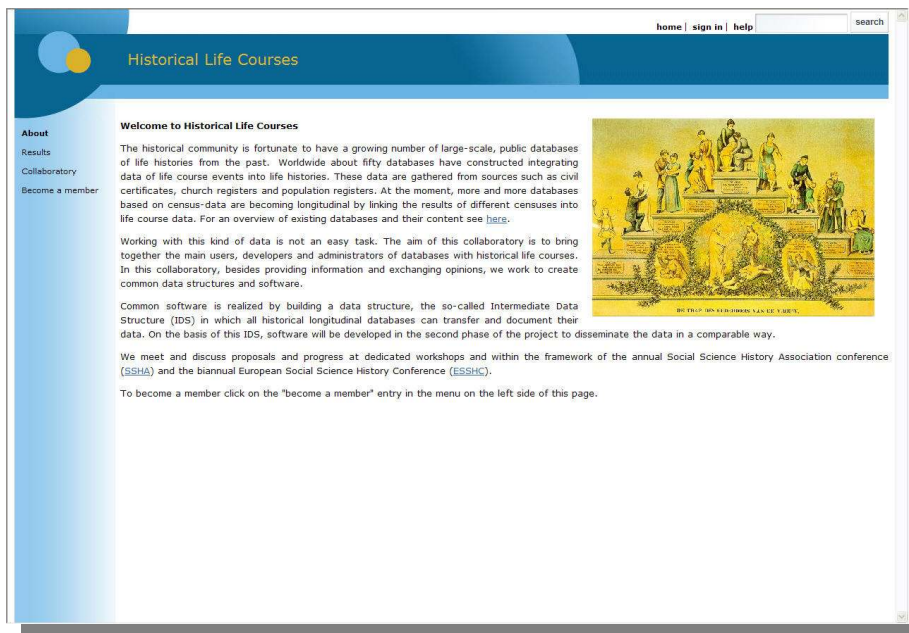
Resultaat: Deliverable 6: Een herschreven handleiding en een verbeterde 'look and feel' (interne deliverable).

Vanwege een interne reorganisatie bij de afdeling Digitale Infrastructuur (voorheen Digitale Projecten) van het IISG en door het onverwachte vertrek (naar SURF Foundation) van de informatie analist die de aanpassing van Liferay zou trekken, is dit werkpakket met duidelijke vertraging uitgevoerd. Uiteindelijk heeft de migratie naar de nieuwe versie plaatsgevonden in de periode 26-28 augustus 2009. De oorspronkelijke geplande demonstratievideo's zijn uiteindelijk niet zelf gemaakt.

Het is gebleken dat voldoende geschikt beeldmateriaal voorhanden is binnen de Liferay community die men via de help-wiki in de collaboratory-omgeving kan bereiken.



Figuur 2: Algemene portal voor Hublab Collaboratories



Figuur 3: Publieke portal voor de Historical Life Courses Collaboratory

WP3. Datasharing

Doel: Het project beoogt inzicht te bieden in de blokkades die bestaan, c.q. de incentives die gegeven kunnen worden aan (internationale) onderzoeksgroepen waarin data sharing centraal staat. Een van de mogelijkheden is het implementeren van open data licenties. Het uiteindelijke doel van dit werkpakket is duidelijkheid te scheppen naar zowel leden van collaboratories, beheerders van repositories en eventuele hergebruikers van de data over eigendomsrechten van gedeelde data.

Aanpak: Voor dit werkpakket zijn verschillende mogelijkheden op het terrein van open data licenties onderzocht. Op basis van dat onderzoek is in de eerste fase van dit project een beredeneerd voorstel gedaan voor een licentieregeling t.b.v. alle collaboratories. Dit voorstel is op 27 mei besproken met de diverse betrokkenen bij het IISG (de trekkers van de diverse collaboratories en vertegenwoordigers van de onderzoeksafdeling en de afdeling digitale projecten). Afgesproken is toen om de ontwikkelingen rond open data licenties te volgen en op basis van het gemaakte voorstel een uitwerking te maken die door de diverse collaboratories gebruikt kan worden (zie bijlage 4).

Eén van de implicaties van het niet helder maken van eigendomsrechten van gedeelde data is dat duurzame opslag moeilijk te realiseren valt. Immers, waarom zou een (data) archief tijd, geld en moeite steken in de verduurzaming van een dataverzameling als niet duidelijk is wie/wat die data in goede staat overdraagt? De vraag naar criteria voor duurzame opslag van digitale onderzoeksdata staat centraal in het SURFShare programma 2007-2010, met name Werkpakket 7, Datacuratie en digitale duurzaamheid. Het IISG heeft vanuit het perspectief van data verzameld binnen collaboratories een aantal richtlijnen opgesteld, die inmiddels zijn verschenen onder de titel *IISH Guidelines for preserving research data: a framework for preserving collaborative data collections for future research* (januari 2010). Hierin wordt een handreiking geboden aan collaboratories rond de voorbereiding van de overdracht van data aan een data-archief, waarin uiteraard afspraken over de eigendomsrechten een eerste stap vormen.

Resultaat: Deliverable 7. Voorstel licentieregeling: “Collaboratories: licentieregeling voor het delen van data” (zie bijlage 4). Extra resultaat. A. Doek, L. Heerma van Voss, K. Hofmeester, J. Kok en T. van der Werf-Davelaar, *IISH Guidelines for preserving research data: a framework for preserving collaborative data collections for future research* (Januari 2010).

WP4. Verbetering communicatie

Doel: Het mogelijk maken van uitgaande email vanuit de Liferay omgeving.

Aanpak: In de vorige versie ontbrak de mogelijkheid om de leden mailberichten te sturen vanuit het Liferay collaboratory platform. Dit betekende in de praktijk dat de beheerders van de collaboratories mailinglijsten opgesteld binnen externe mailprogramma's bleven gebruiken. In dit WP is een component gebouwd die met behulp van de persoonlijke gegevens van onderzoekers in het backend van Liferay mail berichten naar de leden van het collaboratory kan versturen.

Resultaat: Deliverable 8. Portlet voor uitgaande mail. Tijdens de migratie naar de nieuwe versie van Liferay is deze nieuwe functionaliteit naar tevredenheid in werking gesteld.

WP5. Ondersteuning workflows

Doel: Het vergemakkelijken van het werken aan een centrale database.

Aanpak: In dit werkpakket zou binnen Liferay een portlet ontwikkeld worden die toegang verschaft tot één enkele database en de daarbij behorende invoerschermen. Deze database bevindt zich, in technische zin, niet binnen het collaboratory. De database structuur en de invoerschermen zouden door de beheerders met een pakket worden ontwikkeld dat speciaal daarop is toegesneden.

Resultaat: Deliverable 9 (Een portlet voor de communicatie met webformulieren en de achterliggende databases) is uiteindelijk niet gerealiseerd. Het bouwen ervan hing af van de bereidwilligheid van een collaboratory trekker om een pilot te helpen ontwikkelen en te testen. Dit bleek niet haalbaar, ook en vooral omdat het nut in de praktijk ter discussie stond: de meeste mensen voeren data offline in.

WP6. Interoperabiliteit

Doel: Verruiming van de onderzoeksmogelijkheden binnen de collaboratory omgevingen door integratie van alle platforms in het centrale zoekstelsel van het IISG.

Aanpak: Aanvankelijk leek de ombouw van het 'digitale huis' van het IISG de realisering van dit (nog op de oude zoektechniek gebaseerde) werkpakket in de weg te staan (zie tussentijds verslag september 2009). In het najaar werd echter besloten gebruik te gaan maken van de nieuwe techniek SOLR. Wanneer HubLab (en andere) datasets in de nieuwe IISG index applicatie worden geladen (Solr based) zijn deze automatisch voor de wereld toegankelijk te maken. Dat is mogelijk met het SRU protocol. Deze is toegankelijk via api.iisg.nl/srw. SRU is een algemeen bekend zoekprotocol. Liferay kan gebruik maken van SRU via een web portlet. De onzichtbare kant maakt gebruik van SRU. De zichtbare kant presenteert internet

bezoekers met een zoekscherm, resultaatset en view van een record (gebaseerd op Dublin core standaard). Omdat het gaat om een portlet dat gebruik maakt van een protocol SRU is het product inzetbaar voor alle web portals. Tegelijkertijd zal het open sourcen van het te ontwikkelen product derde partijen in staat stellen de portlet te hergebruiken of functionaliteit toe te voegen.

Resultaat: Deliverable 10. Een portlet in Liferay dat bestanden doorzoekbaar maakt. De volgende stap wordt het IISG op zich genomen. In de eerste plaats worden ook andere collecties ontsloten en in de tweede plaats wordt het portlet omgebouwd tot een *widget*, waarmee op alle websites direct zoekacties in de IISG collecties zijn uit te voeren. De broncode hiervan wordt momenteel open source aangeboden: http://svn.xp-dev.com/svn/iisg_portlets/.

Vanwege de late oplevering (december 2009) kon niet meer onderzocht worden in welke mate deze functionaliteit voldoet aan de wensen en welke verbeteringen verder zouden moeten worden aangebracht.

WP7. Kennisdisseminatie

Doel: Verschillende wetenschapsgebieden op de hoogte stellen van de mogelijkheden van de vorming van collaboratories in de (sociale) geschiedenis. In de eerste plaats gaat het daarbij om (historische) onderzoeksgemeenschap, in de tweede plaats om het e-research onderzoek zoals beoefend in de wetenschapsdynamica.

Aanpak: In dit WP zou in de eerste plaats een project website worden ingericht voor de Hublab documentatie. Uiteindelijk is besloten geen aparte website te bouwen voor HubLab maar de mogelijkheden van het collaboratory model en van Liferay te presenteren binnen de website van Global History Hubs (www.clio-infra.eu). In de rapportageperiode zijn verder o.a. twee bijdragen aan de “5th International Conference on e-Social Science” over collaboratories gepresenteerd in Keulen, is een artikel geschreven voor het tijdschrift *Historical Methods*, geparticipeerd in een hoofdstuk over *collaborative work* in een edited volume voor MIT Press, en is een bijeenkomst voor betrokkenen georganiseerd over de nieuwe Liferay versie. Zie verder sectie 3 over disseminatie.

Resultaat: Deliverable 11 (deels gerealiseerd): berichtgeving via discussielijsten, vakbladen, conferenties en workshops.

WP8. Helpdesk

Doel: Zowel de administrators als gewone leden van collaboratories bijstaan in het oplossen van problemen met het platform, en eventueel samen met de Liferay community werken aan bugs en gewenste extra functionaliteiten, voor zover niet genoemd onder WP2, WP4, WP5, en WP6.

Aanpak: De Liferay ontwikkelaars zullen alert reageren op enerzijds de meldingen en vragen die via het algemene leden platform in Liferay binnenkomen, anderzijds op de rapportages van de gebruikersonderzoeker (WP1).

Resultaat: De meeste vragen en problemen zijn online verzameld (via het IISG Support portaal) en adequaat behandeld. Een storende bug in de nieuwe versie van Liferay is nog niet structureel verholpen, waardoor de proxy server met de collaboratory documenten regelmatig onbereikbaar is. Weliswaar is een *workaround* ontwikkeld (dagelijkse automatische herstart server), het is toch zorgwekkend dat de Liferay community niet in staat is gebleken met een werkbare oplossing te komen.

Op 26 november 2009 is een hands-on workshop voor de gebruikers (in het bijzonder de collaboratory administrators) georganiseerd, waarin de voornaamste veranderingen aan het platform zijn doorgenomen, wensen voor verdere verbeteringen zijn besproken.

WP9. Projectmanagement

Doel: Het project aansturen op het voltooien van de beoogde resultaten binnen de geplande projectduur.

Aanpak: De projectleider heeft duidelijke afspraken gemaakt over taken, benodigde inzet, verantwoordelijkheden, overlegfrequentie, communicatie, planning en rapportage conform het projectplan (controlling document).

Resultaat: Het project was een vast onderdeel op de vergaderingen van de stuurgroep Global Hubs, die in 2009 vier keer bijeen is geweest.

3. Kennisdisseminatie en –uitwisseling

De externe informatie uitwisselingen met de SURFshare tender deelnemers vond plaats tijdens de bijeenkomsten, die SURF organiseert ten behoeve van uitwisselingen en afstemming, en per mail, hetgeen gezien de omvang van de

tender (3 groepen) geen probleem is. Tevens is er intensief contact met de VKSC groep uit de tender, vanwege de institutionele affiliatie.

De nieuwe versie van Liferay is op 27 mei gepresenteerd aan alle belanghebbenden. In eerste instantie zijn toen de historici die als beheerder met de omgevingen werken intensief voorgelicht over de (veranderingen aan) de omgevingen. Vanwege de vertragingen rond WP2 is het uiteindelijke overzetten naar een nieuwe versie pas op 28 augustus gebeurd. Tijdens de pre-conference workshop van de Historical Life Courses Collaboratory in Long Beach, Ca. (11 tot en met 15 november) is tevens een introductie gegeven over de nieuwe versie van Liferay voor de leden van de betreffende collaboratory. Op 26 november 2009 is een hands-on workshop voor de collaboratory administrators georganiseerd, waarin de voornaamste veranderingen aan het platform zijn doorgenomen en wensen voor verdere verbeteringen zijn besproken.

De literatuurstudie uit WP1 vormde een belangrijke basis voor de paper "Uncovering Collaboratories in Social and Economic History" (zie bijlage 5) die gepresenteerd is op de internationale e-social science conferentie in Keulen (24-26 juni 2009). De paper is uiteindelijk ook opgenomen in de proceedings van deze conferentie, die werd georganiseerd door het National Centre for E-Social Science, in samenwerking met GESIS (German Social Science Infrastructure Services). Tijdens de conferentie is ook een presentatie gegeven onder de noemer "Collaboratories in social sciences and humanities" tijdens de workshop "Imagining, Identifying, and Investigating e-Research". Tevens zijn ook een tweetal wetenschappelijke publicaties geschreven. Het artikel "An alternative approach to large historical databases. Exploring best practices with collaboratories" is begin december ingediend bij het toonaangevende tijdschrift *Historical Methods* en richt zich specifiek op het verslag doen van onze praktijkervaringen met collaboratories voor een publiek van historici. De tweede publicatie betreft een hoofdstuk voor een edited volume dat medio januari ingediend is bij de editors van de "Inside Technology"-reeks van MIT Press. In dit hoofdstuk, getiteld "Working in virtual knowledge: Affective labour in scholarly collaboration", dient het veldwerk over collaboratories als empirisch materiaal voor een betoog over de veranderende rol van *affective labour* in academische samenwerking.

De Liferay omgeving wordt gebruikt door een nieuwe IISG collaboratory 'Fighting for a living' waarin ongeveer 18 experts in militaire geschiedenis van over de hele wereld samenwerken. De omgeving is voor deze groep ingericht (<https://projects.iisg.nl>). Meer in het algemeen blijft het IISG bij het voornemen het Liferay platform verder te ontwikkelen, onder andere in het kader van geplande dataverzamelingsprojecten (www.clio-infra.eu). Het IISG heeft ook een functioneel beheerder aangesteld, die het platform na de projectfase overneemt.

Een deel van de opgebouwde kennis over het Liferay programma is intussen ook reeds aangeboden aan externen in de vorm van enkele ontwikkelde functionaliteiten: de SOLR portlet (WP6), mailportlet, het sorteren op auteursnamen en datum bij documenten, en sortering van de recent threads message boards (de punten 2 t/m 4 van WP2). Europeana (www.europeana.eu) van de EDL stichting (circa 300 gebruikers begin 2010) heeft deze nieuwe mailfunctionaliteit in haar Liferay platform geïntegreerd.

De literatuurstudie van WP1, alsmede de *draft version* van het artikel voor *Historical Methods* zijn intussen reeds gebruikt als onderwijsmateriaal door Niki Vermeulen in de Mastercursus "Collaboration and competition: on forms and formats of being together in science" en door Rick Duque in de PhD reading class. Beide cursussen zijn gegeven aan het Department for Social Studies, Universität Wien.

Een deel van de kennis opgedaan in het kader van HubLab2 kon daarnaast worden gedeeld met andere SURF projecten, zoals het al besproken WP7 *Datacuratie en digitale duurzaamheid* van het SurfShare programma. Ook is uitgebreid overlegd met mevrouw Lilian van der Vaart in het kader van haar Collaboratory Study voor SurfFoundation en SURFnet.

4. Evaluatie

Het project HubLab2 had een dubbel maar geïntegreerd doel: enerzijds de potentiële bijdrage van een *Virtual Research Environment* (VRE) voor collaboratories zo nauwkeurig mogelijk in kaart brengen, anderzijds de wensen binnen collaboratories (van sociaal-economisch historici) ten aanzien van het VRE implementeren in Liferay-omgeving. De verschillende werkzaamheden zijn uitgewerkt in 9 werkpakketten, met bij elkaar 12 *deliverables*.

De eerste fase van het project behelsde de gedeeltelijke herbouw van de Liferay-omgeving op basis van de wensen geformuleerd aan het einde van het HubLab1-project. Het lag voor de hand om deze herbouw te koppelen aan het *upgraden* naar versie 5.2 van Liferay, waarin al een aantal gewenste functionaliteiten waren aangekondigd. Echter, zowel de oplevering van versie 5.2 door Liferay, het onderzoek door IISG medewerkers naar de mogelijkheden van aanpassing ervan, als het daadwerkelijk inbouwen van de nieuwe functionaliteiten nam meer tijd in beslag dan gepland. Daardoor kwam ook fase 2, het gebruikersonderzoek in het gedrang. Uiteindelijk kon dat onderzoek toch worden afgerond. Een laatste fase, het op grond van het gebruikersonderzoek doorvoeren van een nieuwe ronde verbeteringen kon echter niet meer worden afgerond.

De aanpassingen van het tijdschema zijn tijdig aan SURF voorgelegd en in orde bevonden. In het onderstaande schema ‘Standlijnen en planning’ wordt een overzicht gegeven van de geplande en uitgevoerde werkzaamheden. Bij de bespreking van de afzonderlijke werkpakketten is al aangegeven wat er inhoudelijk is gerealiseerd.

Inhoudelijk

De resultaten kunnen omschreven worden in termen van een paradox. We hebben naar onze mening een qua functionaliteit volledig, gebruikersvriendelijk en krachtig Virtual Research Environment opgebouwd, dat echter maar zeer beperkt een rol speelt in de feitelijke praktijk van collaboratories van historici. Voor een deel kan dit liggen aan eigenschappen van historici in het bijzonder (en geestes- en sociale wetenschappers in het algemeen). Het gaat daarbij om een aversie tegen nieuwe software (zelfs bij enkele trekkers), een geringe vertrouwdheid met groepswork en een nadruk op context-specifieke wetenschappelijke informatie (de inhoud van begrippen hangt vaak af van plaats en tijd) – wat een extra drempel opwerpt voor het online delen van kennis. De rol van *face-to-face* ontmoetingen in workshops blijft dan ook groot. Voor een ander deel kan dit liggen aan de (door ons onderzochte) collaboratories, die zich kenmerken door een hoge mate van vrijwilligheid en gelijkheid. In tegenstelling tot bijvoorbeeld het werk van assistenten in virtuele laboratoria kan deelname aan historische VRE’s niet worden ‘afgedwongen’ door de leiding, voor zover er al van leiding in hiërarchische zin kan worden gesproken.

De conclusie moet zijn dat de verwachtingen ten aanzien van het gebruik van collaborative software door gedistribueerde onderzoekers in de geesteswetenschappen naar beneden moeten worden bijgesteld. Het geringe gebruik als communicatiemiddel kan volgens ons niet worden toegeschreven aan de ‘leercurve’ die elke nieuwe software kent. Die leercurve is met de nieuwe omgeving drastisch beperkt. Voorlopig lijken de platforms vooral te dienen als een handige online *repository* van de groepsdocumenten en de centrale database. Onderlinge discussies spelen zich buiten het forum af, en worden daarmee niet of onvoldoende gedocumenteerd. De communicatiemiddelen binnen het platform vervangen nu eenmaal niet de bestaande, vertrouwde middelen. Het vereist een extra en langdurige inspanning van de groepsleiders om nieuwe communicatiemiddelen ingang te doen vinden in de groep, en dat lijkt bij de onderzochte collaboratories een brug te ver.

