

CLARIAH Interest Groups Work Plan

Versie: 1.0

Datum: 23 oktober 2020

Omschrijving: Eerste volledige versie op basis discussie met Technisch Comité en Interest Group coördinatoren op 15 oktober 2020

Doel: Feedback van CLARIAH bestuur

1. Doelstelling en vormgeving Interest Groups	2
1.1 Doelstelling	2
1.2 Vormgeving	2
1.2 Verhouding IGs tot CLARIAH	3
1.3 IGs en de CLARIAH Roadmap	4
2. Interest Groups	6
2.1 Devops	7
2.2 Preservation/herbruikbaarheid/repliceerbaarheid	7
2.3 Security & Monitoring*	8
2.4 Audiovisual processing	9
2.5 Text processing	9
2.6 Annotation	10
2.7 Linked Open Data	11
2.8 Workflows	13
2.9 Curation	14
2.10 UI/UX	15
2.11 Tools-to-Data*	15
2.12 Data Management & Search*	16
2.11 Overzichten	17
3 Communicatie	18
3.1 Informatie per IG	18
3.2 Dagelijkse communicatie	19
3.3 Vormen van overleg	19

1. Doelstelling en vormgeving Interest Groups

1.1 Doelstelling

Het doel van de Interest Groups (verder IGs) is om technische ondersteuning te bieden bij de inrichting van de CLARIAH infrastructuur, specifiek met betrekking tot onderdelen ervan die werkpakket/domein overstijgend (CLARIAH-breed) zijn.

1.2 Vormgeving

Op basis van informatie over *praktische uses cases* van onderzoekers (requirements/wensen) verzameld in de werkpakketten, onderzoeken IGs:

- welke prioriteiten er zijn met betrekking tot *functies* die de CLARIAH infrastructuur zou moeten bieden,
- welke *technische oplossingen* hiervoor zijn,
- en uiteindelijk wat de beste *keuze* is, gegeven de technische staat van de infrastructuur en best-practices in het veld¹.

Om de IGs zo doelgericht mogelijk te kunnen laten functioneren binnen CLARIAH wordt de scope van de IGs zo klein mogelijk gehouden: liever kleine maar tastbare resultaten in de infrastructuur dan een brede aanpak op papier.

Elk IG heeft een IG coördinator die verantwoordelijk is voor het coördineren van het werk binnen de IG en de verslaglegging ervan richting het Technisch Comité (TC) en CLARIAH breed. De coördinator zorgt voor een werkplan (roadmap) voor de IG en is verantwoordelijk voor het formuleren van technische keuzes vanuit de IGs.

De coördinatoren zijn verantwoordelijk voor de informatievoorziening naar en afstemming met de WP leiders. Dat kan ook een WP leider zijn van een ander WP dan waar de coördinator organisatorisch deel van uitmaakt.

Alle coördinatoren nemen zitting in het TC. Voorgestelde technische keuzes worden in het TC besproken. Besluiten worden via de CTO gerapporteerd aan het CLARIAH bestuur.

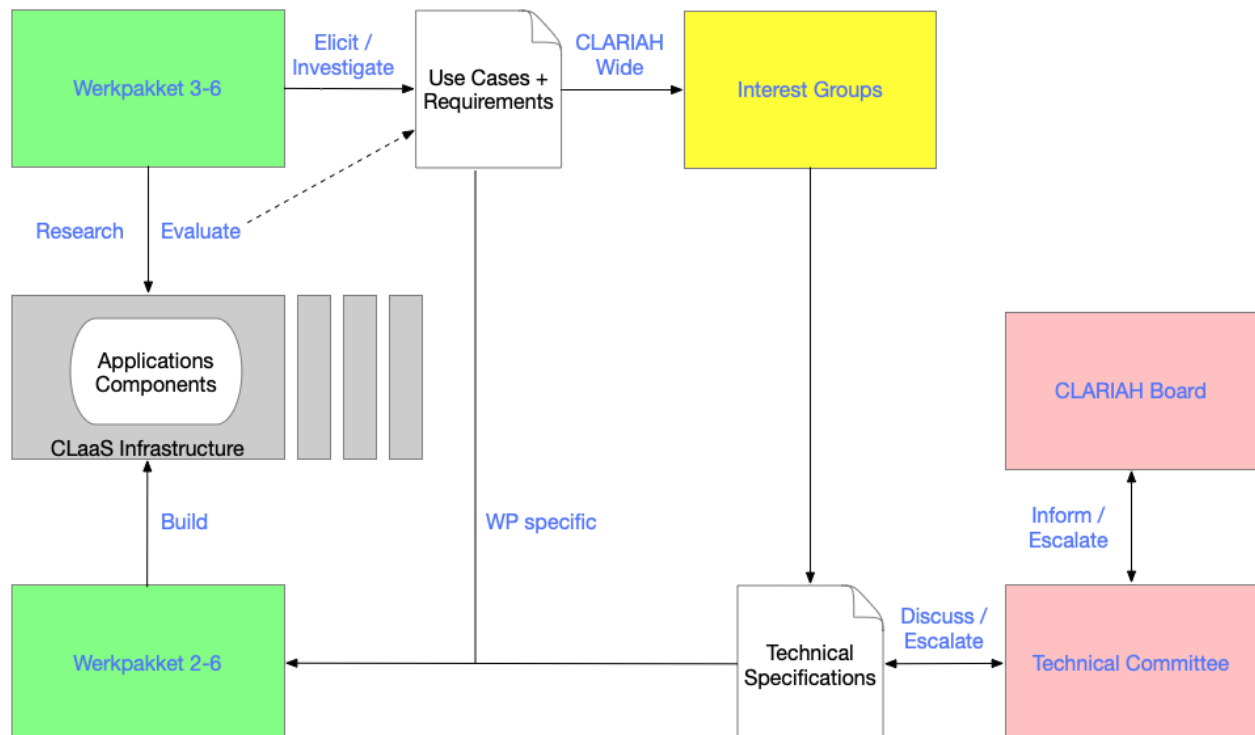
In een IG zitten naast een coördinator in ieder geval de ontwikkelaars (of lead developers) die binnen de werkpakketten aan de betreffende onderwerpen werken. Verder staat het een coördinator vrij om specialisten binnen en buiten CLARIAH te betrekken in een IG. Denk bijvoorbeeld aan specialisten vanuit computer science of vertegenwoordigers van gerelateerde projecten (bijvoorbeeld CLARIN, NDE, ODISSEI).

