

Subsidie / Subsidy : **Nationale roadmap grootschalige onderzoeksinfrastructuur 2011 BOO**

Projectnummer / Project number : **184.032.306**

Hoofdaanvrager / Main applicant : **Prof. dr. J.E.J.M. Odijk**

Projecttitel / Project title : **CLARIAH - Common Lab Research Infrastructure for the Arts and Humanities**

**Adviseur: 1****1. THE LIKELIHOOD OF SCIENTIFIC BREAKTHROUGHS (SCIENCE CASE)****1.a. Please comment on the science case made in the application.**

I believe CLARIAH has every potential to contribute to the production of major breakthroughs in science, provided considerable time and effort is spent on the dissemination of the infrastructure and on providing opportunities for the numerous researchers in Arts and Humanities who are less familiar with digital data to become acquainted with the undeniable advantages of using the resources that will be made available by the Common Lab. Researchers often rely too heavily on opportunistic data that is readily available. By creating a facility that provides unprecedented access to digital texts, databases and audiovisual material, the CLARIAH team will enable researchers working in the fields of Arts and Humanities not only to access incomparably more data, but also to analyse existing data in ways that have not been contemplated before. Because the resources made available by CLARIAH will be shared and can be accessed by anyone, it will be easier for researchers from different institutions to engage in collaborative research, where synergies can be exploited to contribute to new scientific breakthroughs. Another advantage of the Common Lab is that it will enable scholars to criticise or build on previous research using the same data, which is something that we do not see often enough due to the fact that academics often tend to keep their original documents and files to themselves. The possibility of reusing data is a key factor of rigorous empirical research and can have important implications in promoting the progress of science.

**1.b. What is your assessment of the science case?  
Please choose: poor / fair / good / very good / excellent\***

excellent

**2. THE POTENTIAL FOR 'BRAIN GAIN' (TALENT CASE)****2.a. Please comment on the talent case presented in the application.**

As the Common Lab is a virtual facility, it can in theory be accessed from anywhere in the world. It will nevertheless keep Dutch researchers in the Netherlands and attract researchers from abroad if its use is a regular component of the curricula of Dutch universities and Dutch universities establish a reputation for excellence in the training and use of advanced ICT methods in Arts and Humanities. Needless to say, because Dutch is a minority language, an English interface to the Common Lab is highly desirable if one of its goals is to attract international research.

**2.b. What is your assessment of the talent case?  
Please choose: poor / fair / good / very good / excellent\***

excellent

**3. SOCIAL AND COMMERCIAL RELEVANCE (INNOVATION CASE)****3.a. Please give your opinion on the innovation case presented in the application.**

Although it is usually easier to establish links between the hard sciences and industry, the social and commercial relevance of CLARIAH for both the private and public sector is undeniable. The digitized texts, images, databases, audio, video and multimedia material stored and made available through the Common Lab can be of utmost interest to a wide range of enterprises, including museums, libraries, public archives, newspapers, televisions, statistics institutes and companies specializing in natural language processing. Young researchers trained to exploit the facility are likely to develop skills that are becoming increasingly important in today's world, with more and more companies from all sectors of society depending on ICT and digital resources. The human resources factor is therefore of great social and commercial relevance. The impressive number of letters of support from Dutch companies endorsing CLARIAH is further proof of its value for business and industry.

**3.b. What is your assessment of the innovation case?  
Please choose: poor / fair / good / very good / excellent\***

excellent

#### 4. COLLABORATION AND COMPETITION (PARTNERSHIP CASE)

##### 4.a. Please give your opinion on the partnership case.

With partnerships across the Netherlands and as the national counterpart of European Union networks, there is no doubt that the Common Lab is not operating in isolation and is extremely likely to benefit from interaction and exchanges with its associated research centres. The CLARIAH team also appears to have a clear vision of how to coordinate its various bodies.