De *beleidsmatige* verwachtingen ten aanzien van collaboratories blijven echter onveranderd hoog. Dit kunnen we opmaken uit tal van nieuwe projecten in binnen- en buitenland, waarin grote groepen (professionele en amateurs) historici online data gaan delen en verrijken. Het geldt zeker ook voor het beleid van het

Internationaal Instituut voor Sociale Geschiedenis waarvoor de gegevens verzameld in collaboratories een speerpunt vormen in het meerjaren onderzoeksprogramma 'Global Labor History'. Tenslotte geldt het voor economisch historici die wereldwijde dataverzameling een nieuwe impuls willen geven met collaboratories (www.clio-infra.eu). Er zal dus ook in de toekomst tijd en geld worden worden geïnvesteerd in de bestaande en nog op te richten collaboratories. Verdere verbetering van de omgeving zal daarvan een onderdeel zijn.

De in Hublab2 gebouwde omgeving zal een centrale rol blijven spelen in bestaande en nog op te bouwen collaboratories. Omdat de Liferay techniek ook op andere onderdelen van het Digitale Huis van het IISG wordt ingezet (bv inlog procedures), zal blijven worden geïnvesteerd in kennis van het pakket. In die zin lijkt duurzaamheid en implementatie gegarandeerd. De notitie over datasharing speelt een rol in de organisatie van toekomstige collaboratories. In de recente notitie (in het kader van SurfShare) van het IISG over criteria voor duurzame opslag van digitale data wordt voorgeschreven dat duidelijkheid over eigendomsrechten voorafgaat aan het curatietraject. Tenslotte zijn diverse in het kader van HubLab2 gebouwde onderdelen (portlets) als open source aangeboden en geïmplementeerd in andere platforms, zoals dat van Europeana.

Procesmatig

Het project HubLab2 is doelbewust opgezet als een iteratief proces, dat wil zeggen dat er continu uitwisseling werd voorzien tussen het gebruikersonderzoek enerzijds, de ontwikkeling aan het platform anderzijds. In de praktijk is dit toch te ambitieus gebleken, in de eerste plaats omdat er lang gewacht moest worden op de nieuwste release van Liferay en in de tweede plaats omdat de afdeling Digitale Projecten van het IISG menskracht te kort kwam, onder andere door een interne reorganisatie en het vertrek van de voornaamste data-analist (naar SURF). Het gevolg was dat de 'iteraties' beperkt zijn gebleven tot anderhalve ronde: 1) verzamelen wensen t.a.v. de nieuwste versie van Liferay; 2) implementeren van die wensen en 3) gebruikersonderzoek met aanbeveling voor een nieuwe ronde aanpassingen.

Terugblikkend kunnen we stellen dat onze opzet van snelle implementatie en een opeenvolging van feedback rondes te optimistisch is geweest. Het vertrouwen op een open source product als Liferay is toch niet zonder risico's gebleken. Met name blijken nieuwe releases niet grondig getest te zijn, waardoor een deel van het testen, verzamelen van bugs en het aandragen van oplossingen door de community gedragen moet worden. Dit levert natuurlijk risico's op voor een project als Hublab2. Immers, elke bug levert flink verlies aan *goodwill* op bij de toch al terughoudende collaboratories van historici.

Standlijnen en planning

Planning Hublab2 (versie januari 2010)		januari	februari	maart	april	mei	juni	juli	augustus	september	oktober	november	december
WP	Activiteit												
WP1 Gebruikersonderzoek													
d1	verslag lit studie												
d2	enquete				verplaatst >								
d3	diepteinterviews							verplaatst >					
d4	Liferay vs. Sharepoint										verplaatst >		
d5	eindverslag												
WP2 Verhoging gebruikersvriendelijkheid													
d6a	look and feel	verplaatst >											
d6b	handleiding	verplaatst >											
d6c	demonstratievideos	vervallen											
d6d	Helpdesk												
WP3 Datasharing													
d7a	voorstel licentieregeling												
d7b	implementatie							verplaatst >					
WP4 Verbetering communicatie													
d8	integratie email	verplaatst >											
WP5 Ondersteuning workflows													
d9	Tool online werken aan database	vervallen											
WP6 Interoperabiliteit													
d10	Tool voor integratie Search/Lucene							verplaatst >					
WP7 Disseminatie													
d11	webpagina	vervallen											
WP8 Helpdesk													
WP9 Management													
d12	Periodieke projectrapportages							vervallen					

5. Financiële verantwoording

De uiteindelijke projectkosten, inclusief de (geschatte) accountantscontrole en onkosten gemaakt voor het bezoeken van workshops, bedragen 80.332 euro, ruim boven de begrote 71.350 euro uit het controlling document. Het nog in het tussentijds verslag (september 2009) genoemde probleem van onderbesteding is dus in het laatste kwartaal opgelost, en wel door intensivering van het gebruikersonderzoek en de 'nazorg' bij de oplevering van het nieuwe platform , o.a de helpdesk, een hands-on workshop, de handleiding. Tenslotte moet genoemd worden de externe uitbesteding van Werkpakket 6, de SRU portlet voor doorzoekbaarheid van de bestanden binnen de collaborative omgeving.

Bijlagen

Bijlage 1: verslag van literatuurstudie (deliverable 1).

Bijlage 2: eindverslag gebruikersonderzoek (deliverable 2, 3 en 5)

Bijlage 3: vergelijking Sharepoint en Liferay (deliverable 4)

Bijlage 4: voorstel licentieregeling (deliverable 7)

Bijlage 5: NCESS paper (WP7: kennisdisseminatie)

Introduction

Two recently published books on laboratories present case studies in astronomy, space science, physical sciences, biological and health sciences, and earth and environmental sciences (Olson et al., 2008; Shrum et al., 2007). The absence of case studies on laboratories in the social sciences and humanities (SS&H) in both books reflects a widespread bias towards the natural sciences in the literature on laboratories. Accordingly, this literature fails to adequately address the specificities of laboratories in SS&H and the particular challenges facing collaborators in these academic fields. In part, this bias can be explained by the fact that there are substantially more laboratories in the natural sciences compared to SS&H and that these laboratories also have a longer tradition (cf. Finholt, 2003 for an overview). Moreover, funding and support for these laboratories is more substantial than for laboratories in SS&H. However, the recent growth of laboratories in social sciences and humanities has increased the demand for knowledge about the specificities of academic collaboration at a distance in SS&H. This systematic exploration of the literature on laboratories aims to present the current state of the field. At the same time, it sets out to focus on those insights which are most relevant for laboratories in SS&H.²

More academic collaboration

Modern science is increasingly collaborative, Olson et al. (2008: 1) argue. They discuss several patterns to substantiate this claim, like increasing coauthorship and increasing international collaboration. According to Olson et al. (2008: 1), several key factors lie behind these patterns: the urgency, complexity, and scope of unsolved scientific problems; the need for access to new, and often expensive, research instruments and technologies; pressure from funding agencies; and information and communication technologies (ICTs) that facilitate interaction and sharing all play a role in prompting scientists to cooperate with individuals both within and outside their disciplines and institutions. They elaborate on these factors as follows:

Collaborative research makes it possible to tackle research questions that would otherwise not be feasible to address (Thagard 1997; Wray 2002). Researchers work together because there are questions they want to investigate that they cannot undertake alone. In addition, funding agencies, which must respond to the needs of society and the political environment, have encouraged collaborative research. (Olson et al., 2008: 2)

Scientific collaboration, Sonnenwald (2007: 645) argues, can be defined as interaction taking place within a social context among two or more scientists that facilitates the sharing of meaning and completion of tasks with respect to a mutually shared, superordinate goal. Scientists who collaborate may also bring additional, individual goals to a collaboration (Sonnenwald, 2003). Sonnenwald (2007: 643) moreover argues that collaboration has the potential to solve complex scientific problems and promote various political, economic, and social agendas, such as democracy, sustainable development, and cultural understanding and integration. Although it is difficult to deny that collaboration has this *potential*, many of these examples are, in my view, not at the heart of everyday collaborative practices. In contrast, Maienschein's threefold classification of collaborations

² In the following stage of the hublab2-project, this review will serve as theoretical input for a survey among the members of various laboratories in one specific subdiscipline in SS&H: social and economic history.

based on the reasons for collaborating (as cited in Chompalov and Shrum, 1999: 340) seems to be more in tune with mundane collaborative research practices: to promote an efficient division of labour, to enhance credibility, and to build community.

Several authors have argued that the recent increase in academic collaboration is partly attributable to funding regimes (e.g. Cummings and Kiesler, 2005; Hine, 2006; Sonnenwald, 2007). Public and private research funding agencies, such as the National Science Foundation (NSF) in the USA and the Framework Programmes in the European Union, aim for diverse organizational representation and require interdisciplinary, international, and inter-institutional collaboration. Some authors, like Cummings and Kiesler (2005: 704), have argued that dispersed collaborations are more feasible because communication technologies allow scientists to exchange news, data, reports, equipment, instruments, and other resources (Hesse et al., 1993; Kouzes et al., 1996; Finholt, 2002).

In addition to these structural changes with regard to funding regimes and the development of information and communication technologies, Sonnenwald (2007: 650) mentions the opportunity to discover new knowledge and solve complex problems in a timely manner as an important motivation for many academics to consider collaborating. A point which is also raised by Finholt (2003: 8). He argues that greater size and complexity of research tasks is reflected in a higher need for collaboration and that the trend towards collaboration is increasing and appears to be independent of discipline.

Academic collaboration at a distance

Traditionally, academic collaboration entails co-location. Although co-location has many advantages and is even essential for carrying out specific research tasks, there are additional collective costs of co-location (Finholt, 2003: 7). Co-location can play into competitive rivalries among research sites and thus undermine cooperation. Lack of cooperation can subsequently lead to redundant capacity, such as several independent and under-utilized instruments, rather than a single shared and fully utilized instrument. More important, barriers to interaction across sites may slow integration of knowledge required to resolve research questions that exceed the capacity of single sites, or even single disciplines, to answer. According to Finholt, various contemporary large research problems require unprecedented levels of cooperation from communities that have, in the past, worked independently. Although this seems to be first and foremost the case in the natural sciences, I agree with Finholt that this approach relies heavily on information technology to overcome barriers of time and space and – even more importantly so – it suggests a dramatic revision of the historical organization of science.

Communication is unquestionably important in fostering and sustaining successful scientific collaborations. Yet, in the case of long distance collaboration, communication alone – even in an era of electronic mail and the Web – may not be enough to enable a broad range of collaborative activity. In particular, geographically distributed scientists may also need support for data sharing and data exploration, as well as tools that allow remote use of important instruments and facilities. (Finholt, 2003: 5)

Accordingly, a collaboratory can be understood as a combination of enhanced access to data and instruments with improved communication tools. Finholt refers in this respect to the often quoted definition by Wulf, who states that a collaboratory is ‘...a center without walls, in which researchers can perform their research without regard to physical location – interacting with colleagues, accessing instrumentation, sharing data and computational resources, and accessing information in digital libraries’ (Wulf, 1989: 19). The term collaboratory is a hybrid of ‘collaborate’ and ‘laboratory’. Hence, elaborations of the collaboratory concept stress the simultaneous need to solve problems of control and operation of instrumentation over the Internet, of access and distribution of datasets, and of convenient and flexible interaction with colleagues (Finholt, 2003: 5).

A central player in the research on collaboratories is the Science of Collaboratories (SOC) project from the University of Michigan. This project aimed, among other things, to provide the vocabulary, associated principles, and design methods for propagating and sustaining collaboratories across a wide range of circumstances (Olson et al, 2008). In one of the SOC publications, a broad definition of collaboratories is given:

A collaboratory is an organizational entity that spans distance, supports rich and recurring human interaction oriented to a common research area, and fosters contact between researchers who are both known and unknown to each other, and provides access to data sources, artifacts, and tools required to accomplish research tasks. (Bos *et al.*, 2007: 656)

In the same publication, a taxonomy based on a large sample of collaboratories is presented. Seven types of collaboratories are identified and listed in the table below. According to Bos et al. (2007: 668), the field of collaboratories has been observed to move from the top left to the bottom right over time. This implies that the ‘distributed research center’ is considered to be the most challenging form of academic collaboration.

	Tools (instruments)	Information (data)	Knowledge (new findings)
Aggregating across distance (loose coupling, often asynchronously)	Shared Instrument	Community Data System	Virtual Learning Community & Virtual Community of Practice
Co-creating across distance (requires tighter coupling, often synchronously)	Infrastructure	Open Community Contribution System	Distributed Research Center

Here, it should be noted that the SOC project focused exclusively on collaboratories in the natural sciences. This taxonomy thus does not reflect the range of practices in the social sciences and humanities. The collaboratories in social and economic history aim, for instance, to aggregate and co-create across distance and, although they share information, they ultimately focus on the sharing of knowledge and on new findings. Hence, they resemble the distributed research center and as such they require tighter and often synchronous coupling of activities.

According to Finholt (2003: 22), the critical element of collaboratories might be the opportunity they allow for encounters, discussions, and sharing of ideas. Thus, besides being a place where scientists

can access instruments and data independent of time and location (a technology-centric understanding), a collaboratory is also an organizational entity that supports rich and recurring interaction around a common research focus among researchers who are both known and unknown to each other (Finholt, 2003: 22). Sonnenwald (2007: 650) adds to this that scientific collaboration can also help extend the scope of a research project and foster innovation because additional expertise is made available (Beaver, 2001; Cummings & Kiesler, 2003; Lambert, 2003). Hence, it can increase scientific reliability and the probability of success because more than one person is considering the accuracy, quality, and meaning of the results (Beaver, 2001; Thagard, 1997).

Success and failure in collaboratories

If one advice for academic collaboration emerges from the literature on collaboratories, it is not to underestimate the challenges and problems accompanying such an enterprise. Academic collaboration is difficult to organise, it is difficult to sustain and there aren't that many successful examples around. As Bos et al. (2007: 653) argue, only a few of the collaboratories they studied have succeeded in sustaining long-distance participation, solving larger-scale problems, and initiating breakthrough science. Although not all collaboratories set out to achieve similarly ambitious goals, the general assessment in the literature is that academic collaboration is difficult and that collaboratories – because of their specific characteristics – come with additional difficulties. According to Finholt (2003: 18), the experience with collaboratories to date indicates above all the enormous difficulties of supporting complex group work in virtual settings. For him, overcoming these difficulties is the great challenge for the next stage of collaboratory development and use.

In general, three types of barriers can be identified that make scientific collaboration difficult (Bos et al., 2007: 653-654). First, scientific knowledge is difficult to aggregate. While *information* has become very easy to transmit and store over great distances, *knowledge* is still difficult to transfer. Second, scientists work independently. Scientists generally enjoy a high degree of independence, both in their day-to-day work practices as well as in the larger directions of their work. The culture of independence benefits science in many ways, but it also makes it more difficult to aggregate scientists' labour. The third barrier is the difficulty of cross-institutional work. Even when all of the scientists are ready to proceed, collaborations can run into institutional-related problems, especially legal issues, that cannot be resolved.

Although not all collaborative research is cross-institutional, or covers great distances, it always involves multiple scientists and some of these may have different disciplinary backgrounds, work in different institutions, and are not collocated. Hence, additional planning and sufficient time for planning are required. As Olson et al. (2008: 2) argue, collaborations that involve geographically dispersed participants present some additional social and organizational challenges. In particular, they have a higher likelihood of failure or underperformance (Olson and Olson, 2000; Cummings and Kiesler, 2005). In these situations it is more difficult to align goals and incentives, establish common ground, engender and maintain trust, allow for the costs of coordination and communication, and determine an appropriate division of labour and resources (Olson et al., 2008: 2). The literature moreover suggests that research vision, goals, and tasks; leadership and organizational structure; use of ICT; and intellectual property and other legal issues need to be considered in greater detail than in single investigator research (Sonnenwald, 2007: 656-657). In the following pages, many of

these points will be discussed in more detail. In these discussions, it is often implicitly assumed or explicitly stated that distance still matters (Olson and Olson, 2000).³

The loss of common ground

Possibly the most significant barrier to both the design and use of collaboratories is that most group practices and routines assume a shared space (Finholt, 2003: 19). The loss of a common physical setting in collaboratories changes the circumstances for collaboration and may undermine the effectiveness of the collaborative process (Finholt, 2003: 19). According to Bos et al. (2007: 654), distance collaboration challenges coordination and trust building (Jarvenpaa & Leidner, 1999), fosters misunderstandings (Cramton, 2001) and inhibits communication of tacit knowledge (Lawson & Lorenz, 1999) and transactive knowledge, or knowledge of what colleagues know (Hollingshead, 1998). Hence, in the virtual context collaborators must be explicit about information that is normally tacit when co-located. Otherwise, the loss of tacit cues could also result in losing common ground (Clark and Brennan, 1991), where common ground is the shared cognitive understanding that allows collaborators to successfully coordinate their effort to accomplish joint work. At a minimum, then, collaboratories may require more effort, in terms of communicating the additional information required to achieve common ground (Finholt, 2003: 19).

However, more communication does not seem to be enough to overcome the problems discussed above. For some tasks, co-location remains essential and this is, according to Finholt (2003: 20) specifically the case when tasks are tightly coupled, that is, dependent on frequent interaction and feedback among collaborators. In those instances, contemporary communication technologies – such as e-mail, video and audio conferencing, and groupware – do not provide an adequate substitute for co-location (Olson and Teasley, 1996).

Task division and organisation

As mentioned, working together at a distance requires additional coordination and management of the various tasks involved. Obviously, the main tasks in an academic collaboration are research tasks and these can be divisible or conjunctive (Sonnenwald, 2007: 646). Conjunctive tasks require everyone to complete it, like the joint development of data collection instruments, or developing a common codebook for the development of collective datasets. Divisible tasks, on the other hand, can be executed individually and these can be either performed sequentially or concurrently. An example of a divisible task in social science is the separate data collection of the individual members of a group, for instance from similar populations in different regions.

As Cummings and Kiesler (2005: 704-705) argue, a major challenge for dispersed scientific collaborations is coordinating work so that scientists can effectively use one another's ideas and expertise without frequent face-to-face interaction. Although some coordination can be accomplished through project structure, for example, by creating clear lines of authority and division of labour, science is dynamic, and members of the collaboration still must talk out common

³ The specific issue of intellectual property is discussed more elaborately in a proposal on how to deal with copyrights in relation to databases. This proposal is part of the hublab2-project and will be published as part of that project.

problems, discuss shared resources, and monitor and review the work to make joint progress (Malone & Crowston, 1994; Kraut & Streeter, 1995). In part, such discussions can be facilitated by collaboration software, but in many cases face-to-face meetings will be required to organise discussions, or peer review, etc.

Cummings and Kiesler (2005: 717) also observe how communication tends to fall off as the dispersed investigators discover it is easier to work on their own tasks, rather than try to work together. As Hine (2006: 285) found out, dividing the tasks and making most parts of the actual work solitary, proves to be an effective means for organizing the work. With such a task division, tensions are largely being experienced at the level of individuals, and consequently do not threaten the success of the collaboration as a whole. However, handing over too much tasks and responsibilities to individual academics can lead to a diffusion of epistemic and ethical responsibility. When many scholars collaborate without some form of organisation and task division, no one scholar may feel responsible for the work (cf. Sonnenwald, 2007: 651).