¹ Leidraad is hier om te kiezen voor oplossingen die direct kunnen worden ingezet voor onderzoekers, en vanuit een 'minimal viable product' (MVP) in fases verder kunnen worden ontwikkeld.

1.2 Verhouding IGs tot CLARIAH

Het overzicht hieronder schetst de relatie tussen Werkpakketten, de IGs, het Technisch Comité, het bestuur en de CLARIAH infrastructuur.

De domeinspecifieke werkpakketten 3-6 halen wensen en requirements op over de functionaliteiten binnen CLARIAH bij de onderzoekers in hun domein. Hieruit komen *use case beschrijvingen* (hoe zouden onderzoekers binnen CLARIAH willen werken) met hieraan gekoppelde *functionele requirements* (wat is er nodig om dat mogelijk te maken). Wanneer deze requirements ook voor andere werkpakketten/domeinen relevant zijn of een CLARIAH brede aanpak vereisen, zorgen de Interest Groups (waarin uiteraard de werkpakketten vanwaar de requirements komen zijn vertegenwoordigd) voor de vertaling naar een *technische aanpak* (specificaties) die rekening houdt met dit CLARIAH-brede perspectief en de *infrastructurele, niet-functionele requirements*.



Het Technisch Comité (TC) waarin de IGs zijn vertegenwoordigd dient hierbij als platform om keuzes te bediscussiëren en problemen te escaleren. Het TC informeert het CLARIAH bestuur over deze keuzes en wat dat betekent voor de oplevering van gewenste functionaliteiten in de werkpakketten. Ook kan het TC zo nodig escaleren wanneer er technische of financiële knelpunten optreden met betrekking tot de levering van die gewenste functionaliteiten en een beslissing op bestuurlijk niveau vereisen. Die beslissingen worden vervolgens weer teruggekoppeld naar het TC op basis waarvan de technische specificaties kunnen worden aangepast.

De technische specificaties dienen als input voor de ontwikkelteams binnen de werkpakketten die de applicaties en componenten ontwikkelen, met een coördinerende rol voor WP2 wanneer het gaat om CLARIAH brede / infrastructurele componenten of applicaties.

De ontwikkelde applicaties en componenten worden binnen de werkpakketten of met meerdere werkpakketten tegelijk geëvalueerd. Deze evaluatie kan verschillende vormen hebben, zoals de Research Projects uit CLARIAH Core of kleinschalige varianten hierop binnen de WPs, of formele evaluaties waarbij (externe) onderzoekers als eindgebruikers worden betrokken (bijvoorbeeld een pool van geïnteresseerde onderzoekers die kunnen worden geraadpleegd²). De output van de evaluaties wordt gebruikt om de ontwikkelde functionaliteiten te fine-tunen en bugs op te lossen, en om de requirements aan te scherpen of een stap verder te brengen.

Via de route use cases/functionele requirement naar technische specificaties, gefocuste ontwikkeling en evaluatie wordt een circulair ontwikkelproces (co-development/co-creatie) gecreëerd. De eindgebruiker blijft hiermee centraal staan en de regie blijft bij de werkpakketten. Tegelijkertijd helpt deze aanpak om de technische implementatie van de CLARIAH infrastructuur als geheel, effectief en efficiënt in te richten en te monitoren. Door de ontwikkelteams binnen de WPs waar mogelijk 'agile' in te richten kan het circulaire proces verder worden versterkt.