##### 4.b. What is your assessment of the partnership case?

Please choose: poor / fair / good / very good / excellent\*

excellent

#### 5. FINANCIAL ASPECTS (BUSINESS CASE)

##### 5.a.

Please comment on the business case made in the proposal (please notice that even in case no funding is requested for, applicants must fill in subsections 5.1 and 5.2 of the business case)

-

##### 5.b. What is your assessment of the business case?

Please choose: poor / fair / good / very good / excellent\*

-

#### 6. TECHNICAL FEASIBILITY/TECHNICAL CHALLENGES (TECHNICAL CASE)

##### 6.a. Please comment on the technical case made in the proposal.

Since CLARIAH is not a totally new infrastructure, but builds on the groundwork developed in the scope of other projects, the risk of unforeseen technical impediments is greatly reduced. Of course, the aim of interoperability is not an easy one to achieve, but the establishment of clear guidelines and the need for refinement of those guidelines as the need arises is clearly stated. This also applies to other technical risks identified by the applicants, such as failing to grasp the actual needs of users, where the need for close cooperation between developers and users is acknowledged as a key to preventing potential mismatches.

##### 6.b. What is your assessment of the technical case?

Please choose: poor / fair / good / very good / excellent\*

excellent

#### 7. POSSIBLE FOCUS FOR THE NETHERLANDS

##### 7.a. Please comment on the possible focus for the Netherlands.

Concern with the development of digital resources for the humanities and the long-term preservation of digital data and its accompanying software is growing around the world. The fact that CLARIAH builds on other projects in this area where the Netherlands already plays a leading role will only strengthen the position of the country in this respect. It is essential to continue to invest in this kind of infrastructure if Dutch researchers and the Netherlands are to maintain and build up their well-earned reputation in this field.

##### 7.b. What is your assessment of this criterion?

Please choose: poor / fair / good / very good / excellent\*

excellent

#### 8. CRITICAL MASS

##### 8.a. Please give your opinion on the criterion "Critical mass".

The fact that the Common Lab involves higher level education and research institutions from all over the Netherlands means that its outcomes are very likely to serve the needs of scholars across the country. Because of its breadth of scope, in that it includes resources that are relevant to a wide range of interests in the Arts and the Humanities, it is foreseeable that the infrastructure will have many users, and that, among them, will be some of the top researchers in the Netherlands. The fact that the data and tools envisaged are as much as possible open access means people outside the participating institutions and external researchers should not find it a problem to use the infrastructure, which not only promotes scientific progress but also maximizes the investment made.

##### 8.b. What is your assessment of this criterion?

Please choose: poor / fair / good / very good / excellent\*

excellent

## 9. EMBEDDING

### 9.a. Please give your opinion on the criterion "Embedding".

There is absolutely no doubt that the organizational structure CLARIAH is firmly embedded within the Netherlands. This is obvious from the list of institutions involved in the present application for funding. The applicants have also demonstrated that there are clear links between the Dutch institutions involved in CLARIAH and several international networks.

### 9.b. What is your assessment of this criterion?

Please choose: poor / fair / good / very good / excellent\*

excellent

## 10. PROVEN WILLINGNESS TO COLLABORATE

### 10.a. Please give your opinion on the criterion "Proven willingness to collaborate".

By embarking on a project of this dimension, the Dutch institutions applying for funding will necessarily have to collaborate with each other. The very way the present proposal was written suggests collaboration is already underway. Past projects where the same institutions were involved provide further evidence of willingness to work together.

### 10.b. What is your assessment of this criterion?

Please choose: poor / fair / good / very good / excellent\*

excellent

## 11. REFLECTION OF SOCIAL TRENDS

### 11.a. Please give your opinion on the criterion "Reflection of social trends".

By committing itself to the dissemination and long-term preservation of textual, structured and audiovisual data pertaining to the Arts and Humanities, the Common Lab is a major contribution to society. In an age where only a small percentage of facts about our present society are preserved in hard data format, the historical value of enabling digital data to be accessed twenty or fifty years from now is priceless. The findings that can be gleaned out of such a facility, especially the data pertaining to population and economic trends, can have important effects on government policy and the interpretation of history in general.

### 11.b. What is your assessment of this criterion?

Please choose: poor / fair / good / very good / excellent\*

excellent

## SUMMARY AND OVERVIEW BY REVIEWER

.a.