Project management

In the literature, these issues are widely discussed and many authors see good project management as a crucial factor in the success of a collaboratory. However, it is important to stress once more that most collaboratories discussed in the literature are in natural sciences and, in general, these collaboratories tend to have more members and bigger budgets than their counterparts in the social sciences and humanities. Cummings and Kiesler's (2003) assessment that collaborative projects using more coordination and communication mechanisms have been found to be more successful, could therefore be less valid for the smaller collaboratories in SS&H. However, despite these reservations, the discussions on project management and task division in the literature provide some interesting ideas that deserve further scrutiny. For instance, Sonnenwald contends that research tasks should be clearly defined and owned by individual scientists:

This is particularly important when the collaboration occurs across distance (Maglaughlin & Sonnenwald, 2005; Olson et al., 2008). When scientists are collocated, they can informally observe and discuss task progress, but this is more difficult across distances and disciplines; therefore, defining tasks and task responsibility from the outset is important. (Sonnenwald, 2007: 657)

In general, Sonnenwald argues that organisational issues have to be dealt with at the beginning – or even before – the actual scholarly collaboration. It is, according to her, important to avoid misconceptions regarding the resources, including time, required to conduct various research tasks. Such misconceptions could lead to disagreement and conflict during the collaboration (Sonnenwald, 2007: 658). However, taking into account the dynamics of academic collaboration, one can question if it is feasible to organise this in advance. After all, some individuals may behave inappropriately by not honouring some aspect of the plan, failing to complete tasks, withholding needed information from their partners, or not sharing credit appropriately (Sonnenwald, 2007: 665). Besides, any group working together for a longer period will always be challenged by staff turnover or health and family problems among its members (Sonnenwald, 2007: 665). Therefore, successful project management always needs to be prepared for changing circumstances. Some additional best practices include:

sharing information about budgets (Cohen, 2000); ensuring that everyone receives benefits from the collaboration (Olson et al., 2008); developing a shared statement of principles, expected benefits, and mutual obligations (Cohen, 2000); formulating a shared vocabulary (Olson et al., 2008); ensuring that differences in resources are accounted for and aligned (Adessa & Sonnenwald, 2003); and establishing community and scientific advisory boards (Secrest et al., 2004; Sonnenwald, 2003b). (Sonnenwald, 2007: 659)

Depending on the kind of collaboration, some of these issues could be managed by good leadership. Sonnenwald (2007: 659) refers to various studies of successful collaboration which show that scientific, financial, and administrative leadership is important for success (Olson et al., 2008; Schiff, 2002; Stokols et al., 2005). To make the leadership manageable, she furthermore proposes to share the various tasks among several individuals to take advantage of individual strengths. Hence, setting up a project team could provide the required leadership in the collaboratory and it could also help to ensure that no one scientist is overburdened by leadership responsibilities. Nonetheless, extra coordination efforts are often not covered in the budget and many funding agencies do not recognise the actual costs incurred (Cummings and Kiesler, 2005: 717).

A last point with respect to project management in collaboratories mentioned by Sonnenwald, is that, because they are ambitious, visions and goals of collaboratories may require buy-in from other stakeholders as well, including participating institutions, one or more funding agencies, and citizen and community groups. Articulating clear visions and goals that multiple individuals and groups can understand and support is a skill scientists need when initiating large and complex scientific collaborations (Sonnenwald, 2007: 657).

Collaboration readiness

Notwithstanding the importance attached to organisational issues in the literature, good project management and an efficient division of tasks are not the only factors for the success of a collaboratory. Many critical barriers to successful virtual collaboration remain. Olson et al. (2000) characterize these additional barriers in terms of collaboration readiness and collaboration technology readiness (cf. Finholt, 2003: 20). The first term will be discussed in this section, while the latter will be discussed in the following section. Olson et al. (2000) define collaboration readiness as follows:

Collaboration readiness refers to the extent that potential collaborators are motivated to work with each other. In terms of collaboratory introduction, success seems to require a positive orientation toward collaboration, either as a result of incentives or as a result of normative practice. (Finholt, 2003: 20)

One of the ways to improve the extent that potential collaborators are motivated to work with each other, is to develop the collaboratory on the basis of already existing social networks. Sonnenwald (2007: 655) stresses that social networks and personal factors provide a critical foundation for collaboration. Personal factors play a role in establishing and sustaining these social networks and, subsequently, collaborations. Personal compatibility, including similar approaches to science, comparable working styles, mutual respect, trust, and the ability to get along and enjoy one another's company are often used to identify and select collaborators (Creamer, 2004; Hara et al.,

2003; Maglaughlin & Sonnenwald, 2005). In the end, personal factors like these might increase the engagement of all the individual participant during the project. Obviously, those who are involved in multiple projects are likely to be unavailable at least some of the time when their presence is desired. Even so, full engagement of the participants goes beyond physical participation and also involves intellectual or emotional attention to the project (Lawrence, 2006: 405).

In practice, not every collaboratory is based on existing social networks and not every collaboratory consists of participants who share a common approach to science or have a comparable working style. During a project, collaborators may also come to realise they do not have shared norms with respect to students' participation (Cummings & Kiesler, 2003) or sharing information about the research with outsiders (Walsh & Maloney, 2002). Participant of a collaboratory need to negotiate these differences. Otherwise, they might frustrate the whole collaboration process. The same goes for reaching consensus regarding authorship inclusion and order. This may also be a delicate issue, and these difficulties increase as competition intensifies (Atkinson et al., 1998, as cited in Bos et al., 2007: 670). Although the whole process of authorship order does not have the same dynamics in SS&H as it has in natural sciences, the discussion of who publishes on what in a collaboratory is an important one. The expected academic output of a collaboratory is obviously a crucial factor for the engagement and motivation of participants. In the literature, this incentive is actually not as prominent as one might expect. However, it could well be that good publication possibilities are assumed to be a *sine qua non*.

Another 'personal' factor in academic collaboration is scientific reputation and standing. In a critical observation, Finholt (2003: 21) argues that the availability of a means for contact between two scientists does not guarantee that contact will occur. He presents some evidence which suggests that e-research echoes many of the exclusive and selective tendencies that have characterized the historic organization of science. In an examination of an early system that supported network mediated communication among scientists, Hiltz and Turoff (1993) found that elite scientists using the system were more likely to receive messages than non-elite scientists, but that elite scientists were more likely to ignore the messages they received, particularly when those messages were sent by non-elite scientists (Finholt, 2003: 21).

Collaboration technology readiness

In addition to collaboration readiness, collaboration technology readiness is an additional barrier for successful collaboratories identified by Olson et al. (2000). It refers both to the presence of sufficient technology infrastructure and to the availability of local technology expertise, both explicit and implicit (Finholt, 2003: 20). Olson et al. (2000: 12) describe a progression from applications that require minimum training, such as e-mail, to technologies that require greater investment, such as data conferencing tools. Attempts to leapfrog steps in this progression can produce frustration and resistance.

Some technologies and related best practices are mentioned in the literature: systems that are easily integrated into existing work environments are more readily adopted (standalone application vs. browser accessible) (Finholt, 2003: 19); overcoming the difficulties inherent in virtual interaction is partially a matter of elaborating designs and technologies that make virtual settings more like

physical settings (Finholt, 2003: 20); incompatible hardware, software, and data structures can become critical technical hurdles (Zimmerman, 2008: 2); and distributed knowledge centres require standardization of data and long-distance technical support (Bos et al., 2007: 665).

In general, social sciences and humanities tend to be a low-tech environment and applications that require minimum training can already produce some resistance. Hence, some of the best practices mentioned in the literature actually seem too advanced for the type of user one can encounter in SS&H. Instant messaging or video conferencing tools, for instance, might approximate the convenience of face-to-face collaboration and increase workplace awareness (Bos et al., 2007: 665), but they appear to be too high-tech for this type of users. Despite the fact that these technologies could allow distant collaborators to know when others are interruptible, in order to engage in the quick consultations and informal chat that are the glue of co-located interaction, the impression is that they are not compatible with the low-tech environment of SS&H. Hence, they could produce resistance rather than acceptance and in those instances, as Lawrence (2006: 408) shows, low-tech solutions can be more useful and more readily accepted than computer-based tools that are meant to accomplish the same function.

Collaborative software

In recent years, some of the collaboration technologies mentioned above, like instant messaging and data conferencing tools, are integrated into single packages of collaborative software. In commercial organisations, this software is often referred to as enterprise content management (ECM) and it is used to capture, manage, store, preserve, and deliver content and documents. At the moment, most large computer technology corporations, as well as a range of open source organisations, have developed collaborative software. Some prominent examples are listed here below.

- Alfresco (<http://www.alfresco.com/>)
- Basecamp (<http://www.basecamphq.com/?referrer=YT30B7>)
- Drupal (<http://drupal.org/>)
- Confluence (<http://www.atlassian.com/software/confluence/>)
- Liferay (<http://www.liferay.com/web/guest/products/portal>)
- Livelink (<http://www.opentext.com/2/global/sol-products/sol-pro-open-text-ecm-suite.htm>)
- Notes (<http://www-01.ibm.com/software/lotus/products/notes/>)
- OnBase (<https://www.onbase.com/english/index.aspx>)
- Oracle Beehive (<http://www.oracle.com/products/middleware/beehive/index.html>)
- O3spaces (<http://www.o3spaces.com/>)
- Sharepoint (<http://www.microsoft.com/Sharepoint/default.aspx>)
- WebEx Connect (<http://www.webex.com/partners/webex-connect.html>)

In order to be adopted and used, collaborative software should provide benefits over current practices. Sonnenwald (2007: 661) lists some of these benefits: the software has to be compatible with scientists' values, experiences, and needs; it should be easy to try out or to use; and it should have clearly defined results (Rogers, 1995). In comparison to, for instance, mailing lists and e-mail discussions, collaborative software should provide some clear benefits. According to de Moor (2004: 261), mailing lists have proven to be very successful in bringing people together into virtual

communities, but they have many limitations such as information overload, navigation problems, primitive workflow management capabilities, and lack of customization. This negatively affects motivation and the accomplishment of joint objectives. Likewise, e-mail discussions on a particular topic are deemed to be prolonged, divergent, and repetitive (de Moor, 2004: 268). In addition, Finholt (2003: 19) argues that experience to date with laboratories suggests: systems that are easily integrated into existing work environments are more readily adopted (standalone application vs. browser accessible). He also adds that some domains of activity are more naturally inclined towards collaboration (data collection vs. contemplation and idea formation).

Tacit knowledge

One of the aspects which seems to be very important for collaboration at-a-distance, but which is only mentioned by some scholars in the reviewed literature, is the transfer of tacit knowledge among the collaborators. Heimeriks and Vasileiadou (2008: 14) say the following on this:

In many fields the use of ICTs for knowledge generation and exchange requires the transfer of knowledge into certain formats. For this purpose previously 'tacit' or 'implicit' knowledge needs to be captured in an explicit way and therefore to be transformed. (...) This increasing codification of tacit knowledge seems to become more relevant as data in many fields become digitized, more mobile and interchangeable.

This visibility and codification of previously informal and tacit knowledge, and the use of ICTs for recording has another implication as well. The use of ICTs makes the informal processes of research collaboration traceable and visible: processes that took place in the lab corridors before, such as socializing and exchanging pleasantries, decisions about research tasks, disagreements about roles and responsibilities. (Heimeriks and Vasileiadou, 2008: 14)

Although it is unclear to what extent, for instance, disagreements about responsibilities become more visible through the use of ICT, one effect of e-research projects seems to be that it makes visible or explicit many processes that were previously invisible or implicit. As Schroeder and Fry (2007) argue, e-research projects do this partly because computerization necessitates, for example, that data are put into a certain format or stored in a certain way or made accessible across a network. In an e-research setting, things like workflow organization, intellectual property issues, and inter-institutional commitments, require standardization of rules and procedures that may previously have been unstated or unformalized or do not apply to non-digital material (Schroeder and Fry, 2007).

Obviously, there are limits to what one actually can make explicit. Not all knowledge is recordable in easily transferred forms, such as documents (cf. Finholt and Olson, 1997: 28), and as a result it is not easily shared across distance. Nevertheless, in the case of co-creating social scientific data, one can develop elaborate rules for annotation and metadata. In doing so, one can transfer some of the tacit knowledge about a dataset and thus potentially improve interpretations of data by secondary users (cf. Zimmerman, 2008). However, such a system is very time consuming and it is unclear whether collaborators think that this extra effort produces enough extra benefits. Moreover, as Michener and his colleagues (1997, 335) argue, there is no end to metadata: 'There is no unique, minimal, and

sufficient set of metadata for any given data set, since sufficiency depends on the use(s) to which the data are put'.

As a final point, it is important to remember that e-research projects do not only have the potential to make visible or explicit many processes that were previously invisible or implicit, but they also may hide or marginalize other research efforts that are not amenable to the process of rendering research into digital formats (Schroeder and Fry, 2007).

References

Berg, Marc (1997) 'Problems and promises of the protocol ', *Social Science & Medicine* 44/8: 1081-88.

Bos, Nathan, Zimmerman, Ann , Olson, Judith , Yew, Jude , Yerkie, Jason , Dahl, Erik and Olson, Gary (2007) 'From Shared Databases to Communities of Practice: A Taxonomy of Collaboratories', *Journal of Computer-Mediated Communication* 12/2: 652-72.

Chompalov, Ivan and Shrum, Wesley (1999) 'Institutional Collaboration in Science: A Typology of Technological Practice', *Science, Technology & Human Values* 24/3: 338-72.

Cummings, Jonathon N. and Kiesler, Sara (2005) 'Collaborative Research Across Disciplinary and Organizational Boundaries', *Social Studies of Science* 35/703-722.

de Moor, A. (2004) 'Improving the testbed development process in collaboratories', in, *Conceptual Structures at Work, Proceedings*: 261-74.

Finholt, Thomas A. (2003) 'Collaboratories as a new form of scientific organisation', *Economics of Innovation and New Technology* 12/1: 5-25.

Finholt, Thomas A. and Olson, Gary M. (1997) 'From Laboratories to Collaboratories: A New Organizational Form for Scientific Collaboration', *Psychological Science* 8/1: 28-36.

Heimeriks, Gaston and Vasileiadou, Eleftheria (2008) 'Changes or transition? Analysing the use of ICTs in the sciences', *Social Science Information* 47: 5-29.

Hine, Christine (2006) 'Databases as Scientific Instruments and Their Role in the Ordering of Scientific Work', *Social Studies of Science* 36/2: 269-98.

Lawrence, Katherine A. (2006) 'Walking the Tightrope: The Balancing Acts of a Large e-Research Project ', *Computer Supported Cooperative Work* 15/4: 385-411.

Michener, William K. , Brunt, James W. , Helly, John J. , Kirchner, Thomas B. and Stafford, Susan G. (1997) 'Nongeospatial Metadata for the Ecological Sciences ', *Ecological Applications* 7/1: 330-42.

Olson, Gary M., Bos, Nathan and Zimmerman, Ann S. (2008) 'Introduction', in Gary M. Olson, Ann S. Zimmerman and Nathan Bos (eds), *Scientific Collaboration on the Internet* (Cambridge, Massachusetts: MIT Press): 1-12.

Olson, Gary M. and Olson, Judith S. (2000) 'Distance Matters', *Human-Computer Interaction* 15/2&3: 139-78.

Olson, Gary M., Zimmerman, Ann S. and Bos, Nathan (2008) *Scientific Collaboration on the Internet* (Cambridge, Massachusetts: MIT Press).

Schroeder, Ralph and Fry, Jenny (2007) 'Social Science Approaches to e-Science: Framing an Agenda', *Journal of Computer-Mediated Communication* 12/2.

Shrum, Wesley, Genuth, Joel and Chompalov, Ivan (2007) *Structures of Scientific Collaboration* (Cambridge, Massachusetts: MIT Press).

Sonnenwald, Diane H. (2007) 'Scientific Collaboration ', in B. Cronin (ed), *Annual review of information science and technology* (Medford, NJ: Information Today): 643-82.

Zimmerman, Ann S. (2008) 'New knowledge from old data. The role of standards in the sharing and reuse of ecological data', *Science Technology Human Values* 33/5: 631-52.

Eindverslag gebruikersonderzoek

In het kader van het eerste en grootste werkpakket van het Hublab2-project is een uitgebreid gebruikersonderzoek gedaan, waarbij nauwgezet de ervaringen van de gebruikers met de collaboratory-omgeving Liferay in kaart gebracht en geanalyseerd zijn. In dit eindverslag worden de belangrijkste uitkomsten en inzichten van dit onderzoek op een rij gezet, met als uiteindelijke doel om concrete aanbevelingen te geven voor de optimalisatie van het gebruik van virtuele platforms voor geesteswetenschappelijk onderzoek.

In dit verslag wordt ter introductie eerst een kort overzicht gegeven van de functionaliteiten van de nieuwe versie van Liferay zoals die eind augustus 2009 geïnstalleerd is in het kader van werkpakket 2. Daarna wordt het gebruik van Liferay door de gebruikers in kwantitatieve zin in kaart gebracht, de *fact and figures*, om nadien de resultaten te presenteren van de enquête onder een deel van de gebruikers van Liferay. Vervolgens wordt verslag gedaan van het veldwerk, waarin participerende observaties en interviews gebruikt zijn om meer inzicht te krijgen in de rol van de Liferay omgeving in het samenwerkingsproces binnen de diverse collaboratories. Op het einde van dit verslag zullen concrete aanbevelingen gedaan worden voor een optimaal gebruik van collaborative software binnen een collaboratory in de geesteswetenschappen.

Liferay 5.2

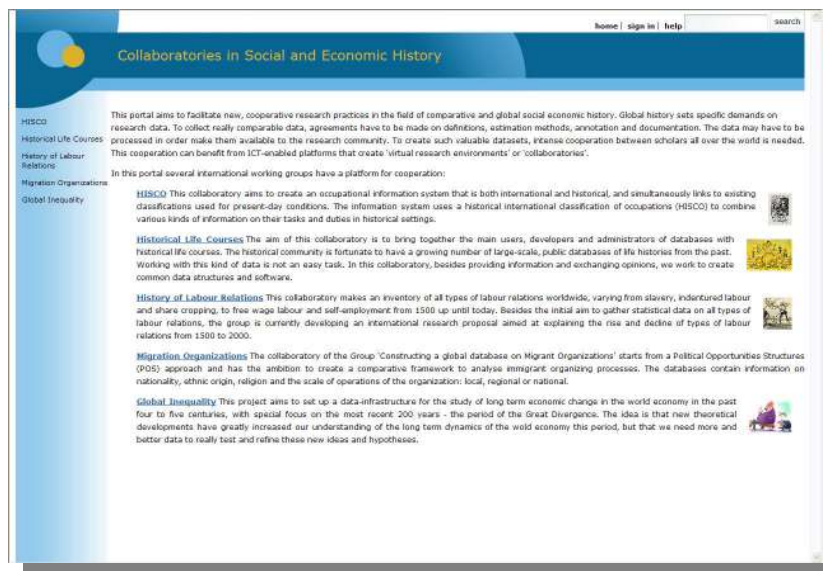
In het kader van werkpakket 2 en op basis van de inzichten uit Hublab1 is gestreefd naar een verhoging van de gebruikersvriendelijkheid van de Liferay omgeving. Nadat in de pilotfase gebruik was gemaakt van een oudere versie van Liferay met een standaard vormgeving (zie Figuur 1) is er voor Hublab2 een nieuwe versie geïnstalleerd en tevens een nieuwe vormgeving ontwikkeld.

The screenshot displays the 'Document Library' interface. At the top, there is a header with the title 'History of Labour Relations' and a user greeting 'Welcome, stefan dormans!'. Below the header is a navigation menu with options: 'About the Collaboratory', 'Events', 'Documents' (selected), 'Forum', and 'Public'. The main content area is titled 'Document Library' and includes a 'Folders' section with 'My Documents' and 'Recent Documents' tabs. There is an 'Add Subfolder' button and a 'WebDAV URL' field containing 'https://collab.iisg.nl:443/tunnel-web/secure/webdav/liferay.cc'. Below this is a 'Documents' section with a search bar and buttons for 'Search File Entries', 'Add Document', and 'Add Shortcut'. A table shows 6 results with columns for Document, Size, Downloads, and Locked. Each row has a 'View' link.