1.3 IGs en de CLARIAH Roadmap

Een technologische roadmap is een flexibele planningstechniek ter ondersteuning van strategische en langetermijnplanning, door korte- en langetermijndoelen te koppelen aan specifieke technologische oplossingen (wikipedia). Om dit binnen CLARIAH vorm te geven starten we met een *planning* van de technologische oplossingen³ die we vanuit de IGs willen vormgeven.

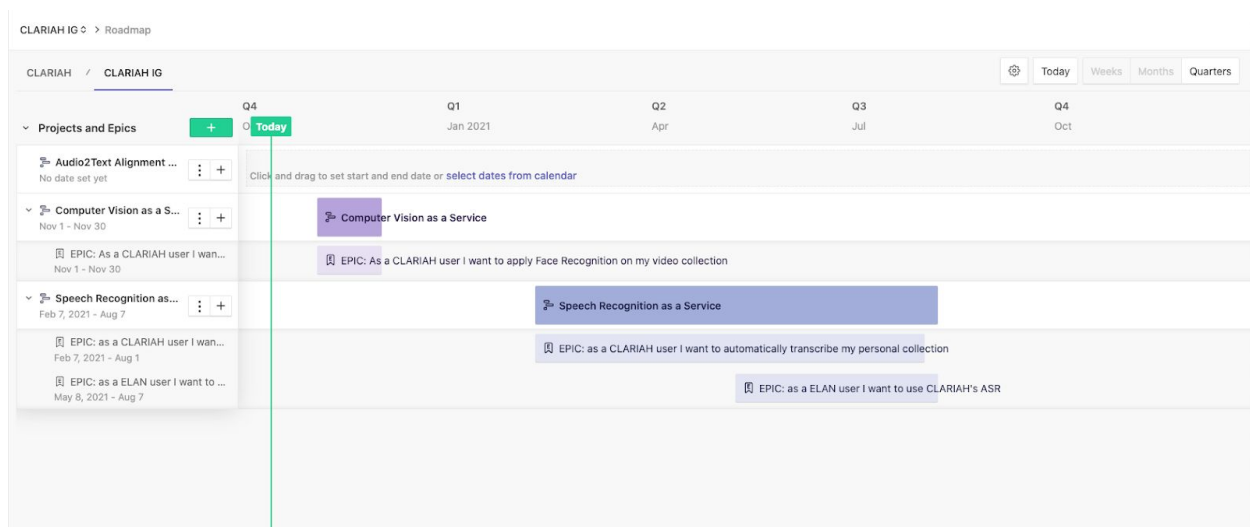
Hieronder een screenshot van een (fictief) stukje roadmap van een enkele IG (IG-AVprocessing), gemaakt met de ZenHub applicatie van GitHub die een Roadmap functie biedt. De roadmap is opgedeeld in projecten (in dit voorbeeld zijn er drie projecten in IG-AVProcessing: alignment, spraakherkenning en computer vision). Onder elk project hangen EPICs die in brok werk in het project beschrijven vanuit gebruikersperspectief. Bijvoorbeeld "Als een CLARIAH onderzoeker wil ik spraakherkenning inzetten op mijn persoonlijke collectie". Het werk binnen zo'n EPIC kan dan weer opgedeeld worden in kleinere brokken waar ontwikkelteams (bijvoorbeeld via tickets in een Agile omgeving) aan kunnen werken.

In de context van de IGs zou het al nuttig zijn om de roadmap van alle IGs onder elkaar te zetten zodat we een beter beeld hebben van wat er CLARIAH breed op de planning staat, beschreven tot op het niveau van een EPIC zoals hierboven. De inrichting van de verdere uitwerking is aan de

² who would be organizing such a sounding board? Especially of users across the WP's? (RZ)

³ If this signifies that choices for (some) specific technological solutions were already made, an aggregated list of such choices would be good as part of this doc (DB)

individuele teams in de WPs. Aan de andere kant valt er wat voor te zeggen om de onderwerpen die in de IGs worden opgepakt, ook centraal (vanuit WP2?) te coördineren als het gaat om de implementatie. Met teams van ontwikkelaars vanuit de verschillende werkpakketten, tijdelijk gekoppeld aan een taak binnen een IG, kan dan gericht worden gewerkt aan centrale applicaties of componenten. Met de inrichting van de security rond inloggen met Satosa is in het verleden al een beetje op deze manier gewerkt. In de huidige situatie met IGs zou dat betekenen dat de requirements rond inloggen mbv input uit de WPs zijn verzameld door de IG-Security. Ook de implementatie ervan (prioriteren requirements, vaststellen van meest geschikte technische oplossing) wordt daar voorbereid, in samenspraak met WP2. Vanuit WP2 wordt de implementatie zelf gecoördineerd en vormgegeven met behulp van ontwikkelaars in WP2 en (in dit geval) WP5. De beschikbaarheid van ontwikkelaars uit alle WPs is hierbij cruciaal.