Document	Size	Downloads	Locked
conceptmanual.doc	39.5k	2	No
Definitions Labour Relations/def.doc	26.0k	0	No
Source book of the database.doc	37.0k	0	No
taxonomy.doc	101.0k	1	No
User Guide Collabs.doc	1.197.5k	0	No
UsingHISCO.doc	24.0k	2	No

Figuur 1: Document library (pilotversie)

Uitgangspunt bij al deze vernieuwingen was dat de interface meer intuïtief zou moeten worden, de uitstraling professioneler, en het menu eenvoudiger. In de pilotversie werd – mede omdat gebruik werd gemaakt van een reeds bestaande standaard opmaak – met name de verhouding tussen het publieke en het afgesloten deel als verwarrend ervaren. Zodra men in het afgesloten deel van Liferay kwam, verdwenen de tabbladen die navigeren in het publieke deel mogelijk maakte (zie Figuur 1). Dit had o.a. tot gevolg dat veel *community owners* (de functionele beheerders van een collaboratory) informatie uit het publieke deel kopieerden naar het afgesloten deel en vice versa. In de nieuwe versie zijn het publieke en het afgesloten deel dan ook op een meer overzichtelijke manier naast elkaar gezet. In aanvulling op de vernieuwde opzet van het menu is ook een eigen visuele stijl ontwikkeld die in vorm meer aansluit bij de inrichting van de website van het IISG, dus met een menu aan de linkerkant en enkele functies boven de *banner* aan de rechterbovenzijde van de pagina. Er is overigens wel gekozen om in het kleurgebruik uitdrukkelijk niet de huisstijl van het IISG te kopiëren, omdat diverse collaboratories met internationale partners samenwerken die – zo was de indruk – een te sterke visuele aanwezigheid van het IISG in het project niet zouden waarderen. Een punt dat hier tenslotte ook nog genoemd dient te worden is dat er reeds vanaf de pilotfase gestreefd is naar het aanbieden van een beperkt aantal functionaliteiten. Onze indruk was – en is nog steeds – dat *collaborative software* (zeker bij een groep van reeds samenwerkende onderzoekers) in eerste instantie die functionaliteiten moet aanbieden die een duidelijke meerwaarde hebben ten opzichte van reeds gebruikte software of communicatiemiddelen. Alleen als een samenwerkingsproject vanaf het begin beschikt over *collaborative software* kan gekeken worden of er gebruik gemaakt kan worden van een ruimer aantal functionaliteiten. In dat laatste geval is er echter wel een duidelijke visie vanuit het projectteam vereist over de noodzaak om die functionaliteiten (zoals bijvoorbeeld een gezamenlijk forum in plaats van een mailinglijst) als groep te gebruiken.



Figuur 2: Collaboratory portal (collab.iisg.nl)

De portal

De portal vormt de algemene toegang tot de diverse collaboratories zoals die door het IISG gehost worden (zie Figuur 2). In 2009 waren dit in totaal vijf collaboratories, die elk met een korte

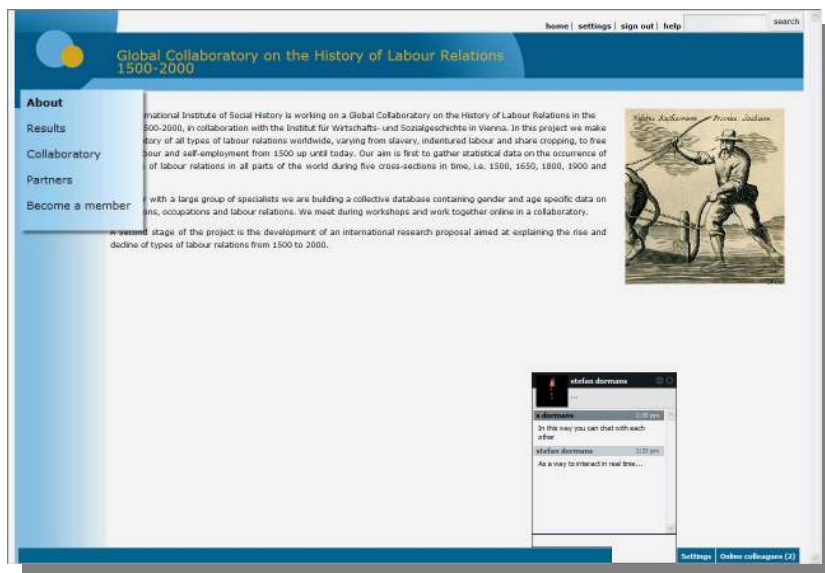
introducerende tekst gepresenteerd werden op de portal. Naast toegang tot deze collaboratories biedt deze portal enkele standaard functionaliteiten die verder ook op elke pagina binnen Liferay (met uitzondering van de *control panel*) zichtbaar zijn in de hoek rechtsboven van het scherm: 'home', 'sign in', 'help' en 'search'. Indien iemand aangemeld is in het systeem, verandert 'sign in' overigens in 'sign out' en verschijnt er een extra functionaliteit: 'settings'. In Tabel 1 staat kort aangeven waar elk van deze functionaliteiten voor staat.

Functionaliteit	Omschrijving
Home	Navigeert terug naar portal (collab.iisg.nl)
Sign in/out	Aanmelden of afmelden in het systeem, hetgeen vervolgens alleen toegang geeft tot het afgesloten deel van een collaboratory als iemand ook daadwerkelijk lid is van die collaboratory
Help	Geeft toegang tot de help-wiki, alwaar helpteksten staan voor zowel eindgebruikers als voor individuele beheerders
Search	Zoekfunctie, voor het doorzoeken van alle pagina's en documenten
Settings	Geeft toegang tot de 'control panel', alwaar de gebruiker de individuele informatie en instellingen kan bepalen en de beheerder de collaboratory kan aansturen en inrichten

Tabel 1: Standaard functionaliteiten

Publieke deel

De publieke pagina's van elke collaboratory zijn zowel via de *portal* als via een specifieke URL of via een zoekmachine te vinden. Het menu aan de linkerkzijde van de pagina biedt een bezoeker van deze pagina's een vijftal opties: 'about', 'results', 'collaboratory', 'partners', en 'become a member' (zie Figuur 3). Dit menu is constant zichtbaar, vanaf elke plek in Liferay, behalve in het 'control panel' gedeelte dat in een standaard Liferay vormgeving is opgemaakt.



Figuur 3: Publieke homepage Labour Relations-collaboratory (menu vergroot en chat geactiveerd)

De 'become a member'-optie is dynamisch: diegene die reeds lid is van de betreffende collaboratory ziet deze optie niet meer. Voor anderen biedt deze optie de mogelijkheid om een verzoek voor lidmaatschap in te dienen bij de beheerder van de collaboratory. De 'about'-sectie is bedoeld als een algemene introductie tot de collaboratory in kwestie en biedt informatie over de achtergronden,

doelen en opzet van de samenwerking. De 'results'-sectie kan reeds behaalde resultaten van de samenwerking tonen. Afhankelijk van de fase waarin de samenwerking zich bevindt kan deze pagina publicaties, databases, notulen, *working papers* of *position papers* bevatten. Toegang tot documenten kan met hyperlinks gecreëerd worden. De 'partners'-sectie biedt de mogelijkheid om het netwerk aan betrokkenen te presenteren (al dan niet inclusief links naar de relevante homepages). De 'collaboratory'-optie, tenslotte, biedt toegang tot het afgesloten deel van de software. Indien iemand lid is van de collaboratory (en ingelogd is in het systeem) dan wordt deze toegang automatisch verschaft.

Afgesloten deel

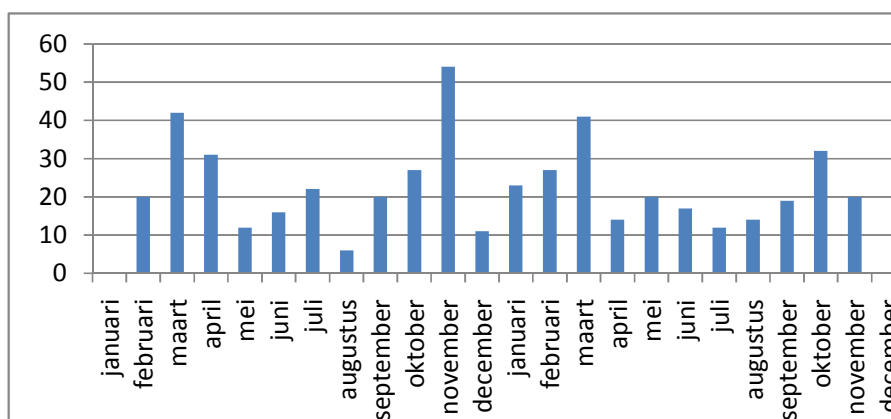
Het afgesloten gedeelte van de collaboratory bestaat uit een viertal opties die als submenu van het standaard menu aan de linkerkzijde van het scherm constant getoond worden als iemand in de betreffende collaboratory zit (zie Figuur 4). De 'news'-sectie is de homepage voor de leden van de collaboratory: hier kunnen ze elkaar op de hoogte houden van recente ontwikkelingen en aankomende bijeenkomsten of deadlines. De 'documents'-sectie bevat een folderstructuur die naar eigen inzicht ingericht kan worden en waarin documenten en andere bestanden geplaatst, bekeken en gedownload kunnen worden. Ook is het mogelijk om in deze *document library* gezamenlijk aan documenten of databestanden te werken door middel van versiebeheer. De 'forum'-sectie bevat een *message board* waarin op verschillende zelfgekozen thema's discussies gevoerd kunnen worden. De 'members'-sectie toont alle leden van de collaboratory en geeft tevens de mogelijkheid om al die leden (of een selectie van hen) direct een mailbericht te sturen. Naast deze verschillende secties met bijbehorende functionaliteiten, biedt het afgesloten deel ook de mogelijkheid om te chatten met andere leden van de collaboratory. Het aantal 'online colleagues' – dat wil zeggen, andere leden van de collaboratory die ook ingelogd zijn in het systeem – is permanent te zien beneden in de rechterhoek en door de naam van een *online* lid aan te klikken verschijnt een chatbox om één-op-één te chatten (zie Figuur 3).



Figuur 4: Nieuwspagina afgesloten deel HLC-collaboratory (menu vergroot)

Kwantitatieve analyse

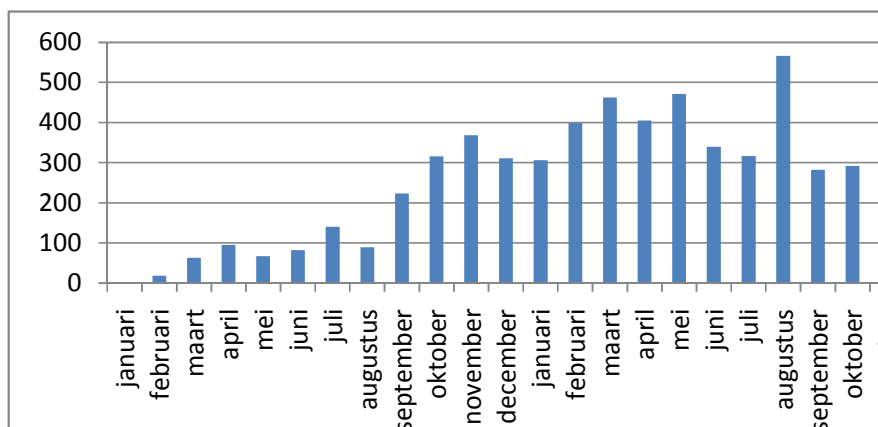
Om het gebruik van de Liferay software te kunnen volgen is in het kader van Hublab2 AWSTAT-software geïnstalleerd. Deze software maakt het mogelijk om het bezoek en gebruik van het publieke en het afgesloten deel van de websites in kwantitatieve zin te monitoren. In Tabel 2 en Tabel 3 zijn het aantal bezoekers opgenomen die in de periode 2008-2009 het publieke en het afgesloten deel van Liferay bezocht hebben. Bij deze gegevens dient aangetekend te worden dat er een zekere vertekening plaatsvindt door het feit dat er in bepaalde periodes veel verkeer op de websites was in verband met het ontwikkelen van nieuwe functionaliteiten of het implementeren van een nieuwe versie. Dit interne gebruik verklaard o.a. de piek in augustus 2009 in het afgesloten deel van de website, omdat toen de nieuwe versie van Liferay geïnstalleerd is.



Tabel 2: Unieke bezoekers publieke deel (<http://collab.iisg.nl>) in periode januari 2008 – december 2009.¹

Zoals uit Tabel 2 en Tabel 3 blijkt is er een groot verschil tussen het aantal bezoekers van het publieke deel (tussen de 0 en 60 unieke bezoekers per maand) en het afgesloten deel (vanaf zomer 2008 tussen de 200 en 500 unieke bezoekers). Deze gegevens reflecteren het feit dat Liferay vooralsnog primair gebruikt wordt als platform voor interne samenwerking binnen de collaboratories en nog nauwelijks als extern uithangbord voor de diverse projecten. Uiteindelijk zijn er hooguit enkele personen per week die de publieke delen van de website bezoeken, waarbij de kans groot is dat een deel van die bezoekers reeds bij één van de collaboratories betrokken is en dus niet echt een externe bezoeker is. Wat verder opvalt aan deze gegevens is dat het bezoek fluctueert en dat er bijvoorbeeld niet sprake is van een structurele groei van het bezoek aan beide delen van Liferay. Ons inziens hangt dit deels samen met de verhoogde activiteiten in voorbereiding op de diverse workshops. In november 2008, alsmede in maart 2009 en november 2009 zijn internationale workshops gehouden, waarbij het materiaal grotendeels via Liferay is gedeeld. Verder tonen de gegevens dat sinds in afwachting van de installatie van de nieuwe versie van Liferay en ook in de periode daarna het gebruik van Liferay minder is dan bijvoorbeeld in het eerste deel van 2009.

¹ Zonder aanwijsbare oorzaak is de data van gebruik na 17 november 2009 niet opgeslagen, vandaar de lage scores voor de laatste twee maanden. Voor het afgesloten deel van de website is wel data beschikbaar, zie Tabel 3. Zie voor meer data <http://hygeia.iisg.nl/awstats/awstats.pl?config=collab&configdir=/etc/awstats>.



Tabel 3: Unieke bezoekers afgesloten deel (<https://collab.iisg.nl>) in periode januari 2008 – december 2009.²

De gegevens die via de AWSTAT-software verzameld zijn geven vooral inzicht in het bezoek van de websites, maar geven minder zicht op het daadwerkelijke gebruik. Daarom zal in het onderstaande een aanvullend kwantitatief beeld geschetst worden van het gebruik van Liferay. In Tabel 4 is het aantal leden gegeven dat de vijf collaboratories begin 2010 hadden.³ Dit zijn dus personen die toegang hebben tot Liferay en die lid zijn van één of meerdere collaboratories. In totaal zijn er momenteel tussen de 150 en 200 leden die toegang hebben tot deze collaborative software. Het aantal actieve leden is uiteindelijk kleiner, omdat sommige collaboratories vooralsnog ook geïnteresseerden lid laten zijn.

	Aantal leden
HISCO	28
Historical Life Courses	37
Labour Relations	88
Migrant Organisations	9
Global Inequality	8

Tabel 4: Aantal leden collaboratories op collab.iisg.nl (meetpunt: 7 januari 2010)

Deze leden hebben toegang tot het afgesloten deel van hun collaboratory en beschikken aldaar over een aantal functionaliteiten voor het delen van documenten en voor het communiceren met elkaar. De *document library* is in deze fase de meest intensief gebruikte functionaliteit, waarbij de mogelijkheid wordt geboden voor het uploaden, downloaden, en bewerken van documenten (inclusief versiebeheer). De diverse collaboratories gebruiken de *document library* momenteel met name voor de uitwisseling van onderzoeksdata, presentaties tijdens workshops, de programma's van

² Zie voor meer data <http://hygeia.iisg.nl/awstats/awstats.pl?config=collabssl&configdir=/etc/awstats>.

³ Voor de collaboratories 'Migrant Organisations' and 'Global Inequality' geldt momenteel dat men nog in een opstart-fase zit waarin de collaboratory wel al is opgezet en deels voorzien van *content*, maar nog niet actief gebruikt wordt in de ondersteuning van samenwerking in het onderzoeksproject. In het geval van de MO-collaboratory is intussen overigens een subsidie verstrekt om de samenwerking – waarbij het platform een centrale rol speelt – verder te ontwikkelen en te intensiveren.

workshops, de planning, methodologische stukken en *position papers*. De nadruk ligt hierbij op informatie die gerelateerd is aan workshops.

Zoals de gegevens in Tabel 5 aangeven, wordt de *document library* het meest actief gebruikt in de collaboratory met de meeste leden: de 'Global Collaboratory on the History of Labour Relations 1500-2000'. Niet alleen omvat deze collaboratory circa drie maal zo veel documenten als de andere vier collaboratories tezamen, maar ook worden deze documenten veel vaker gedownload dan in de andere collaboratories. Een punt dat in dit verband wel nog genoemd dient te worden is dat een groot deel van de documenten die in deze collaboratory geplaatst zijn, niet door de leden zelf zijn geupload, maar door een tweetal leden van het projectteam. Zij verzamelden de documenten – bijvoorbeeld in aanloop naar een workshop – en zorgden ervoor dat deze in de *document library* werden gezet.

	Aantal documenten	Aantal downloads
HISCO	6	69
Historical Life Courses	27	119
Labour Relations	247	2640
Migrant Organisations	14	76
Global Inequality	49	65

Tabel 5: Gebruik Document libraries (meetpunt: 7 januari 2010)

Daar waar het gebruik van de *document library* intensief genoemd kan worden, is het gebruik van andere functionaliteiten, zoals het discussieforum, hier duidelijk bij achtergebleven. Zoals de gegevens uit Tabel 6 duidelijk maken zijn er relatief weinig berichten (*posts*) geplaatst in de diverse discussiefora. Wel zijn deze berichten over het algemeen regelmatig bekeken (gemiddeld bijna 15 *views* per *post*). De aard van de berichten varieerde van voorzetten voor discussies en updates over de voortgang van het project, tot algemene mededelingen.

	Aantal posts	Aantal views
HISCO	5	54
Historical Life Courses	11	161
Labour Relations	3	96
Migrant Organisations	1	9
Global Inequality	4	23

Tabel 6: Gebruik Forum (meetpunt: 7 januari 2010)

Aangezien er van enkele andere functionaliteiten, zoals de chat en de mail, niet automatisch een log wordt bijgehouden is niet te achterhalen hoe actief deze gebruikt worden. Uit gesprekken met andere gebruikers en uit eigen ervaring komt het beeld naar voren dat deze functionaliteiten vooralsnog weinig gebruikt worden. Dit kan mede samenhangen met het feit dat het pas sinds eind augustus 2009 mogelijk is om binnen deze collaboratories te chatten en een mail te versturen. Uit de resultaten van de enquête blijkt tevens dat de gebruikers vinden dat deze functionaliteiten voor hen geen toevoegde waarde heeft. Men blijft liever meer 'traditionele' communicatiemiddelen gebruiken zoals email en mailinglijsten.

Enquête

In december 2009 is onder de leden van de twee meest actieve collaboratories, namelijk 'Historical Life Courses' en de 'Global Collaboratory on the History of Labour Relations in the period 1500-2000', een enquête over de ervaringen met Liferay afgenomen. De keuze om alleen leden van deze twee collaboratories te enquêteren was gebaseerd op het feit dat de vragenlijst een volledig ingerichte en actief gebruikte collaboratory veronderstelde. In totaal zijn 50 participerende onderzoekers benaderd en uiteindelijk hebben 9 respondenten op de enquête gereageerd. In de betreffende enquête werden een achttal korte vragen gesteld over de algemene indruk die men van Liferay had, de frequentie waarmee men Liferay bezocht, de functionaliteiten die men al dan niet gebruikte, de gebruiksvriendelijkheid en ondersteuning, alsmede de bruikbaarheid van Liferay voor het samenwerkingsproces (zie Tabel 7 voor de volledige vragenlijst). Voor de verspreiding van de enquête is gebruik gemaakt van het Forum en de mailfunctie van Liferay. In het vervolg van deze sectie zal puntsgewijs een impressie van de verschillende antwoorden gegeven worden.