2. Interest Groups

Hieronder een voorlopige lijst van beoogde Interest Groups. Het zijn er nu 12 maar de komende weken zal worden besproken in hoeverre elke IG echt nodig is. Een aantal IGs hebben een sterretje (*) om aan te geven dat over het bestaansrecht van deze IG nog een discussie loopt. Er zit vanzelfsprekend soms overlap in de onderwerpen van verschillende IGs en ook ligt het leggen van praktische relaties (overleg, afstemming) voor de hand. Nog niet alle IGs zijn nog volledig ingevuld, hier wordt de komende weken aan gewerkt. De IGs:

1. DevOps
2. Preservation
3. Security & Monitoring*
4. Audiovisual Processing
5. Text Processing
6. Annotation
7. Linked Open Data
8. Workflows
9. Curation*
10. UI/UX
11. Tools2Data*
12. Data Management*

Voor de beschrijving van de IGs houden we onderstaand format aan:

IG-Titel	Hoe de IG wordt genoemd
Coordinator	Wie is verantwoordelijk voor output IG en coordinatie ervan
Leden	Wie zijn betrokken vanuit (andere) WPs. Niet elk WP hoeft vertegenwoordigd. Leden houden eigen WP leiders op de hoogte.
Doel	Zo specifiek mogelijke beschrijving van het onderwerp waar de IG zich mee bezig houdt, liefst begrijpelijk voor niet-techneuten. Welk (type) probleem helpt dit IG mee oplossen.
Scope	Wat valt expliciet <i>buiten</i> deze IG
Gerelateerd	Welke activiteiten, tools, producten waaraan nu in CLARIAH wordt gewerkt zijn (sterk) gerelateerd aan dit IG (om ook een beetje context te bieden bij het doel)
Relatie tot andere IGs	Met welke IGs is afstemming / samenwerking noodzakelijk voor het succesvol kunnen opleveren van output
Onderwerpen (prio's)	Shortlist(!) van urgente topics die op de rol staat en waaraan de komende tijd (komende half jaar?) gaat worden gewerkt.

2.1 Devops

IG-Titel	DevOps
Coordinator	Mario Mioldijk (WP2)
Leden	Hennie Brugman (WP6) Martijn van de Donk (WP5) Maarten van Gompel (WP3)
Doel	Met DevOps ⁴ wordt bedoeld op het efficiënt organiseren van softwareontwikkeling en -beheer door een integrale aanpak die zowel rekening houdt met flexibiliteit (agile ontwikkeling) als met betrouwbaarheid (release management, monitoring) binnen een keten van IT-processen. Binnen CLARIAH is dat te vertalen naar het op een beheersbare manier ontwikkelen en beschikbaar maken (inclusief SLAs) van componenten van de infrastructuur die op diverse plaatsen binnen CLARIAH worden ontwikkeld. ⁵
Scope	(wat valt buiten deze IG: nog nader te bepalen)
Gerelateerd	<ul style="list-style-type: none"> • Het gebruik van Kubernetes om de implementatie, schaling en het beheer van software te automatiseren (WP2) • De virtualiseerbare NLP omgeving LaMachine (WP3) • CLAM (WP3) • Video analyse pijplijn DANE, spraakherkenning, Media Suite (WP5) • Toolschairs van INT en VU (WP6) • Tools2Data scenario (WP6)
Relatie tot andere IGs	
Onderwerpen (prio's)	<ol style="list-style-type: none"> 1. A 2. B 3. C

2.2 Preservation/herbruikbaarheid/repliceerbaarheid

IG-Titel	Preservation
Coordinator	Vacant
Leden	Roeland Ordeman (WP5) Herbert van der Sompel (WP3)

⁴ Zie ook

<https://www.computable.nl/artikel/opinie/development/5048283/1509029/de-zin-en-onzin-van-devops-ontrafeld.html>

⁵ I think all WP's ought to be represented in IG DevOps, or actually all IG's! (RZ)

	Jerry de Vries (WP3), Joris van Waesberge (WP2), Richard Zijdemans (WP4) Lucien van Wouw
Doel	In Preservation gaat het over het bewaren en preserveren van door gebruikers gegenereerde (meta)data: hoe die zich verhouden tot institutionele data (bv collecties van IISG, KB, NIBG), welke formaten hiervoor worden gebruikt en waar die data voor de langere termijn kunnen worden opgeslagen. Het gaat hierbij ook over strategieën om 'intermediate' data op te slaan (bv in VREs) die vervolgens weer kunnen worden gekoppeld aan lange termijn opslag.
Scope	(wat valt buiten deze IG: nog nader te bepalen)
Gerelateerd	<ul style="list-style-type: none"> • CMDI, een framework voor hergebruik van metadata (WP3) • Het voor een eindgebruiker inzichtelijk en bruikbaar maken van repositories voor preservatie, afhankelijk van de use cases in de verschillende WPs
Relatie tot andere IGs	<ul style="list-style-type: none"> • LOD: provenance trail from source to LOD to Query result
Onderwerpen (prio's)	4. A 5. B 6. C