1. What is your general impression of the online platform of the IDS project (see website)?
2. How frequently have you visited the platform until now?
3. Which of the following functions did you use the platform for?
 - a. Share data or documents with others.
 - b. Review data or documents by others.
 - c. Communication with other members of the group.
 - d. Getting up to date with the progress on the project.
4. Could you indicate why you use some of these functions and why you do not use other functions?
5. Have you encountered any problems in accessing or using the platform and have you been able to solve these problems easily?
6. Are there certain functionalities you miss in the current version of the platform?
7. Could you give some reasons why you think that this platform is - or is not - a useful tool to support your collaboration on the Intermediate Data Structure?
8. Are there any other relevant issues with regard to the online platform you would like to mention?

Tabel 7: vragenlijst enquête (voor leden van de HLC-collaboratory)

Algemene indruk en bruikbaarheid

De algemene indruk waar in vraag 1 naar geïnformeerd wordt is zonder uitzondering positief. Zowel het concept van een collaboratory als het daadwerkelijk platform worden gunstig geëvalueerd en als een duidelijke meerwaarde voor de onderzoekspraktijk gezien. Een respondent noemt expliciet dat de nieuwe versie een duidelijke verbetering is ten aanzien van de navigeerbaarheid – die wordt meer intuïtief gevonden – en ook de verhouding van het publieke en afgesloten deel is veel helderder geworden. De enige kanttekening die verschillende malen geplaatst wordt, is dat er vooralsnog relatief weinig gebruik wordt gemaakt van Liferay, waardoor een echte evaluatie van de potentie van de software moeilijk is. Het idee lijkt te zijn dat bij intensief gebruik, Liferay een belangrijke bijdrage aan de samenwerking zal leveren. Verder geeft één respondent aan dat hij/zij nog weinig gebruik heeft gemaakt van Liferay, omdat het ingewikkelder lijkt om het forum te bezoeken dan om alle betrokkenen een groepsmail te sturen.

In aanvulling op de algemene indruk van vraag 1, is in vraag 7 gevraagd naar de bruikbaarheid van Liferay in de ondersteuning van het samenwerkingsproces. Ook hier waren de reacties positief en

werd Liferay als een goede mogelijkheid gezien om het eigen werk te vergelijken met het werk van anderen en zo het eigen onderzoek te stimuleren. De toegankelijkheid van data, snelheid, flexibiliteit, en het bieden van een discussieplatform worden in dit kader expliciet genoemd. Ook hier wordt echter weer het punt genoemd dat intensiever gebruik van Liferay van belang is om voor een echte meerwaarde te zorgen. Eén van de respondenten stelt dan ook dat het probleem bij de gebruikers ligt en niet bij het platform *an sich*. Daarnaast stelde een andere respondent dat Liferay niet los gezien kon worden van de plenaire bijeenkomsten, die onmiskenbaar ook een toegevoegde waarde hebben: “collab-platform en plenaire bijeenkomsten vormen een twee-eenheid bij onderzoeksprojecten van dit formaat”. Een laatste punt dat hierbij genoemd is, is dat het moeilijk is om overzicht te krijgen van wat wel en wat nog niet besproken is. Voor iemand die later bij een collaboratory betrokken is geraakt, is het onduidelijk of de vragen waarmee hij/zij zit al bediscussieerd zijn. Dit zou minder een probleem zijn als de collaboratory al vanaf het begin intensief gebruikt was en bijvoorbeeld alle discussies in het Forum terug te vinden waren, maar daarvan is in deze gevallen nooit sprake geweest. Een respondent spreekt in dit kader nog de verwachting uit dat mogelijk aan het einde van het project intensiever gebruik van Liferay gebruikt zal worden, als eenieder de eigen data dient af te ronden en alle individuele sets verder op elkaar afgestemd dienen te worden. Als dat het geval is dan zou dat nog een extra kans bieden om de bruikbaarheid van Liferay te observeren.

Intensiteit gebruik

De frequentie van het gebruik varieert, volgens de antwoorden van de respondenten, tussen eens per twee of drie maanden tot enkele malen per maand. Het bezoek hangt logischerwijs af van de werkzaamheden die men voor het betreffende project dient te verrichten. Eén respondent geeft bij vraag 2 een uitgebreide verklaring waarom hij/zij tot nu toe in beperkte mate gebruik maakt van het platform: “De focus is op mijn eigen data en hun specifieke problemen. Uiteraard is het mogelijk om door middel van e-mailcontact algemene problemen aan de orde te stellen. Bij mezelf bespeur ik echter de neiging en voorkeur om ‘algemene zaken’ tijdens plenaire workshops aan de orde te stellen. Het grote voordeel daarvan is dat ‘algemene zaken’ ieders aandacht heeft en besproken kan worden, terwijl bij het rondmailen de respons maar moet worden afgewacht (afgezien van de non-respons).” In aanvulling hierop geeft de betreffende respondent aan dat hij/zij het gebruik van MS Access in plaats van MS Excel als een barrière ervaart: “Met Excel kan ik veel gemakkelijker overweg en daardoor ook makkelijker cijfers berekenen en interpreteren”.

Gebruik diverse functies

Bij vraag 3 is gevraagd naar het gebruik van een viertal verschillende functies die het platform biedt. Het delen van data en documenten wordt door de meerderheid van de respondenten gedaan met behulp van Liferay, alsmede het bekijken van data en documenten van anderen. Het communiceren met andere leden wordt daarentegen door de meeste respondenten niet via Liferay gedaan. Dit zou mede verklaard kunnen worden door het feit dat pas sinds eind augustus 2009 de mogelijkheden op dit vlak verbeterd zijn door de installatie van de nieuwe versie van Liferay (incl. chat en mail). Bij de vierde en laatste functie, het op de hoogte blijven van de ontwikkelingen rond de collaboratory, wordt weer door de meeste respondenten aangegeven dat ze het platform daarvoor gebruiken. Hierbij dient vermeld te worden dat in de huidige situatie bij de twee onderzochte collaboratories afzonderlijke leden geen nieuwsberichten op het platform kunnen plaatsen. Dit wordt gedaan door leden van het project team. Natuurlijk zouden de leden wel een bericht kunnen *posten* in het forum.

In antwoord op vraag 4 gaven enkele respondenten vervolgens aan waarom ze bepaalde functies wel of niet gebruiken. Opvallend is dat nagenoeg alle antwoorden betrekking hebben op communiceren via Liferay. Eén respondent geeft aan dat de afwezigheid van levendige discussies hem/haar weerhoudt van actieve participatie, terwijl twee respondenten aangeven dat email volstaat als communicatiemiddel of zelfs de voorkeur geniet. Liferay lijkt voor deze respondenten geen meerwaarde te hebben ten opzichte van bestaande en reeds bekende communicatiemiddelen, terwijl één respondent op zich wel graag via Liferay zou willen discussiëren maar geremd wordt door de huidige inactiviteit. Hij/zij geeft aan het gevoel te hebben dat een bericht in het forum op zich wel gezien zou worden, maar dat er uiteindelijk toch geen discussie op gang zou komen.

De indruk die deze antwoorden geven is dat de diverse mogelijkheden die de *document library* biedt voor het delen van data en documenten actief gebruikt worden en dus ook een zekere meerwaarde lijken te hebben ten opzichte van de bestaande onderzoekspraktijk. Liferay lijkt dan ook een belangrijke functie te vervullen in het samenwerkingsproces als een dynamische *repository*. Eén respondent geeft aan dat hij/zij nog niet de gelegenheid heeft gehad om samen *online* te werken aan hetzelfde bestand met behulp van versiebeheer. Aangezien niemand anders deze functionaliteit noemt – en op basis van onze algemene observatie van het gebruik van Liferay – is de indruk dat ook andere gebruikers dit nog niet gedaan hebben. Voor wat betreft de interne communicatie lijkt Liferay vooralsnog geen duidelijke meerwaarde te hebben ten opzichte van bestaande communicatiemiddelen zoals email en mailinglijsten. Voor het uitwisselen van nieuws over de voortgang van het project lijkt Liferay ten slotte wel weer een duidelijke meerwaarde te hebben. Op dit moment is het echter nog niet zo dat individuele leden hun eigen nieuws op Liferay (kunnen) zetten.

Ondersteuning

De vraag of men problemen had ondervonden in het gebruik van Liferay en of deze problemen vervolgens snel werden opgelost, leverde slechts enkele korte antwoorden op. Eén respondent gaf aan dat er server problemen waren, waardoor er geen toegang tot bepaalde documenten was. Na het rapporteren van dit probleem was er echter ook snel een oplossing. Een andere respondent noemde compatibiliteitsproblemen met het gebruik van Apple's Safari browser, maar na het op eigen initiatief overstappen naar Firefox leken er geen problemen meer te zijn. Tevens is later alsnog de compatibiliteit met Safari verbeterd. Het geringe aantal reacties op dit punt kan samenhangen met het feit dat deze groep respondenten Liferay incidenteel bezocht. In ons eigen gebruik van Liferay, dat intensief was afgelopen jaar, zijn we juist vaak tegen *bugs* aangelopen en was permanente en deskundige ICT-ondersteuning vereist. Het feit dat één van de ontwikkelaars op het IISG veel ervaring had met Liferay bleek in dit kader cruciaal voor een goede ondersteuning.

Extra functionaliteiten en overige punten

Als laatste is nog gevraagd of er functionaliteiten zijn die volgens de respondenten toegevoegd dienen te worden (vraag 6) of dat er volgens hen relevante onderwerpen onbesproken zijn gebleven (vraag 8). Op beide vragen zijn geen noemenswaardige reacties gekomen. Naast het feit dat de enquête volgens deze respondenten alle relevante punten aangehaald lijkt te hebben, lijkt Liferay dus ook de functionaliteiten te bieden die men bij *collaborative software* verwacht.

Participerende observaties en interviews

Gedurende de hele duur van het project zijn voor het gebruikersonderzoek de diverse collaboratories van het IISG geobserveerd, zowel online als offline. De online observaties betrof het volgen van mailinglijsten, emailcorrespondenties, en het verkeer op het Liferay platform, alsmede het doornemen en analyseren van alle relevante documentatie die via deze wegen in 2009 gedeeld zijn. De offline observaties vonden plaats tijdens workshops van collaboratories in binnen en buitenland, tijdens vergaderingen over collaboratories en Liferay binnen het IISG, en buiten die formele bijeenkomsten om is gedurende het hele jaar nauw contact geweest met de onderzoekers, het managementen de ICT afdeling van het IISG. Het feit dat de Virtual Knowledge Studio een gebouw deelt met het IISG faciliteerde deze continue informele observaties natuurlijk. Tijdens de gehele onderzoeksperiode zijn tevens diverse etnografische interviews gehouden met gebruikers en ontwikkelaars van het Liferay platform.

In het kader van de participerende observaties en interviews is o.a. de vierde workshop bezocht van de "Global Collaboratory on the History of Labour Relations" (13 en 14 maart), alsmede een workshop van de Historical Life Courses Collaboratory in Long Beach, Ca. (11 tot en met 15 november). Tevens werd in Long Beach de jaarlijkse conferentie bezocht van de American Social Science History Association. Diverse andere collaboratories gebruiken deze conferentie om hun resultaten te tonen. Er is o.a. gesproken met de projectleiders van de Historical Violence Database (<http://cjrc.osu.edu/researchprojects/hvd/>) en van Founders and Survivors. Australian Life Courses in Historical context 1803-1920 (<http://www.foundersandsurvivors.org/>). Vanwege het verblijf in Long Beach en het besloten karakter van de betreffende bijeenkomsten zijn een aantal andere workshops van de Labour Relations collaboratory niet bezocht, te weten: de 'Middle East and North Africa Session' (12 maart in Amsterdam), de 'Second Ottoman Empire, Middle East and North Africa Workshop' (25 en 26 september in Istanbul), de 'Africa Workshop' (6 en 7 november in Berlijn), en de 'Fifth Workshop' (12-14 November in Düsseldorf). Wel is van al deze workshops het relevante materiaal doorgenomen en zijn de resultaten besproken met betrokkenen. Tevens zijn gedurende het jaar diverse vergaderingen georganiseerd en bijgewoond, zoals een introductie van de nieuwe versie van Liferay (27 mei) en een Liferay cursus voor *administrators* (26 november).

De diverse observaties en interviews zijn uiteindelijk met name gebruikt als input voor de diverse publicaties die uit het Hublab2-project voortvloeien, zoals de NCESS paper (zie bijlage 5) en het artikel voor *Historical Methods* (ingediend). Beide publicaties richten zich op het gehele proces van samenwerking in de diverse collaboratories en dus niet alleen op de rol van Liferay in de dagelijkse onderzoekspraktijk. Eén van de bevindingen die in deze artikelen besproken wordt en die in belangrijke mate gebaseerd is op het veldwerk in het kader van dit project, is de verhouding tussen *conjunctive* en *divisible tasks*.⁴ Zoals het volgende fragment laat zien, dienen de diverse taken in een collaboratory deels individueel en deels gezamenlijk uitgevoerd te worden.

⁴ Zie ook Sonnenwald, D. H. 2007. Scientific collaboration. In Annual review of information science and technology, edited by B. Cronin, 643-82. Medford, NJ: Information Today.

Divisible tasks can be executed individually and conjunctive tasks are those where each group member has to contribute in order to complete them. In the laboratories in social and economic history, an example of a divisible task is the collection of parts of the data by the individual members of a group. This sort of data collection can be done concurrently, while some divisible tasks have to be executed sequentially. The analysis of the data, for instance, can obviously only be done after the actual collection of the data. Both tasks can, in principle, be done individually, but have to be done after each other. In addition, some tasks have to be carried out conjunctively. If a laboratory uses a taxonomy to categorise its data, this taxonomy is preferably the result of a common effort. Although it is also an option that some members of a laboratory take the initiative in developing a taxonomy, eventually all members of a laboratory need to agree on the content and design of such a collectively used research tool.

In praktijk blijken de *conjunctive tasks* alleen tijdens workshops opgepakt te worden. Liferay wordt, zoals eerder ook aangegeven, niet gebruikt om gemeenschappelijk bepaalde taken uit te voeren, ondanks dat die functionaliteiten wel aangeboden (kunnen) worden. In plaats van het online samenwerken in bijvoorbeeld een *wiki*, verkiezen de leden van de laboratories in overgrote meerderheid voor het uitvoeren van gemeenschappelijke taken tijdens de workshops. Onze inschatting is dat een combinatie van online en offline samenwerken essentieel blijft voor de succesvolle uitvoering van onderzoeksprojecten. Voor de *divisible tasks* wordt Liferay momenteel wel gebruikt. Immers, individuele onderzoekers presenteren hun eigen data via het platform aan hun collega's en gezien het grote aantal downloads worden die datasets ook door andere onderzoekers bekeken en mogelijk ook gebruikt ter inspiratie en/of vergelijking.

Een ander punt dat in de publicaties besproken wordt, is de zogenaamde *collaboration technology readiness*.⁵ Dit begrip verwijst enerzijds naar de aanwezigheid van een adequate technologische infrastructuur en anderzijds naar de beschikbaarheid van lokale technische expertise, zowel impliciet als expliciet. Aangezien de onderzochte laboratories leden van over de hele wereld hebben, kan de lokaal aanwezige infrastructuur sterk verschillen. Zo heeft een lid in India slechts incidenteel toegang tot internet. Daarnaast is de technische expertise van de ondersteuning in het IISG adequaat, maar verschillen de onderzoekers sterk in de mate waarin ze technisch onderlegd zijn. Het vergeten van een paswoord bleek soms al een voldoende blokkade te zijn om het platform helemaal niet meer te bezoeken. Vaak worden documenten van deze leden uiteindelijk door leden van het project team in Liferay geplaatst. Ons inziens zullen meer geavanceerde functionaliteiten zoals *chat*, die op zichzelf een meerwaarde zouden kunnen hebben ten opzichte van de regulier gebruikte software, door een belangrijk deel van de gebruikersgroep dan ook niet gebruikt worden.

Een laatste punt dat we hier nog specifiek willen aanhalen is de rol van de projectleider en/of van het projectteam. Naar onze mening is een pro-actieve houding en een duidelijke visie op de rol van *collaborative software* in een laboratory van cruciaal belang voor het succesvolle gebruik van een platform als Liferay. In de praktijk hadden veel onderzoekers in eerste instantie een afwachtende houding ten opzichte van de Liferay software, mede vanwege het feit dat niet duidelijk was welke meerwaarde dit zou kunnen hebben, maar nadat op dit laatste punt meer duidelijkheid bestond was men uitgesproken positief en enthousiast over de mogelijkheden. Dat men uiteindelijk toch maar

⁵ Olson, G. M. and J. S. Olson. 2000. Distance matters. *Human-Computer Interaction* 15: 139-78.

incidenteel het platform bezocht had in belangrijke mate te maken met het feit dat men nauwelijks activiteit kon waarnemen op het platform. Een helder uitgedragen visie van het projectteam op de rol van Liferay in het samenwerkingsproject, in samenhang met het geven van het juiste voorbeeld – zoals het actief initiëren en modereren van discussies in het forum en het *up to date* houden van de berichtgeving – zouden ons inziens een noodzakelijke eerste stap zijn om de aanwezige bereidheid van de onderzoekers om te zetten in actiever gebruik van de software.

Conclusie

De ambitie van het Hublab2-project was om bestaande groepen van internationaal samenwerkende sociaal en economische historici te voorzien van een optimale online onderzoeksomgeving en communicatieplatform. Uitgangspunt daarbij was dat de onderzoekspraktijk leidend was en dat de online onderzoeksomgeving diende aan te sluiten bij de dagelijkse praktijken van de betrokken onderzoeksgroepen. Om dat te bereiken is een uitgebreid gebruikersonderzoek verricht en is het Liferay platform uitgebreid aangepast. Kijkend naar die dagelijkse praktijken – en de rol van Liferay daarin – dan kan geconstateerd worden (1) dat Liferay vooralsnog primair voor interne samenwerking gebruikt wordt, (2) dat Liferay door het merendeel van de onderzoekers slechts incidenteel gebruikt wordt, (3) dat de mogelijkheden voor het delen van bestanden het meest gebruikt worden en tevens de meeste meerwaarde lijken te hebben ten opzichte van andere software en/of communicatiemiddelen, (4) dat de rol van *face-to-face* ontmoetingen in de onderzoekspraktijk onverminderd groot blijft, en (5) dat in het algemeen de verwachtingen ten aanzien van het gebruik van *collaborative software* door gedistribueerde onderzoekers in de geesteswetenschappen naar beneden moeten worden bijgesteld.

Op dit moment wordt Liferay in de dagelijkse onderzoekspraktijk primair gebruikt om grote bestanden te delen en te bewaren. Op dat punt biedt het platform duidelijk een functionaliteit die tot voor kort onvoldoende beschikbaar was voor geografisch gedistribueerde onderzoekers. Het uitwisselen van bestanden per mail liep immers aan tegen de beperkte omvang die bijlage in dat soort programma's mogen hebben. Dergelijke beperkingen kent Liferay niet en het intensieve gebruik van de *document library* toont ook aan dat hier behoefte aan was.

Liferay ambieert – evenals vergelijkbare *collaborative software* – echter meer te zijn dan een *online document repository*. Liferay biedt juist al die functionaliteiten die nodig zijn voor online samenwerken. Voor elk van deze afzonderlijke functionaliteiten geldt echter dat de onderzoekers in de praktijk al gewend zijn aan andere tools (zoals e-mail en mailinglijsten), die vaak ook nog eens gebruiksvriendelijker en adequater zijn. Zo heeft de forum-functie in Liferay de beperking dat alleen leden van de *collaboratory* bereikt kunnen worden en die dienen dan ook nog eens geabonneerd te zijn op de betreffende *thread*. Gezien het infrequente gebruik van de *collaborative software* in de onderzoekspraktijk, kan verder niet verwacht worden dat functionaliteiten die afhankelijk zijn van frequent en intensief bezoek, zoals bijvoorbeeld de *chat*-functie en de kalender, daadwerkelijk gebruikt zullen worden.

Het feit dat al deze functionaliteiten in de onderzoekspraktijk niet gebruikt worden, kan te maken hebben met de egalitaire structuur van de onderzochte *collaboratories*. Gebruik kan niet

afgedwongen worden. Verder is het ritme van de groepen dermate onregelmatig dat continu intensief gebruik niet waarschijnlijk noch nodig is. Toch zien we ook dat er onder een deel van de leden van de collaboratories een duidelijk enthousiasme was voor dit soort software en dat er meer gebruik van gemaakt zou kunnen worden als de projectleiders op dat vlak het juiste voorbeeld zouden geven. Ook dan blijft het echter de vraag of een platform als Liferay werkelijk een meerwaarde zal hebben voor de onderzoekspraktijk. Immers, gecommuniceerd wordt er nu ook al, niet in de laatste plaats tijdens workshops en andere bijeenkomsten. Het belang van daarvan blijft onverminderd groot en het veldwerk dat in het kader van het Hublab2-project verricht is, versterkt wat ons betreft alleen maar het beeld dat gedistribueerde samenwerking staat of valt bij regelmatige *face-to-face* ontmoetingen en niet bij de potentie van de gebruikte *collaborative software*.

Aanbevelingen

Op basis van de inzichten uit het gebruikersonderzoek komen we tot enkele aanbevelingen ten aanzien van het optimale gebruik van *collaborative software* in de geesteswetenschappen en de sociale wetenschappen. Hierbij dient nogmaals aangetekend te worden dat onze inzichten gebaseerd zijn op specifieke samenwerkingsprojecten, waarbij de geografisch gedistribueerde onderzoekers gezamenlijk werken aan de creatie van historische datasets. Deze collaboratories kenmerken zich door vrij egalitaire verhoudingen – waardoor er bijvoorbeeld geen gewenst gedrag afgedwongen kan worden – en een onderzoekspraktijk waarin bijdragen aan de samenwerking eerder incidenteel – bijvoorbeeld in aanloop naar workshops – dan structureel zijn. Deze punten indachtig, komen we tot de volgende zes aanbevelingen:

- Op dit moment is de grootste meerwaarde van *collaborative software* voor de dagelijkse onderzoekspraktijk de mogelijkheid die het biedt om grote bestanden te delen en te bewaren. De *document library* dient dan ook een centrale plaats in het platform te krijgen, stabiel te zijn, en eenvoudig te bedienen en te vinden zijn.
- Voor de communicatie tussen onderzoekers zal *collaborative software* niet snel de bestaande communicatiemiddelen vervangen. In aanvulling op deze software blijft het dan ook nodig om na te denken over de rol van e-mail en mailinglijsten – alsmede *face-to-face contact* - in het samenwerkingsproces.
- Gezien het interpretatieve karakter van geesteswetenschappelijk onderzoek vormt *collaborative software* een te smalle basis voor gedistribueerde samenwerkingsprojecten van dit formaat. In de opzet van dergelijke projecten dient dan ook altijd voldoende rekening gehouden te worden – in praktische en financiële zin – met plenaire bijeenkomsten.
- De implementatie en het beheer van *collaborative software* vereist een goede institutionele inbedding, waarbij gezorgd wordt voor stabiele financiële en technische ondersteuning van het pakket. Zelfs een relatief licht *open source* pakket als Liferay vereist veel ICT-ondersteuning.
- Om de bereidheid onder onderzoekers te vergroten om gebruik te maken van *collaborative software* is een overzichtelijke, herkenbare en intuïtieve interface van belang. Uitgangspunt dienen daarbij niet de mogelijkheden van het pakket te zijn, maar de wensen van de onderzoekers.

- Projectleiders dienen een duidelijke visie op de rol van de *collaborative software* in het gehele samenwerkingsproces te hebben en tevens zelf het goede voorbeeld geven. Zonder initiatief en visie van sommigen zal de meerderheid van de gebruikers niet geneigd zijn om actief gebruik te maken van de software.

Liferay en Sharepoint in de onderzoekspraktijk

Het gebruik van Liferay (www.liferay.com/web/guest/products/portal) als *collaborative software* voor de collaboratories in sociale en economische geschiedenis is vanaf het begin een keuze geweest in een Nederlandse academische context waarin Microsoft's Sharepoint (sharepoint.microsoft.com) de meest dominante speler is. Onze analyse van het gebruik van het platform Liferay in het kader van het Hublab2-project roept dan ook juist vragen op over een vergelijking met ervaringen met Sharepoint. In dit verslag zullen we aan de hand van enkele bronnen die vergelijking maken, waarbij de nadruk zal liggen op de bruikbaarheid en functionaliteit van beide platforms voor de onderzoekspraktijk en bijvoorbeeld niet op de technische mogelijkheden of beperkingen van Liferay of Sharepoint.

Ter introductie zal eerst kort teruggekeken worden op de vergelijking tussen beide platforms in het kader van Hublab1. Vervolgens zal kort het gebruik van Liferay in Hublab2 besproken worden om daarna de vergelijking te maken met het gebruik van Sharepoint. Daarvoor wordt geput uit een tweetal bronnen: allereerst is er gesproken met de leidster van de collaboratory rond Gilden, die aanvankelijk met Liferay werkte maar is overgegaan op SharePoint; tevens gebruiken we (Dormans en Kok) als leden van de Virtual Knowledge Studio Collaboratory zelf ook Sharepoint in onze dagelijkse onderzoekspraktijk.

Hublab1: Liferay vs. Sharepoint

In het kader van het eerste Hublab-project is Liferay al eens vergeleken met Sharepoint (en Sakai). Dit was primair een technisch-functionele vergelijking en keek nadrukkelijk naar de vraag hoe het betreffende pakket ondersteund zou moeten en kunnen worden door de ICT-afdeling van het IISG. Eén van de evaluatie-criteria was dat het pakket met relatief weinig ondersteuning geïmplementeerd zou kunnen worden. Ook is gekeken in hoeverre het mogelijk was om zelf extra functionaliteit toe te voegen en of het pakket eenvoudig te beheren was.

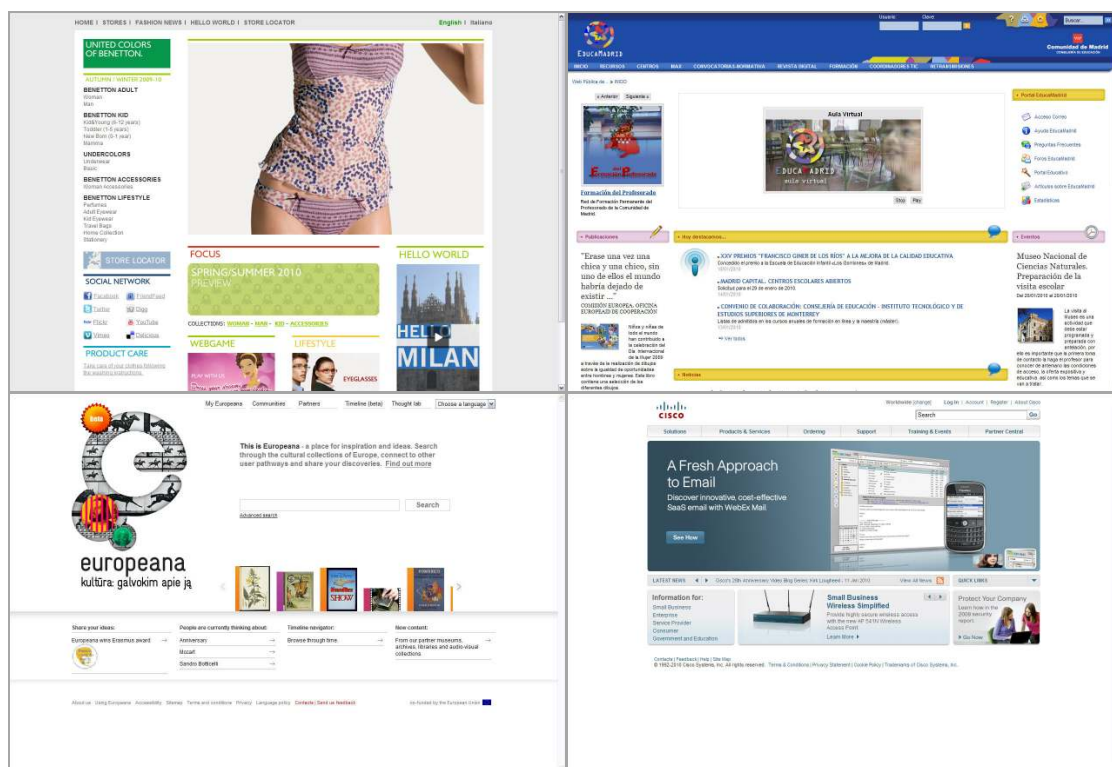
Sharepoint bleek in de test een zeer uitgebreid pakket te zijn, met een gedegen uiterlijk dat voor de meeste gebruikers van Microsoft producten erg herkenbaar zou zijn. Het pakket had echter ook hoge hardware eisen en implementatie zou opleiding van beheerders noodzakelijk maken. Liferay werd ook als een zeer compleet pakket beoordeeld, dat vanwege de eenvoudige opzet snel eigen gemaakt zou kunnen worden. Het beheer van dit pakket werd als eenvoudig aangeduid en alleen de integratie binnen de kantoorautomatisering van het IISG werd als een potentieel probleem gezien.

Uiteindelijk is, de tijdsdruk en gestelde eisen in ogenschouw nemend, gekozen voor Liferay. De potentie van Sharepoint werd toen (begin 2008) hoger ingeschat. De talrijke (potentiële) functies zouden het mogelijk maken om Sharepoint uiteindelijk als een breder instrument in te zetten. De kosten verbonden aan het pakket bleken echter te hoog en Liferay voldeed aan de wensen zoals die er toen lagen. Intussen is na een pilotperiode van ruim een jaar niet alleen besloten om Liferay aan te houden als *collaborative software* voor de collaboratories die gelieerd zijn aan het IISG

(<https://collab.iisg.nl>), maar ook wordt Liferay nu in het instituut gebruikt voor interne samenwerkingsprojecten (<https://projects.iisg.nl>).

Hublab2: gebruik Liferay

De aanvankelijk beperkt geachte functionaliteit van Liferay – in vergelijking met Sharepoint – wordt nu niet meer als een tekortkoming gezien. Niet alleen ontwikkelt het pakket zich snel en heeft de nieuwste versie veel nieuwe functionaliteiten, maar ook is gebleken dat de doelgroep van historici bij voorkeur met eenvoudige en overzichtelijke software werken. Het feit dat veel grote organisaties in de private en publieke sector momenteel Liferay gebruiken (zie Figuur 1) lijkt daarnaast de continuïteit en innovativiteit van het pakket alleen maar te ondersteunen.

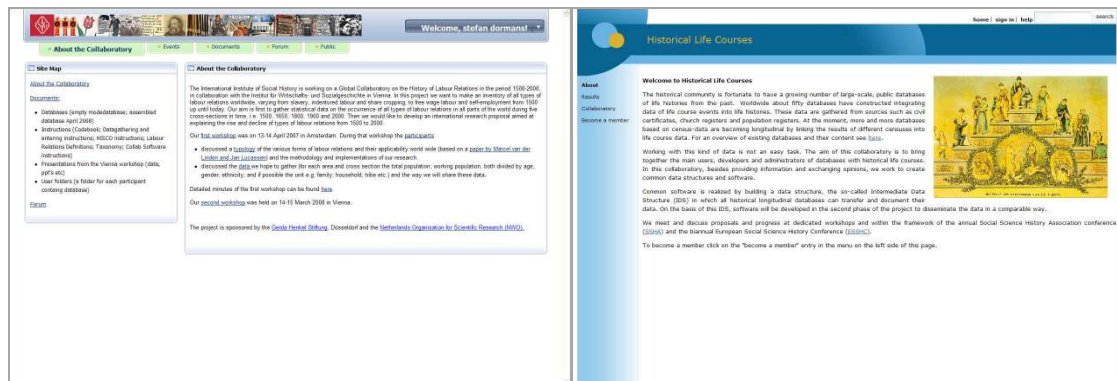


Figuur 1: Websites op basis van Liferay¹

Mede omdat Liferay een *open source* pakket is, kan de interface eenvoudig en nagenoeg volledig aangepast worden aan de eigen eisen. De uiteindelijke interface van de nieuwe versie van Liferay lijkt dan ook in niets meer op de pilotversie (zie Figuur 2) en verschilt daarnaast ook zeer sterk van de standaardopmaak van het pakket. Bij de herinrichting van Liferay hebben we ons uiteindelijk laten leiden door de *usability* (bruikbaarheid) van Liferay. *Usability* gaat over efficiëntie, effectiviteit en het gemak waarmee iets te leren is. In de praktijk betekent dit dat we onder andere hebben gekeken naar: doorzoekbaarheid en begrijpelijkheid; de helderheid van de tekstuele en visuele

¹ Zie www.benetton.com, www.educa.madrid.org, www.europeana.eu, en www.cisco.com

informatie om de weg te vinden; de structuur: of deze intuïtief doorzien kan worden; in hoeverre teksten, structuur en vormgeving aansluiten bij de doelgroep; in welke mate de doelgroep zich aangesproken voelt. In bijlage 2 zijn de daadwerkelijke functionaliteiten alsmede het gebruik van deze nieuwe versie van Liferay reeds uitgebreid besproken. Hier zal nu vooral aangegeven worden hoe we getracht hebben de *usability* van de nieuwe versie van Liferay te vergroten.

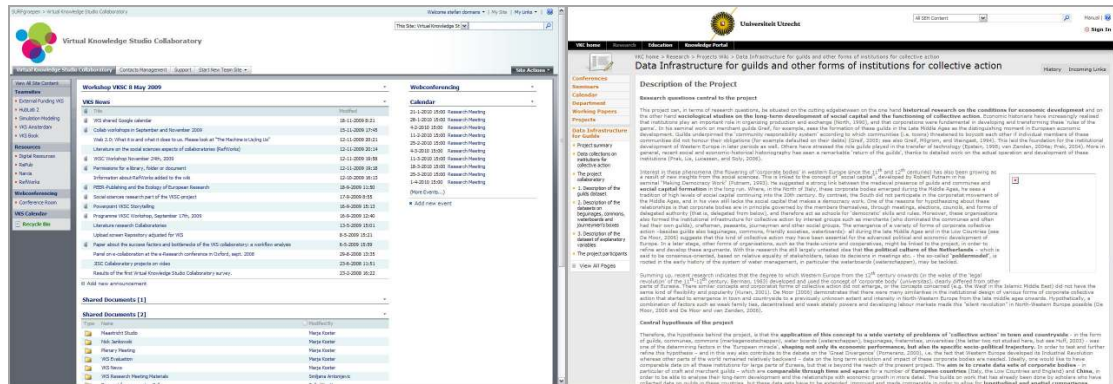


Figuur 2: Pilotversie en huidige versie Liferay (collab.iisg.nl)

Om de doorzoekbaarheid en begrijpelijkheid van Liferay te vergroten is gekozen voor een vormgeving die in structuur vergelijkbare is met de site van het IISG (een site die de meeste collaboratory-leden regelmatig bezocht hebben): het hoofdmenu staat aan de linkerkant van het scherm, een *banner* bovenaan het scherm toont waar je je bevindt, en een submenu met algemene functionaliteiten staat rechtsboven de *banner*. In tegenstelling tot de IISG-site is er verder gekozen voor twee in plaats van drie kolommen, waardoor de centrale tekst of ‘portlet’ (zoals de *document library* of het *forum*) nog prominenter in beeld zijn. Ten aanzien van de helderheid van de tekstuele en visuele informatie om de weg te vinden, en de bijbehorende vraag of er voldoende feedback is, is er voor gekozen om zo weinig mogelijk keuzes in ieder scherm aan te bieden. Door een overzichtelijk menu met maximaal acht keuzes en enkele standaard opties (zoals ‘sign in’, ‘home’, ‘help’ en ‘search’) aan te bieden, is getracht om de gebruiker niet te verwarren door een teveel aan keuzes. Tevens is er gekozen voor een hoofdmenu met een stabiele locatie en een stabiele indeling, een punt dat in de pilotversie nog slechts uitgewerkt was, zodat de gebruiker niet snel zal verdwalen. Dit alles is gedaan met de ervaringen van Hublab1 en onze kennis van de doelgroep in het achterhoofd. De sociale en economische historici die gebruik maken van collab.iisg.nl hebben over het algemeen een relatief geringe expertise over en interesse in dergelijke *collaborative software*. Tevens bleek tijdens de eerste fase al snel dat met name de intuïtiviteit van de pilotversie van Liferay te wensen overliet. Op basis van commentaren van gebruikers is voor de nieuwe versie dan ook juist gestreefd naar een structuur die intuïtief doorzien kan worden. Daarbij is het menu aangepast aan begrippen die door de doelgroep gebruikt worden en hebben de onderzoekers zelf voor de content gezorgd. Het algemene streven naar een zakelijke, *no-nonsense* uitstraling, met alleen de meest basale functionaliteiten en de afwezigheid van visuele of functionele tierelantijnen, was eveneens gebaseerd op directe feedback van de doelgroep.

Hublab2: gebruik Sharepoint

Voor inzicht in het gebruik van Sharepoint ten tijde van het Hublab2-project, kijken we hier nader naar een tweetal gebruikservaringen met Sharepoint: ten eerste, de ervaringen en overwegingen van de Gilden-collaboratory om over te stappen van Liferay naar Sharepoint, en ten tweede, onze (Dormans en Kok) eigen ervaringen in het gebruik van Sharepoint als samenwerkingstool binnen de Virtual Knowledge Studio. In Figuur 3 is de interface van de homepage van beide op Sharepoint draaiende websites te zien.



Figuur 3: VKS-collaboratory en Gilden-collaboratory in Sharepoint²

In de loop van 2008 besloot de Gilden-collaboratory om over te stappen van de toenmalige pilotversie van Liferay – die was ontwikkeld in het kader van het Hublab1-project – naar Sharepoint – zoals die op dat moment werd ontwikkeld en aangeboden door de Universiteitsbibliotheek van Utrecht in het kader van het *Virtual Knowledge Centre*-project. Volgens de projectleider van de Gilden-collaboratory was deze overstap gebaseerd op het feit dat Sharepoint op drie punten beter scoorde dan Liferay: duidelijkheid, gebruiksvriendelijkheid en stabiliteit. Duidelijkheid en overzichtelijkheid hadden in deze vergelijking met de manier van inrichten te maken en de keuzes die daarbij gemaakt werden. Daarbij werd specifiek aangegeven dat het taalgebruik en de inrichting van Sharepoint meer gericht was op gebruik en tevens overzichtelijker was, terwijl Liferay vooral gericht leek te zijn op de ontwikkelaar. Het derde punt, stabiliteit, was primair gerelateerd aan het feit dat Liferay een *open source* pakket was. Van een dergelijk pakket kon men verwachten dat er vaker sprake was van instabiliteit en in de pilotfase bleek dit ook verschillende keren (o.a. tijdens een presentatie op een conferentie) het geval te zijn. In dergelijke gevallen bleek de snelheid van de technische ondersteuning onvoldoende. Na de overstap naar Sharepoint was met name het rechtensysteem niet altijd even helder. Verschillende gebruikers met een gelijke status bleken uiteindelijk niet altijd dezelfde rechten te hebben. De exacte oorzaak daarvan was vervolgens onduidelijk. Het navigeren leverde verder geen problemen op. Het huidige gebruik van de VKC (januari 2010) wordt door de projectleider als zeer intensief omschreven. Twaalf mensen werken momenteel wereldwijd en zonder problemen aan het project met behulp van Sharepoint. Tevens

² Zie https://www.surfgroepen.nl/sites/VKS_C/default.aspx en <http://vkc.library.uu.nl/vkc/seh/> (de VKS-website wordt overigens alleen voor interne doeleinden gebruikt en is dus niet toegankelijk voor externen).

worden nieuwe mogelijkheden ontwikkeld door de UU, zoals een koppeling met een Dataverse-omgeving waar gemeenschappelijke datasets veilig en duurzaam kunnen worden opgeslagen opslaan. Bij het publiekelijk toegankelijke deel van de Gilden-collaboratory valt overigens op dat de meest recente updates van april 2009 zijn en dat diverse plaatjes niet meer zichtbaar zijn, hetgeen voor een buitenstaander een zekere inactiviteit suggereert.