2.3 Security & Monitoring*

IG-Titel	Security & Monitoring
Coordinator	Jan Pieter Kunst (WP2)
Leden	Jaap Blom (WP5) Daan Broeder (WP3) Hennie Brugman (WP6)
Doel	Het bepalen van CLARIAH brede security requirements (ivm copyright en privacy aspecten van data en onderzoekers), het onderzoeken van opties die aansluiten bij het (inter)nationale speelveld en de implementatie hiervan op een manier die CLARIAH breed kan worden ingezet.
Scope	(wat valt buiten deze IG: nog nader te bepalen)
Gerelateerd	<ul style="list-style-type: none"> • Satosa • Voorbeeld van implementatie diverse use cases in Media Suite • Copyrights (?); • GDPR (?); • Logging/Analytics

Relatie tot andere IGs	<ul style="list-style-type: none"> IG-LOD: ownership and rights to edit/share/view data
Onderwerpen (prio's)	7. A 8. B 9. C

2.4 Audiovisual processing

IG-Titel	AV Processing
Coordinator	Roeland Ordeman (WP5)
Leden	Henk van den Heuvel (WP3) Nanne van Noord (WP5) Arjan van Hessen
Doel	Het CLARIAH-breed beschikbaar maken van technieken om audiovisuele data te analyseren (of te acquireren), zoals met behulp van spraakherkenning, tekst/spraak-alignment, sprekerherkenning, en diverse vormen van beeldherkenning.
Scope	(wat valt buiten deze IG: nog nader te bepalen)
Gerelateerd	<ul style="list-style-type: none"> DANE (WP5) LaMachine (WP3)
Relatie tot andere IGs	<ul style="list-style-type: none"> IG-DevOps (schaalbaarheid en beheer AV-analyse services) IG-workflow (analyse pijplijns, bv LaMachine) IG-Text (bv output van spraakherkenning in text-analyse tools) IG-Security & Monitoring (beveiligen van AV-analyse services) IG-Annotation (relatie tot handmatige annotatie's)
Onderwerpen (prio's)	

2.5 Text processing

IG-Titel	Text processing
Coordinator	Jesse de Does (WP3) en Maarten van Gompel (WP3)

Leden	Hennie Brugman (WP2/6) Roeland Ordelman Martin Reynaert Liefst nog iemand van VU, FA (bijvoorbeeld Eduard Drenth) en Groningen?
Doel	Het CLARIAH-breed beschikbaar maken en aanbevelen van technieken voor automatische tekst-verrijking (al dan niet linguïstisch), text-mining & tekst-analyse, taal modellering en vertaling.
Scope	(wat valt buiten deze IG: nog nader te bepalen)
Gerelateerd	<ul style="list-style-type: none"> • Frog (WP3), DeepFrog (WP3) • Alpino (dependency tagging, pos tagging) • ucto (WP3) (tokenisatie) • PICCL/TICCL (WP3/2) (normalisatie/postcorrectie) • Langcat, FoLiA-langcat, folialangid, colibri-lang (taalidentificatie) • Gecco/valkuil (spellingcorrectie) • Oersetter (vertaling) • Colibri Core • T-scan, Stylen • Nederlab pipeline, VU Reading Machine Pipeline • spaCy, coreNLP, stanza, FreeLing, UDPipe, Huggingface Transformers
Relatie tot andere IGs	<ul style="list-style-type: none"> • IG-Annotation (relatie tot manuele annotatie, annotatie modellen/formaten)
Onderwerpen (prio's)	<ul style="list-style-type: none"> • Overeenstemming over het gebruik van formaten (TEI, FoLiA, NAF,) • Inventarisatie van gedeelde behoeften bij de werkpakketten en aanbieden van oplossing(en) daarvoor zodat dubbel werk vermeden wordt, bijvoorbeeld (prioriteit TBD in overleg met de werkpakketten) <ul style="list-style-type: none"> ◦ Named entity recognition ◦ PoS tagging en lemmatiseren ◦ Normalisatie van historische tekst, tekst van sociale media, OCR,... ◦ Syntactische analyse ◦ • Gemeenschappelijke aanpak van taken die substantiële rekenkracht vragen (bv syntactische analyse van groot corpus; bouwen van modellen op grond van een groot corpus)

2.6 Annotation

IG-Titel	Annotation
Coordinator	Marijn Koolen / Hennie Brugman
Leden	<ul style="list-style-type: none"> • Hennie Brugman (WP6)