Ons (Dormans en Kok) eigen gebruik van Sharepoint in de VKS-collaboratory begon in principe reeds in 2008 met enkele workshops en introducties, maar kwam pas echt van de grond in mei 2009 toen we tijdens een workshop een teamsite hebben ingericht voor het Hublab2-project. Via die teamsite hebben we enkele centrale documenten van het project gedeeld, maar vooral ook gezamenlijk een artikel geschreven voor het tijdschrift *Historical Methods*. Daarnaast hebben we allebei in de loop van 2009-2010 geparticipeerd in een boekproject van de VKS, waarvoor we ieder een hoofdstuk schrijven met andere collega's binnen de VKS. In het kader van het boekproject werd naast alle conceptversies van hoofdstukken ook de meer algemene projectinformatie via Sharepoint gedeeld. In beide gevallen – Hublab2 en het boekproject – gebruikten we Sharepoint echter niet als veel meer dan een *document sharing environment*. Dit kwam deels omdat meer geavanceerde functies, zoals versiebeheer, in praktijk niet altijd even helder en stabiel werkten. Uiteindelijk is altijd gekozen om nieuwe versies van documenten een nieuwe naam te geven en slechts in enkele instanties zijn er meerdere versies van één bestand gemaakt. In dat opzicht lijkt google-docs een eenvoudiger en meer laagdrempelige variant te zijn. Tevens bleek het verplaatsen van een document niet eenvoudig, zeker als de documentnaam niet voldeed aan de richtlijnen (zoals geen spaties). Het navigeren vergt tenslotte veel geklik, waarbij het met name onpraktisch is dat een documentenmap op een homepage niet dezelfde mogelijkheden (zoals inzicht in de *version history*) biedt als een documentenmap in een *shared documents*-pagina. De drukke interface zorgt in dit verband voor veel verwarring en ondanks de aanwezigheid van *bread crumbs* voel je je als gebruiker regelmatig verloren in Sharepoint.

Vergelijking Liferay en Sharepoint

Het uitgangspunt voor deze vergelijking tussen Liferay en Sharepoint was de vraag in hoeverre het pakket aansluit bij de onderzoekspraktijk. Een algemeen punt dat in dit kader eerst gemaakt dient te worden is dat in geen van de gevallen die wij hebben meegemaakt, de *collaborative software* intensief gebruikt werd in de onderzoekspraktijk. Verreweg de meeste bezoekers bezochten Liferay of Sharepoint incidenteel en het gebruik was vervolgens in de meeste gevallen basaal: beide platforms worden primair gebruikt voor het delen van bestanden.

Kijkend naar onze ervaringen met Liferay dan is in het kader van Hublab2 juist getracht de *usability* te vergroten, waarbij specifiek rekening gehouden werd met de eisen van de primaire gebruikers: onderzoekers in de sociale en economische geschiedenis. Deze doelgroep had expliciet behoefte aan een duidelijke, intuïtieve en simpele omgeving – mede vanwege het incidentele en basale gebruik van de software – en in de nieuwe versie van Liferay was het relatief eenvoudig om deze te ontwikkelen. Het bleek redelijk eenvoudig voor onze ontwikkelaars om aan de meeste wensen van de gebruikers tegemoet te komen.

We hebben zelf geen ervaring met het verbeteren van de *usability* van Sharepoint, dus op dat punt kunnen we geen vergelijking maken, maar in de SURF-versie van Sharepoint die we zelf hebben gebruikt in de VKS-collaboratory was juist de *usability* een van de mindere punten. De structuur die Sharepoint standaard biedt is complexer dan de interface die we intussen gewend zijn van Liferay en dat is met name lastig bij incidenteel gebruik; altijd is het weer even opnieuw zoeken hoe het ook alweer werkt. In de onderzoekspraktijk is dit punt, wat ons betreft, van doorslaggevend belang. Het succes van het gebruik van Sharepoint (of Liferay) lijkt in deze pioniersfase niet bepaald te worden door de veelheid aan functionaliteiten die het gebundeld aanbiedt, maar de *usability* van het platform. Zelfs een vrij geavanceerde groep gebruikers binnen de sociale wetenschappen en geesteswetenschappen als de leden van de VKS gebruiken Sharepoint vooral als een *document sharing environment*. Op dat gebied lijkt dit soort *collaborative software* een duidelijke meerwaarde te hebben voor de onderzoekspraktijk. De manier waarop Sharepoint momenteel ingericht is, lijkt echter niet uit te gaan van dat gegeven en het kost dan ook nog steeds relatief veel moeite om in Sharepoint snel bij de juiste documenten(folders) te komen. Ondanks dat de versie van Liferay waar wij mee werken op dit punt beter scoort, blijft ook daar het punt dat het platform primair een virtuele ruimte voor het uitwisselen van documenten blijft.

Collaboratories: licentieregeling voor het delen van data

Binnen het Internationaal Instituut voor Sociale Geschiedenis (IISG) is er een groeiende behoefte aan duidelijke richtlijnen inzake eigendomsrechten rond het delen van data. In het kader van het Hublab2-project zijn verschillende mogelijkheden op het terrein van open data licenties onderzocht. Op basis van dat onderzoek is in de eerste fase van dit project een beredeneerd voorstel gedaan voor een licentieregeling ten behoeve van alle collaboratories. Dit voorstel is op 27 mei 2009 besproken met de diverse betrokkenen van het IISG (de trekkers van de diverse collaboratories en vertegenwoordigers van de onderzoeksafdeling en de afdeling digitale projecten). Vervolgens zijn eind 2009 door het IISG enkele richtlijnen uitgewerkt, naar aanleiding waarvan een gekozen licentieregeling door de diverse collaboratories geïmplementeerd kan worden.

Uitgangspunt van dit voorstel is dat co-creatie van historische databases, zoals die momenteel vorm krijgen in de diverse collaboratories, duidelijke afspraken omtrent intellectueel eigendomsrecht vereist. Een ander uitgangspunt is dat, op basis van de wensen van de gebruikers, de doelstellingen van het IISG en de beschikbare licenties, zal gekozen worden voor een open data licentie om de verzamelde data publiek toegankelijk te maken. Inhoudelijk sluit dit voorstel expliciet aan bij de publicatie "Hergebruik van materiaal in onderwijs- en onderzoekomgevingen" van Paul Keller (Creative Commons Nederland) en Wilma Mossink (SURF).¹

Databankenrecht

De rechten die er op data en databanken rusten vormen in eerste instantie natuurlijk een bescherming van de makers van de betreffende data(bank). Deze rechten kunnen in een onderzoeksomgeving echter ook als een belemmering ervaren worden omdat ze een obstakel voor (her)gebruik kunnen vormen. Zoals onderstaand opsomming² aangeeft kunnen er diverse rechten verbonden zijn aan databanken. Hierbij dient overigens aangetekend te worden dat de betreffende databank voldoende 'creatief en origineel' dient te zijn en dat er sprake is van een substantiële investering in de samenstelling en/of het onderhoud van de databank om daadwerkelijk aanspraak te kunnen maken op deze rechten.³

- Copyright: both for the selection and arrangement of the database contents and over the contents of the database itself (the data), though factual information will generally not be protected by copyright;
- Database rights: the European Union's Database Directive requires member states to implement a "sui generis database right" covering the extraction and re-utilisation of the contents of protected databases;

¹ <http://www.surffoundation.nl/smartsite.dws?bw=&ch=def&id=14142>

² <http://www.osbr.ca/ojs/index.php/osbr/article/view/516/475>

³ Het is vooralsnog niet duidelijk in hoeverre de digitalisering van een bestaande dataset ook als een *substantiële investering* wordt beschouwd.

- Contract: contractual obligations about what users can and can't do with a database and its contents can also be used to provide for protection;
- Other rights: rights such as trade secrets and laws of unfair competition can also protect databases.

In Europa is de juridische context die van het databankenrecht. Officieel is dit een Europese richtlijn (uit 1996), die in de diverse nationale wetgevingen op diverse manieren is uitgewerkt. In principe draait het databankenrecht om bescherming van de rechten van de producent (o.a. exclusief recht tot exploitatie) en geldt het voor 15 jaar. Iedereen die een creatieve bijdrage aan de databank levert, wordt beschouwd als een medeproducent van die databank en heeft dus de bijbehorende rechten over zijn of haar deel van de databank. In de VS en andere delen van de wereld kent men geen vergelijkbaar databankenrecht en zijn de producenten van databanken hooguit beschermd door *copyright*.

Rechthebbende(n)

Bij de ontwikkeling van een dataset zijn vaak meer dan één wetenschapper betrokken. Nu is het zo dat, in principe, eenieder die een creatieve bijdrage heeft geleverd aan de totstandkoming van een dataset rechten bezit op de zelf gemaakte data. Praktisch gezien betekent dit dat, indien de productie of publicatie van de afzonderlijke datasets minder dan 15 jaar geleden is, afspraken gemaakt dienen te worden met alle wetenschappers die creatief hebben bijgedragen aan de dataset. 'Creatief' is hier logischerwijs het cruciale woord: het invoeren van de primaire bronnen in een door andere bedachte database structuur lijkt hier niet onder te vallen, maar het mede bedenken van de structuur van de database waarschijnlijk wel.

Als is vastgesteld dat iemand een creatieve bijdrage heeft geleverd aan een dataset, dan is de vraag of deze persoon ook een rechthebbende inzake de dataset is nog niet beantwoord. Zo is het niet vanzelfsprekend dat individuele wetenschappers alle rechten bezitten op de data die zij verzameld hebben. Een voor de hand liggend voorbeeld is het exploitatierecht, dat in sommige gevallen in het bezit is van een uitgever. Het kan echter ook zo zijn dat de data in handen zijn van de organisatie voor wie de wetenschapper de data verzameld heeft. In Nederland zijn er geen duidelijk afspraken over het eigendomsrecht. Bij buitenlandse universiteiten en instituten kunnen hier echter wel afspraken over bestaan en mogelijk moet men in die gevallen afspraken op organisatorisch niveau maken.

Aggregatie en co-creatie

Het is belangrijk om een onderscheid te maken tussen data die in het verleden verzameld is en data die in het kader van de *collaboratory* verzameld gaat worden. In dat laatste geval kunnen er binnen de groep afspraken gemaakt worden hoe men in toekomst met de datarechten zal omgaan. Bij reeds eerder verzamelde data (*secondary use*) dient men, zoals eerder aangegeven, de producenten van de data te achterhalen en ad hoc afspraken te maken. Het is hierbij van belang om boven tafel te krijgen of de betrokken wetenschappers de rechten op een dataset ook daadwerkelijk bezitten of dat ze alleen bezit claimen. Een manier om de rechten rond data te regelen – zowel binnen een *collaboratory* als bij *secondary use* – is met behulp van open data licenties.

Open data licenties

Het achterliggende idee van open data licenties is dat wetenschappelijke kennis – die grotendeels met publieke middelen gefinancierd wordt – voor een breed publiek vrij toegankelijk te maken. Vandaar dat dit soort licenties zich vooral richten op het afstaan van rechten op de data. Afgezien van het gebod to correct citeren van de maker van de data⁴, streven open data licenties vooral naar het rechtenvrij maken van datasets. In dat kader is Creative Commons de Zero Waiver (CC0) aan het ontwikkelen.

CC0 is a protocol that enables people to WAIVE to the fullest extent possible under applicable copyright law all rights they have and associate with a work so it has no (or minimal) copyright or neighboring rights restrictions attached to it. To the extent the waiver is not legally effective in any jurisdiction, then the protocol takes the form of a nonexclusive worldwide license to exercise all copyright and neighboring legal rights in the work.⁵

Open data licenties zijn een vrij recent fenomeen en nog volop in ontwikkeling. Volgens het rapport van Keller en Mossink over hergebruik van materiaal in onderzoeksomgevingen, werden tot 2007 Open Content licenties voldoende geschikt geacht om toegang en vrij (her)gebruik van wetenschappelijke onderzoeksdata te waarborgen. Science Commons heeft een dergelijke aanbeveling echter in 2007 ingetrokken en een 'Protocol for Implementing Open Access Data' gepubliceerd. Doel van dit protocol is het realiseren van interoperabiliteit van open wetenschappelijke data. Licenties die aan de voorwaarden van dit protocol voldoen kunnen van Science Commons het keurmerk 'Science Commons Open Access Data Mark' verkrijgen. Data(banken) die onder een licentie met dit keurmerk beschikbaar gesteld zijn, kunnen vrij gecombineerd worden met alle andere data(banken) die onder een licentie met dit keurmerk beschikbaar gesteld zijn.

Op dit moment zijn er nochtans nog geen licenties (of andere juridische tools) aan te wijzen die het keurmerk hebben verkregen, aldus Keller en Mossink. Wel zijn er twee juridische instrumenten die conform het protocol ontwikkeld zijn en naar verwachting voor het keurmerk in aanmerking zullen komen: de Creative Commons Zero Waiver 1.0 Universal (CC0 Waiver) en de Open Data Commons Public Domain Dedication and Licence (ODC PDDL). Beide instrumenten voldoen aan de kernvoorwaarden van het Science Commons protocol, maar in hun uiteindelijke advies zijn Keller en Mossink enigszins ambivalent. In eerste instantie geven ze aan dat de ODC PDDL beter geschikt is voor gebruik in de Nederlandse onderzoekssituatie dan de CC0 Waiver (2008: 24), terwijl ze in de uiteindelijke aanbeveling het volgende stellen:

Voor het (her)gebruik van data beveelt SURFdirect een licentie aan die voldoet aan het Science Commons Protocol for Implementing Open Access Data waardoor de interoperabiliteit van open wetenschappelijke data wordt gerealiseerd. Hiervoor zouden twee licenties in aanmerking kunnen komen: de Open Data Commons Public Domain Dedication and Licence, en de Creative Commons

⁴ De voorwaarden omtrent naamsvermelding voldoen omvatten niet alleen het vermelden van de naam van de maker of een andere door de licentiegever aangewezen persoon of instelling (zoals Universiteit of Journal) maar ook de verplichting om afgeleide werken duidelijk als zodanig aan te wijzen.

⁵ <http://wiki.creativecommons.org/CCZero>

Zero Waiver. Gezien de nauwe verwevenheid met de Creative Commons licentie die SURFdirect aanbeveelt voor het hergebruik van onderwijs- en onderzoeksmateriaal meent SURFdirect dat deze laatste licentie geschikt aanbevelenswaardig zou kunnen zijn voor het gebruik met wetenschappelijke data. Echter omdat de licentie slechts in concept beschikbaar is zal SURFdirect in een later stadium deze aanbeveling opnieuw bezien en over de uitkomsten hiervan het hoger onderwijsveld nader informeren. (Keller en Mossink, 2008: 31)

Richtlijnen IISG

Begin 2010 heeft het IISG een rapport⁶ gepubliceerd waarin richtlijnen worden gegeven voor het bewaren van onderzoeksdata. Voortbouwend op het Hublab2-voorstel voor een licentieregeling ten behoeve van alle collaboratories van het IISG, die op 27 mei 2009 was besproken met de belanghebbenden, besteedt het rapport specifiek aandacht aan de licentiekwesties rond data die verzameld worden in een collaboratory. Zoals uit onderstaand citaat blijkt, is het advies aan de onderzoekers om binnen collaboratories afspraken te maken op basis van de Zero Waiver van Creative Commons.

In the above, we have assumed that the research team in a collaboratory has pooled the data and has, at some point, transferred the common dataset and the various releases to a trusted repository or data archive. The archive provides access to the datasets for re-use by researchers who can enhance and redistribute the data themselves. This assumption may be too idealistic and naïve. Several hurdles, associated with data rights, have to be taken. In principle, everybody who has made a 'creative' contribution to a dataset has rights on the data he or she contributed. 'Creative contribution' is to be seen here as participation in the construction of the data structure, not just data-entry. In the case of pooling already existing data from different countries, the (not fully crystallized) international laws on databases has to be taken into account. According to the European Directive regarding the protection of databases (96/9/EC), for data which have been produced or published less than 15 years ago, consent has to be given by the original creators of the data. To complicate things further, rights on data may belong to the organizations the individual scholars work for. In the Netherlands, however, the individuals who created the data own by default, the right to the data. When collaboratories consisting of international teams start pooling data, it is strongly advised that they check and clear the ownership of the rights on the data. A new situation emerges, when collaboratories create – jointly - a new dataset. This gives more leeway to reach an agreement on how to handle data rights. We advise strongly that – after a period of embargo in which the team can collect and harmonize the data and do research on them – an open data license is adopted. In such a license, right on data are waived and data are provided freely to the public, on the condition of proper citation. Although not fully developed yet, the Zero Waiver of Creative Commons (CC0) seems most appropriate to data collected within collaboratories.

⁶ Doek et al. (2010) IISH Guidelines for preserving research data: a framework for preserving collaborative data collections for future research, International Institute of Social History, Amsterdam. Dit rapport is geschreven in het kader van de SURF Call "Collectioneren van data" van het SURFshare programma "Datacuratie en digitale duurzaamheid".

Uncovering Collaboratories in Social and Economic History¹

Stefan Dormans

The Virtual Knowledge Studio for the Humanities and Social Sciences (Royal Netherlands Academy of Arts and Sciences), the Netherlands.

stefan.dormans@vks.knaw.nl

Abstract. Research on collaboratories is mainly focussed on the natural sciences. This paper introduces an ethnographic project which studies collaboratories in the social sciences and humanities, i.e. in the subdisciplines of social and economic history. The focus of the paper is on the ways in which e-research projects like collaboratories make explicit many processes that were previously implicit and the problems associated with this in the daily practices of social and economic historians.

Introduction

Two recently published books on collaboratories present case studies in astronomy, space science, physical sciences, biological and health sciences, and earth and environmental sciences (Olson et al., 2008; Shrum et al., 2007). The absence of case studies on collaboratories in the social sciences and humanities (SS&H) in both books reflects a tendency to the natural sciences in the literature on collaboratories. In part, this can be explained by the fact that there are substantially more collaboratories in the natural sciences compared to SS&H and that these collaboratories also have a longer tradition (Finholt, 2003; Sonnenwald, 2007). Moreover, funding and support for collaboratories in the natural sciences is more substantial than for collaboratories in SS&H. However, the recent growth of collaboratories in social sciences and humanities has increased the demand for knowledge about the specificities of academic collaboration at a distance in these academic fields.

This paper looks at the subdisciplines of social and economic history, and more precisely at a number of collaboratories which are related to the International Institute of Social History (IISH) in Amsterdam, the Netherlands. The paper is based on the research project 'Socio-technological aspects of collaboratory projects in social and economic history' of the Virtual Knowledge Studio for the Humanities and Social Sciences. This project entails an ethnographic study of the practices, risks and opportunities of the implementation of the collaboratories of the IISH. In this paper, I discuss some of the specific features of collaboratories in SS&H and aim to enrich the current debate with case studies beyond the confines of the natural sciences. In doing so, special attention will be paid to

¹ Dormans, S. 2009. Uncovering collaboratories in social and economic history. Proceedings of the 5th International E-Social Science Conference of the National Centre for E-Social Science (Cologne, Germany, June 24-26).

the ways in which e-research projects make explicit many processes that were previously implicit and the problems associated with this.