	<ul style="list-style-type: none"> • Jaap Blom (WP5) • Maarten van Gompel (WP3) • Roeland Ordelman (WP5) • Dirk Roorda • Peter Boot (WP6) • Lodewijk Petram (WP6) • Han Sloetjes (WP5)
Doel	<ul style="list-style-type: none"> • Foster discussion and knowledge sharing regarding annotation infrastructure • Develop and share best practices • Inform development of CLARIAH annotation tools and services
Scope	<ul style="list-style-type: none"> • Inventory of need for manual annotation support in CLARIAH • Support of the creation and use of annotations on any media type • Support of manual annotation processes • Support of manual correction of automatic annotation • Inventory of and recommendations for annotation models, formats and converters
Gerelateerd	<ul style="list-style-type: none"> • Exchange format for CLARIAH MediaSuite and ELAN (WP5) • Generic Web Annotation server (Elucidate server, SWA-S) (WP2) • FoLiA format for linguistic annotation and the annotation tool FLAT (WP3)
Relatie tot andere IGs	<ul style="list-style-type: none"> • IG-LOD: communication on annotated entities • IG-Curation: communication on vocabs used to describe entities and such • IG -Text: relatie tot automatische annotatie van teksten • IG -AVProcessing: relatie tot automatische annotatie van audio/video
Onderwerpen (prio's)	

2.7 Linked Open Data

IG-Titel	LOD
Coordinator	Richard Zijdemann
Leden	<ul style="list-style-type: none"> • Richard Zijdemann (International Institute of Social History, group coordinator) • Wouter Beek (VU University, Triply) • Victor de Boer (VU University) • Antske Fokkens (VU University)

	<ul style="list-style-type: none"> • Willem van Hage (eScience Center) • Chiara Latronico (University of Amsterdam) • Carlos Martinez-Ortis (eScience Center) • Enno Meijers (National Library of the Netherlands, Dutch Digital Heritage Network) • Willem Melder (invited, Sound & Vision) • Harmen Nijboer (Huygens ING) • Jacco van Ossenbrugge (CWI, VU University) (invited) • Ronald Siebes (DANS)(invited, tbc) • Joe Raad (VU University) • Sjors de Valk (Dutch Digital Heritage Network) • Thomas Vermaut (KNAW Humanities Cluster) • Jerry de Vries (DANS) • Mari Wigham (Sound & Vision) • Leon van Wissen (University of Amsterdam) • Menzo Windhouwer (KNAW Humanities Cluster) • Maarten van Gompel (DI, KNAW)
Doel	The key goal of each IG is to research, negotiate, propose, and implement technical choices that become a standard requirement in CLARIAH. This repository is intended to organize the work, output and documentation of the IG on Linked Open Data (LOD).
Scope	(wat valt buiten deze IG: nog nader te bepalen)
Gerelateerd	<ul style="list-style-type: none"> • find compatibility between various solutions (e.g. Anansi/Druid, CoW/LD Wizard) within the ClaaS framework; • improved out-of-the-box availability of existing tools (e.g. DIVE, brwsr) within the CLAAS framework; • improve compatibility with inter- and national initiatives (e.g. within Pelagios, Odissei and NDE); • to attract and interact with a larger less technical audience (e.g. via LD Wizard, Data Stories, Notebooks) • what is the CLAAS way to store & retrieve triples (file storage, triple stores, media fragments) • how/when/why can/should partners share LD services & data in the CLAAS the infrastructure (using Docker/Kubernetes) • performance (of querying, updating data, etc) • the use of grlc (and further development on it) • link with NDE principles • link with NDE termennetwerk • distributed search functionality for vocab and ontology resources • recommendation on metadata describing (documenting) a dataset within Clariah; ex: title (name), description, ownership, license, provenance, temporal and/or spatial coverage, keywords, etc. Which properties should be made mandatory? Which vocabulary (schema, dc, void, etc.)?
Relatie tot andere IGs	<ul style="list-style-type: none"> • IG-Annotation (and text and AV?) • See specifically IG-Curation and the awesome humanities ontologies list.

	<ul style="list-style-type: none"> IG-Datamanagement&Search: The need for a good vocabulary to describe metadata as related to linksets datasets (VOID?)
Onderwerpen (prio's)	<ul style="list-style-type: none"> Prio: LOD storage, browsing and retrieval; Prio: LOD workflow (user perspective) Prio: Entity linking <p>LOD Browsers</p> <ul style="list-style-type: none"> Anansi Brwsr Druid LOD Storage Anansi Druid DIVE (deprecated). Media Suite aims for a hybrid approach where LOD can provide an auxiliary mechanism for connecting information sources. Linked Data Converters CoW (CSV to RDF) CMDI2RDF (CMDI to RDF) LDWizard (CSV to RDF) CodemetaPy (divers naar JSON-LD) - voor software metadata MMiscellaneous FoLiA Set Definitions (...) Codemeta in LaMachine (...) Network-of-terms Comunica OOutreach datastories (tool for creating webpages based on sparql queries)

2.8 Workflows

IG-Titel	Workflows
Coordinator	Daan Broeder en Maarten van Gompel (WP3)
Leden	<ul style="list-style-type: none"> Jauco Noordzij Jaap Blom Hennie Brugman Martijn van de Donk Roeland Ordelman Joe Raad