In the following, I first discuss some of the main characteristics of the VKS ethnographic study and then I go into the 'Global Hubs for Global History' strategy of the IISH. This long term strategy aims to establish global networks of experts on specific topics in historical research and the collaboratories are an essential element in this strategy. Consequently, I elaborate on the particularities of the subdisciplines of social and economic history and show how these relate to the current efforts of developing historical collaboratories. In the penultimate section, the discussion focuses on how collaboratories require historians to be explicit about practices which traditionally are kept implicit and the problems associated with this.

An ethnographic approach to collaboratories

Various scholars have argued that it is important to acknowledge disciplinary differences in integrating e-research infrastructures and tools in scholarly practices (e.g. Fry, 2006; Jankowski, 2007). Furthermore, research has shown that the introduction of infrastructures and tools which do not complement, or are incompatible with, existing policies and practices will not increase scientific collaboration (Sonnenwald, 2007: 660). Consequently, it is not only essential to understand the differences between collaboratories in the natural sciences and in SS&H, but also to explore how in specific disciplines within SS&H scholars from different locations and institutes work together on common projects. The ethnographic approach used here to study the collaboratories in social and economic history, makes it possible to zoom in on the specificities of these subdisciplines, while it simultaneously enables a close examination of the daily academic practices of the collaborators.

The fieldwork for this project covers the period from early 2008 until the end of 2009 and combines various ethnographic techniques, like non-participant observation, text analysis and interviewing. An important part of the fieldwork is the attendance of both the formal and the informal parts of the international workshops organised by several collaboratories. In addition, various members of the collaboratories and other people related to the Global Hubs project are interviewed. Furthermore, the online interaction between the members of the various collaboratories is studied closely. Finally, all the relevant documents on and by the collaboratories are examined. Since my own institute, the Virtual Knowledge Studio (VKS), is located in the building of the IISH, a lot of my fieldwork can be done relatively unobtrusive. However, my role is not only that of the non-participating observer. In line with the wider policy of the VKS, the project aims to integrate critical analysis and design-oriented research in close cooperation with researchers and information technology experts. One of the core features of the Virtual Knowledge Studio is its ambition to support researchers in the humanities and social sciences in the creation of new scholarly practices, as well as in their reflection on e-research in relation to the development of their fields (Beaulieu and Wouters, forthcoming). This is done in an integrated approach which combines critical analysis with design-oriented research. So, on the one hand the VKS aims to reflect on e-research practices in the humanities and social sciences, and on the other hand, it has the specific aim to intervene in these practices. In the case of this project, this results, among other things, in helping the collaboratories to develop a

policy on (open) data licences and in supporting the process of selecting, using and upgrading their collaborative software.

Global Hubs for Global History

In recent years, the IISH has decided to strengthen its already leading role in the field of social and economic history by developing international collaborative projects on Global Labour History and on Global Economic History. Under the heading 'Global Hubs for Global History', a strategy is developed to establish a network of collaboratories around diverse topics related to life courses, labour and migrant organisations. These collaboratories revolve around groups of experts in these fields who aggregate and co-create datasets for international comparative research. A total of five collaboratories is studied more closely in the framework of the above mentioned ethnographic project: (1) HISCO, (2) Towards a Global History of Life Courses, (3) Global Collaboratory on the History of Labour Relations 1500-2000, (4) Labour Conflicts and New ways in Strike Research, and (5) Constructing a Global Database on Migrant Organizations.² As an illustration, two of these collaboratories are discussed here in somewhat more detail.

The 'Global Collaboratory on the History of Labour Relations 1500-2000' makes an inventory of all types of labour relations worldwide, varying from slavery, indentured labour and share cropping, to free wage labour and self-employment from 1500 up until today. The members of the collaboratory gather statistical data on the occurrence of all types of labour relations in all parts of the world during five cross-sections in time, i.e. 1500, 1650, 1800, 1900 and 2000. These data, including critical annotation and other metadata, are gathered by regional experts from different disciplines with the eventual aim to explain the rise and decline of types of labour relations in the period 1500-2000. At the moment, this collaboratory already covers many areas and periods, but new partners are still added to the group. The HISCO collaboratory has its origins in the mid 1990s and is built around an international classification of historical and contemporary occupations. Comparative research in this field is severely hampered by confusion regarding occupational terminology across time and space, within as well as between languages. To overcome this, the HISCO collaboratory creates an occupational information system that is both international and historical, and simultaneously links to existing classifications used for present-day conditions. Other researchers are invited to join the collaboratory and they can use as well as expand the current classification system.

Just like the other collaboratories in social and economic history, these two groups are, compared to the majority of collaboratories in the natural sciences, relatively small and egalitarian and this is reflected in both their management and their budget. The fact that the collaboratories are based on the complementarities of the individual expertise of the members – each member covers a specific period, topic, or region – distinguishes them from their more hierarchical counterparts in the natural

² Visit <https://collab.iisg.nl> to view the portal to the various collaboratories of the IISH.

sciences.³ The data they collect are of heterogeneous nature, in different stage of digitization and collected and archived in different epistemic and cultural environments. Some collaboratories recently started looking for participants and are in the process of finding common ground, while other collaboratories have already collected data from various regions and periods and are currently analysing and comparing these data. This difference can partly be explained by the fact that some groups have a history which goes back to the period before the formation of the collaboratory. Some collaboratory members, for instance, started working together in the early 1990s. Since the beginning of 2008 they formed a collaboratory and to facilitate the collaboration process, they started using a virtual workspace for data storage and online communication. In the future, the same workspace might also be used for data manipulation and data visualization.

An additional feature of these collaboratories is that the work is done on a voluntary basis. Participation of the individual academics is not based on financial incentives, but mainly on social bonds and academic opportunities. Many members are part of the wider network of the IISH and often have a longer tradition of cooperating with researchers at the institute. What's more, the projects as they are pursued by the different collaboratories all relate to the recently growing subdiscipline of world history. As such, the collaborative projects are at the forefront of contemporary developments in the field of history and have the accompanying academic appeal. The fact that the collaboratories are initiated by the IISH is an additional argument to participate, because of the leading position of the institute in the fields of social and economic history.

Collaboratories in social and economic history

If one studies the particularities of collaboratories in social and economic history, it is important to be aware of the fact that the subdisciplines social history and economic history are not typical for the discipline of history as a whole. On the one hand, social and economic historians are trained as historians and, like most humanities scholars, their research practices are for a relatively large share solitary. On the other hand, they use social scientific theories and methods to gather their largely quantitative historical data and accordingly their practices are more like those of social scientists than those of historians. In the following, both sides of the social and economic historian – the humanities and the social science side – are discussed in more detail, as well as the implications both sides have for collaborative projects.

Although collaborations have long been a feature of historical research, the 'lone historian' remains the dominant paradigm for describing and evaluating the research practices in this academic field. This is reflected most notably in the publication practices of historians. As Steckel (2007: 13-15) shows in his discussion of coauthorship in humanities journals in general and a historical journal in particular, articles in history are predominantly single authored. In the period 1975-2000, the development of the average number of authors per article shows an almost flat time trend and does

³ As Shrum et al. (2007: 106) argue, only particle physics has a distinctive style of collaboration that is described by its members as highly egalitarian. All other collaboratories in the natural sciences they studied had more hierarchical and formalised organisational structures.

not exceed the 1.2 mark. Likewise, Cronin (2003: 9-10) argues that the essentially individualistic model of scholarly production in history – which above all relies on monographs – continues to be the basis for career advancement.

Nevertheless, just like in other academic disciplines, there is a growing trend and increasing pressure towards collaboration in social and economic history. General trends like funding regimes which promote academic collaboration (Cummings and Kiesler, 2005; Hine, 2006) and the rapid development and increased use of information and communication technologies (ICTs) that facilitate interaction and sharing (Sonnenwald, 2007), also influence the subdisciplines of social and economic history. Moreover, social and economic historians also seem to subscribe to the idea that in collaborative research more ground can be covered, more minds are brought to bear, more cross-checking and comparing can be done, more burdens are shared out, and wider perspectives are taken (cf. Woods et al., 2000: 85-86).

In their collaborative practices, the humanities and social science side of the social and economic historians combine in a particular constellation. Just like most historians, social and economic historians typically have a specific territory, theme and/or period of expertise and they collect their data on the basis of this expertise. Consequently, they mainly produce relatively small datasets which are often not easily comparable to other sets on similar topics, because the design of the dataset reflects a distinctive process of collection and analysis. Trying to understand such datasets without expertise on this territory, theme or period – or without knowledge of the particularities of the data collection process – is very difficult. Roth et al. (2008: 82) articulate this idea eloquently in their discussion of a collaborative research project on the history of violence:

Historians must be aware of ambiguities in both the original evidence and the judgments past historians have made about that evidence, or they cannot construct the past. (...) Unless the evidence is available in its original form or in careful extracts transcribed by previous scholars, historians will not have the information they need to reinterpret the events they have not researched personally.

Likewise, de Moor and van Zanden (2008: 67) argue that more is needed to practice collaborative history than raw data alone. They refer in this respect to the work of Allen et al. (2005), who claim that, in addition to indispensable metadata, working with data from sources that originated in different parts of the world requires other knowledge: the context in which the information was collected, the method of collection, and possible problems in the interpretation of the data. Consequently, it is necessary to have intensive communication and interaction among researchers when co-creating large databases. According to de Moor and van Zanden (2008: 68), collaboratories have the potential to facilitate the mutual exchange, peer review, and improvement of data quality: 'not only does it become possible to amass data on the same theme, thus creating new opportunities for comparative research, but interaction with other researchers also offers more background information on the data and their possible uses'. So, collaboratories go beyond mere scale advantages, these authors argue. It is not only about collecting historical data on different periods and different regions. In their view, collaboratories also have the potential to improve the quality of the data through extensive peer review and through intensive communication and interaction among researchers.

Making the implicit explicit

At the heart of this idea of intensive communication and interaction is the need to exchange the knowledge of the experts on the collection, construction and analysis of their individual dataset. As such, this particular interpretation of what a collaboratory can and should do, acknowledges and reinforces the idea that data always remains inextricably bound to the contexts and rationales of the researcher. Historical data cannot be disconnected from the scholar who retrieved them from the archives and who consequently represented and analysed them.

Traditionally, this process of data collection and analysis is implicit and invisible in social and economic history. In contrast to sociologists and other social scientists, social and economic historians do not elaborately discuss the steps taken during the research process, nor do they consistently reflect on the methods or methodologies used to collect and analyse their data. The introduction of the collaboratory model challenges this tradition, since the use of ICTs for knowledge generation and exchange requires the transfer of knowledge into certain formats (cf. Heimeriks and Vasileiadou, 2008: 14). For this purpose previously 'tacit' or 'implicit' knowledge, like the expert's knowledge on the collection and analysis of the data, needs to be captured in an explicit way and therefore to be transformed. This visibility and codification of previously informal and tacit knowledge, and the use of ICTs for recording has another implication as well, according to Heimeriks and Vasileiadou (2008: 14): 'The use of ICTs makes the informal processes of research collaboration traceable and visible: processes that took place in the lab corridors before, such as socializing and exchanging pleasantries, decisions about research tasks, disagreements about roles and responsibilities.'

Although it often remains unclear *if* and *to what extent*, for instance, disagreements about responsibilities become more visible through the use of ICT, one effect of e-research projects is that it has the potential to make visible or explicit many processes that were previously invisible or implicit. As Schroeder and Fry (2007) argue, e-research projects do this partly because computerization necessitates, for example, that data are put into a certain format or stored in a certain way or made accessible across a network. In an e-research setting, things like workflow organization, intellectual property issues, and inter-institutional commitments, require standardization of rules and procedures that may previously have been unstated or unformalized or do not apply to non-digital material (Schroeder and Fry, 2007).

The explicit discussion of intellectual property issues in the collaboratories in social and economic history is one apparent example of this. Before, this issue was rarely discussed by historians, but the collaboration between researchers from different institutes and the sustainable storage of their data put intellectual property rights and (open) data licences on the historical research agenda. Because the collaboratories studied here primarily worked together on a voluntary basis, other issues – like inter-institutional commitments – turned out to be less relevant.

Metadata and moments of engagement

In practice, two main methods are used for discussing the collection and analysis of the individual datasets on the one hand, and the coordination of the collective dataset on the other hand. The first method is the development of elaborate rules for annotation and metadata. The second method is the organisation of moments of engagement. In the first instance, one can transfer some of the tacit knowledge about a dataset and thus potentially improve interpretations of data by other researchers (cf. Zimmerman, 2008). However, such a system is very time consuming and it remains unclear whether collaborators think that this extra effort produces enough extra benefits. Until now, some laboratories have developed guidelines on metadata, but most members have not used them for elaborating on their own dataset. An additional problem with metadata is, as Michener and his colleagues (1997, 335) argue, that there is no end to metadata: 'There is no unique, minimal, and sufficient set of metadata for any given data set, since sufficiency depends on the use(s) to which the data are put'. Moreover, Mauthner and Doucet (2008: 981) assert that there are more fundamental limitations on the extent to which contextual, tacit, embodied knowledge and multi-sensory field-based experiences can be articulated, recorded and exchanged in discursive form.

The organisation of moments of engagement takes two different, but related forms. The first way of doing this is the organisation of recurring workshops. However, but the international composition of the laboratories related to the IISH does not allow for more than one or two meetings a year. The various groups have members working in such different countries as India, Germany, Turkey, Great-Britain, Spain, and Uruguay. Bringing these people together on a regular basis is beyond the budget of these laboratories, but having a few workshops a year is insufficient to create systematic discussions on the collection, construction and interpretation of the data. In theory, the virtual workspace are intended to function as an additional space of engagement, but up till now the members of the various laboratories rarely discuss their data and interpretations online. Although there are some positive exceptions, the virtual workspaces are still primarily used for uploading and downloading documents.

Conclusion

Despite the current lack of explicit discussion in these laboratories, developments are still on its way. Most of the groups discussed here are in an early stage of their life cycle and seem to adapt slowly to the demands of the collaborative and geographically dispersed setting. If some kind of discussion would develop in the near future, this could potentially contribute to a significant change of the research practices in social and economic history. As said, these discussions could uncover the often implicit process of collecting, constructing and interpreting historical data and make the whole process more visible and transparent. Consequently, through explicating that which was implicit, through making public what was private, patterns of practice become open for scrutiny and contestation (cf. Berg, 1997: 1086).

However, making the research process more publicly inspectable is difficult to achieve and, as Schroeder and Fry (2007) remind us, e-research projects do not only have the potential to make visible or explicit many processes that were previously invisible or implicit, but they also may hide or

marginalize other research efforts that are not amenable to the process of rendering research into digital formats. Not all knowledge is recordable in easily transferred forms, such as documents (cf. Finholt and Olson, 1997: 28), and as a result it is not easily shared across distance. Therefore, there are limits to what one actually can make explicit.

Acknowledgments

This research has been funded by the 'Strategiefonds' of the Royal Netherlands Academy of Arts and Sciences (www.knaw.nl) and by the SURF, the higher education and research partnership organisation for Information and Communications Technology (ICT). For more information about SURF, please visit www.surf.nl. Furthermore, I want to thank my colleagues at the Virtual Knowledge Studio for the discussions and their useful comments on previous versions of this paper.

References

- Allen, R.C. Bassino, J.P., Ma, D., Moll-Murata, C. and van Zanden, J.L. (2005): *Wages, prices, and living standards in China, Japan, and Europe, 1738–1925*, International Institute of Social History, http://www.iisg.nl/research/jvz-wages_prices.pdf (accessed on 18 May 2009).
- Beaulieu, A. and Wouters, P. (forthcoming): 'E-research as intervention', in: N. Jankowski (ed.) *E-Research: Transformations in Scholarly Practice*, Routledge, London.
- Berg, M. (1997): 'Problems and promises of the protocol', *Social Science & Medicine*, vol. 44, no. 8, pp. 1081-1088.
- Cronin, B. (2003): *Scholarly communication and epistemic cultures*, paper presented at Scholarly tribes and tribulations: How tradition and technology are driving disciplinary change, October 17, Washington.
- Cummings, J.N. and Kiesler, S. (2005): 'Collaborative Research Across Disciplinary and Organizational Boundaries', *Social Studies of Science*, vol. 35, no. 5, pp.703-722.
- De Moor, T. and van Zanden, J.L. (2008): 'Do ut des (I Give So That You Give Back). Collaboratories as a New Method for Scholarly Communication and Cooperation for Global History', *Historical Methods*, vol. 41, no. 2, pp. 67-78.
- Finholt, T. (2003): 'Collaboratories as a new form of scientific organisation', *Economics of Innovation and New Technology*, vol. 12, no. 1, pp. 5-25.
- Finholt, T. and Olson, G. (1997): 'From laboratories to collaboratories: A new organizational form for scientific collaboration', *Psychological Science*, vol. 8, no. 1, pp. 28-36.
- Fry, J. (2006): 'Scholarly research and information practices: a domain analytic approach', *Information Processing and Management*, vol. 42, pp. 299-316.

Heimeriks, G. and Vasileiadou, E. (2008): 'Changes or transition? Analysing the use of ICTs in the sciences', *Social Science Information*, vol 47, pp. 5-29.

Hine, C. (2006): 'Databases as Scientific Instruments and Their Role in the Ordering of Scientific Work', *Social Studies of Science*, vol. 36, no. 2, pp. 269-298.

Jankowski, N. (2007): 'Exploring e-science. An introduction', *Journal of Computer-Mediated Communication*, vol. 12, no. 2.

Mauthner, N. and Doucet, A. (2008): 'Knowledge Once Divided Can Be Hard to Put Together Again: An Epistemological Critique of Collaborative and Team-Based Research Practices', *Sociology*, vol. 42, no. 5, pp. 971-985.

Michener, W., Brunt, J., Helly, J., Kirchner, T. and Stafford, S. (1997): 'Nongeospatial Metadata for the Ecological Sciences', *Ecological Applications*, vol. 7, no. 1, pp. 330-342.

Olson, G., Zimmerman, A., and Bos, N. (eds.) (2008): *Scientific Collaboration on the Internet*, MIT Press, Cambridge, MA.

Roth, R., Eckberg, D.L., Dayton, C.H., Wheeler, K., Watkinson, J., Haberman, R. and Denham, J.M. (2008): 'The Historical Violence Database. A Collaborative Research Project on the History of Violent Crime, Violent Death, and Collective Violence', *Historical Methods*, vol. 41, no. 2, pp. 81-97.

Schroeder, R. and Fry, J. (2007): 'Social science approaches to e-science: Framing an agenda', *Journal of Computer-Mediated Communication*, vol. 12, no. 2.

Shrum, W., Genuth, J., and Chompalov, I. (2007): *Structures of Scientific Collaboration*, MIT Press, Cambridge, MA.

Sonnenwald, D. (2007): 'Scientific Collaboration', in B. Cronin (ed.): *Annual review of information science and technology*, Information Today, Medford, NJ, pp. 643-682.

Steckel, R. (2007): 'Big Social Science History', *Social Science History*, vol. 31, pp. 1-34.

Wasser, J. and Bresler, L. (1996): 'Working in the Interpretive Zone: Conceptualizing Collaboration in Qualitative Research Teams', *Educational Researcher*, vol. 25, no. 5, pp. 5-15.

Woods, P., Boyle, M., Jeffrey, B. and Troman, G. (2000): 'A research team in ethnography', *Qualitative Studies in Education*, vol. 13, no. 1, pp. 85-98.

Zimmerman, A. (2008): 'New knowledge from old data. The role of standards in the sharing and reuse of ecological data', *Science, Technology & Human Values*, vol. 33, no. 5, pp. 631-652.