	<ul style="list-style-type: none"> Adam Belloum
Doel	<ul style="list-style-type: none"> Delen en aanbevelen van CLARIAH-brede oplossingen voor workflow systemen Inventariseren welke workflow oplossing er gebruikt worden in CLARIAH Inventariseren welke workflows gewenst zijn vanuit onderzoekers Interoperabiliteit tussen tools binnen CLARIAH bevorderen
Scope	(wat valt buiten deze IG: nog nader te bepalen)
Gerelateerd	<ul style="list-style-type: none"> Workflow systemen als Nextflow, Luigi, Airflow, Apache Tavera Common Workflow Language (CWL) Bestaande workflow oplossingen in CLARIAH zoals: <ul style="list-style-type: none"> PICCL/TICCL (Nextflow) De tweede VRE poging VU Reading Machine?
Relatie tot andere IGs	<ul style="list-style-type: none"> IG text IG devops (voor deployment)
Onderwerpen (prio's)	

2.9 Curation

IG-Titel	Curation
Coordinator	Sebastian Derks
Leden	<ul style="list-style-type: none"> Willem Melder (WP5) Ruben Schalk (WP4) Mari Wigham (WP5)
Doel	
Scope	(wat valt buiten deze IG: nog nader te bepalen)
Gerelateerd	<ul style="list-style-type: none">
Relatie tot andere IGs	
Onderwerpen (prio's)	

2.10 UI/UX

IG-Titel	UIUX
Coordinator	Bas Doppen (WP6)
Leden	<ul style="list-style-type: none"> • Sebastiaan Fluitsma (WP1) • Bram van den Hout (WP4) • Roeland Ordeman (WP5) • Paul Trilsbeek (WP3)
Doel	Het uitwerken van strategieën rond UI design die enerzijds zorgen voor een uniform 'CLARIAH smoel' waar dat kan, maar anderzijds rekening houdt met het gedistribueerde karakter (en gedeelde eigenaarschap) van de portals binnen de CLARIAH infrastructuur. Ook onderzoekt deze IG modellen voor UI/UX die onderzoekers helpen hun weg te vinden in de complexe wereld van data en tools, zowel op CLARIAH-breed niveau als binnen de individuele domeinen/portals.
Scope	(wat valt buiten deze IG: nog nader te bepalen)
Gerelateerd	<ul style="list-style-type: none"> • stories.datalegend.net (data stories)
Relatie tot andere IGs	
Onderwerpen (prio's)	

2.11 Tools-to-Data*

IG-Titel	Tools2Data
Coordinator	
Leden	<ul style="list-style-type: none"> • Maarten van Gompel (DI, KNAW) [LaMachine]
Doel	Methoden en strategieën uitwerken om CLARIAH tools en Jupyter Notebooks toe te kunnen passen op vanwege copyright of privacy afgeschermd data collecties (bv KB, NIBG) in een 'secure environment' (data vault).
Scope	(wat valt buiten deze IG: nog nader te bepalen)

Gerelateerd	<ul style="list-style-type: none"> • Jupyter Notebooks • Tool repositories • LaMachine, DANE, CLAM
Relatie tot andere IGs	<ul style="list-style-type: none"> • IG-Security • IG-DataSearch • IG-TextProcessing • IG-AVProcessing
Onderwerpen (prio's)	

2.12 Data Management & Search*

IG-Titel	DataManagementSearch
Coordinator	
Leden	Sebastiaan Fluitsma
Doel	Ontwikkelen van een algehele CLARIAH strategie voor data management gegeven de diverse data repositories van verschillend signatuur binnen CLARIAH (bv CLARIAH Data Registry) inclusief de mogelijkheid om over collecties heen (bv via data repository) en binnen collecties te kunnen zoeken: individueel (zie bv Delpher), geaggregeerd (zie bv Media Suite), dan wel gefedereerd. Het verwijzen binnen/tussen collecties (via LOD of anderszins) valt hier ook onder.
Scope	(wat valt buiten deze IG: nog nader te bepalen)
Gerelateerd	<ul style="list-style-type: none"> • nederlab.nl • mediasuite.clariah.nl • mediasuitedata.clariah.nl (CKAN gebaseerde registry) • druid.datalegend.nl • datasets.iisg.amsterdam • whgazetteer.org • github.com/clariah/jobHoard •
Relatie tot andere IGs	<ul style="list-style-type: none"> • IG-Preservation • IG-LOD • IG-Tools2Data • IG-UI/UX
Onderwerpen (prio's)	<ul style="list-style-type: none"> • Verzamelplekken voor metadata lijsten van entities zoals plaatsen, personen, beroepen, producten, organisaties, etc.

2.11 Overzichten

WP needs for IG topic

WP\IG	1	2	3	4	5	6	7	8	9	10	11	12
1												
2												
3	x	x	x	x	x	x	x	x		x	x	x
4					x		x	x		x		x
5	x	x	x	x	x	x	x	x		x	x	x
6	x		x		x	x	x	x		x	x	x

coordinators per WP

WP\IG	1	2	3	4	5	6	7	8	9	10	11	12
1									x			
2	x		x									
3					x			x				
4							x					
5				x								
6		x				x				x		

3 Communicatie

3.1 Informatie per IG

Elke IG is zelf verantwoordelijk voor het verzamelen, up-to-date houden en publiceren van informatie rond de onderwerpen die spelen in de IG. Een aantal IGs hebben al de voorkeur uitgesproken om dit te doen in GitHub, zie de voorbeelden hieronder. Ervan uitgaande dat de IT mensen gevoeglijk bekend zullen zijn met GitHub en het de voorkeur heeft om één methode aan te houden voor alle IGs, willen we dit voor alle IGs doorvoeren. We zien daarbij het gebruik van GitHub *niet* als middel om informatie te delen binnen heel CLARIAH, maar als documentatie-strategie binnen/tussen de IGs en het technisch comité, en om persistent naar te kunnen verwijzen als achtergrondinformatie bij stukken die *we*/ CLARIAH breed worden gedeeld (zie ook hieronder).

- <https://github.com/CLARIAH/IG-Annotation>
- <https://github.com/CLARIAH/IG-Workflows>
- <https://github.com/CLARIAH/IG-AVProcessing>
- <https://github.com/CLARIAH/IG-Text>
- <https://github.com/CLARIAH/IG-LOD>

Een overkoepelende repository die naar alle bovenstaande linkt staat hier:

<https://github.com/CLARIAH/IG>

Een voorstel voor een werkwijze met de git repositories en motivatie daarvoor (contribution guidelines) is hier gedaan: <https://github.com/CLARIAH/IG-Annotation/pull/1> . Een beknopte samenvatting:

- Alle relevante documenten in de desbetreffende openbare git repository zodat alles op één plek te vinden is, het transparant is wie wat wijzigt, en ieder met tools naar eigen keuze kan editen.
- Gebruik zoveel mogelijk plain text, of specifieker: markdown syntax. Dit houdt de dingen makkelijk leesbaar/wijzigbaar/converteerbaar.
- Groepsdiscussie primair via de issue tracker (ticketing) en openbaar. Elk issue heeft idealiter een eigen duidelijke scope. Men is niet per se gebonden aan de interface en kan gewoon via mail reageren.
- Alle deelnemers aan de groep moeten zorgen dat ze de repository 'watchen' om notificaties (bv. per mail) te krijgen.
- Wijzigingen via pull requests kunnen een goede manier zijn om discussie te openen en goede peer review te hebben binnen de groep.

Elke IG verzorgt regelmatig een CLARIAH-brede post over de vorderingen van het IG, waarin in ook voor niet-technen in begrijpelijke taal wordt uitgelegd waar het IG mee bezig is. Het proces wordt afgestemd met WP1.

Het zou mooi zijn als IGs interessante bevindingen publiceren via workshop/conference papers.

Elk IG geeft roulerend een presentatie over wat de IG doet tijdens de maandelijkse Tech Dag.

3.2 Dagelijkse communicatie

Dagelijkse communicatie (chatfunctie, dus niet het primaire kanaal voor doorgeven van officiële berichten zoals vergadering e.d.) binnen en tussen IGs vindt plaats met behulp van [Slack](#). Er wordt een slack instantie aangemaakt voor CLARIAH Tech, waar elk IG en het Technische Comité een eigen kanaal heeft (bv. #IG-Annotatie). De CLARIAH Tech slack is voor iedereen binnen CLARIAH (op aanvraag, evt ook externen) toegankelijk. WP1 doet het beheer van deze slack. Elke coordinator is admin van zijn eigen kanaal.

De coördinatoren zijn verantwoordelijk voor de informatievoorziening naar en **afstemming met de WP leiders**. Indien relevant kan dat ook een WP leider zijn van een ander WP dan waar de coördinator in zit!

Het Technisch Comité i.c. de CTO van CLARIAH is verantwoordelijk voor de **informatievoorziening rond de IGs richting het CLARIAH bestuur**.

3.3 Vormen van overleg

Er vindt 1x per maand op een vast moment (bv laatste donderdagmiddag van de maand) een **Tech Dag** plaats om kennis uit te wisselen over technische onderwerpen. De doelgroep is in principe alle IG deelnemers en ontwikkelaars maar staat in principe open voor iedereen binnen CLARIAH met de aantekening dat van een “technisch publiek” wordt uitgegaan. Vaste punten bij de Tech Dag is een presentatie van één of meerdere IGs (in overleg, roulerend) en een update van de CTO over relevante meta-onderwerpen die spelen, zoals samenwerkingen met andere infrastructuren en netwerken. Communicatie over de Tech Dag wordt verzorgd door WP1. Het bestuur ontvangt een kort verslag van elke Tech Dag.

Er vindt 1x per maand op een vast moment (bv tweede donderdagmiddag van de maand) een **Technisch Comité** overleg plaats dat bestaat uit de IG coördinatoren en de CTO, aangevuld met de technisch coördinatoren van WPs voor zover die geen IG coördineren.

De coördinatoren organiseren zelf hun overleg binnen de IG.