What is your opinion on the entire application? Please justify your overall assessment by summarizing or briefly commenting on the strengths and weaknesses of the proposal.

The proposal presented by the Common Lab for Research in the Arts and Humanities team is sufficiently detailed to allow a careful appreciation of its merits. CLARIAH addresses a real need that, if successful, will undoubtedly have positive implications for science and society in general and the Netherlands in particular. The applicants are experienced researchers, many of whom have already worked together in the past, and seem to be well aware of existing work in the field and the challenges that lie ahead of them. I have no doubt that the present proposal focusing on the Arts and Humanities will have a favourable impact on the National Roadmap for Large-Scale Research Facilities.

### .b. What is your overall assessment of the entire proposal?

Please choose: poor / fair / good / very good / excellent\*

excellent

**Adviseur: 2****1. THE LIKELIHOOD OF SCIENTIFIC BREAKTHROUGHS (SCIENCE CASE)****1.a. Please comment on the science case made in the application.**

Allowing humanities researchers to properly exploit digital resources has the potential for radical change, transforming all aspects of research in those fields. If CLARIAH is at all successful, it will be part of a real paradigm shift. These points are well made in the proposal and the science case in this application is thus extremely strong.

There are some potential scientific benefits which are not extensively discussed in the proposal. eScience has helped biomedical researchers by allowing workflows to be published - among other thing, this makes it possible for experiments to be replicated very easily. A similar benefit should be possible from this resource, which would be of considerable importance, since it is often difficult to replicate findings based on ad-hoc investigations of corpora or databases. It would require explicit support in CLARIAH, so I hope this could be considered. Another aspect is the potential for allowing research to be carried out at an earlier point in students careers than is currently possible. If the facilities were sufficiently easy to use, it should be possible for beginning undergraduates to have some hands-on experience of investigating hypotheses in humanities subjects.

**1.b. What is your assessment of the science case?**

Please choose: poor / fair / good / very good / excellent\*

excellent

**2. THE POTENTIAL FOR 'BRAIN GAIN' (TALENT CASE)****2.a. Please comment on the talent case presented in the application.**

This is another especially strong aspect of the proposal. There is huge untapped potential for exploitation of digital resources in the humanities and a resource like CLARIAH would be extremely attractive to researchers, especially in areas where the Netherlands is strong and in related areas. I would expect there to also be great interest in the process of development, since it seems likely that researchers in other countries would want to set up their own version of such a resource and being involved in the construction of CLARIAH would be extremely useful from this perspective.

A minor point: the discussion of training users to exploit CLARIAH resources centers on students. However, I would expect it to be necessary to offer training to established researchers if CLARIAH is going to contribute to making the Netherlands an attractive location for researchers.

**2.b. What is your assessment of the talent case?**

Please choose: poor / fair / good / very good / excellent\*

excellent

**3. SOCIAL AND COMMERCIAL RELEVANCE (INNOVATION CASE)****3.a. Please give your opinion on the innovation case presented in the application.**

There is evidence in the proposal of interest from several companies in different aspects of the facility. Probably the most direct and immediate benefit would be to publishers, since the use of metadata to facilitate enhanced publications has already been very beneficial in many hard science subjects, and should also add value to humanities publications.

The historical database component of the project has a direct relevance to current social issues. The text related aspects could also be relevant, e.g., because of the possibility of automatic analysis mentioned in the context of opinion mining. This is a less strong link, but if it were possible to develop sophisticated text analysis tools which were easily usable, it would have enormous potential for impact on the ability of users to investigate political issues, which clearly has a social relevance as well as an academic one.

**3.b. What is your assessment of the innovation case?**

Please choose: poor / fair / good / very good / excellent\*

excellent

**4. COLLABORATION AND COMPETITION (PARTNERSHIP CASE)****4.a. Please give your opinion on the partnership case.**

There is evidence of excellent collaboration between the partners themselves and also between the partners and institutions outside the Netherlands. The proposed management structures are sensible.

**4.b. What is your assessment of the partnership case?**

Please choose: poor / fair / good / very good / excellent\*

excellent

**5. FINANCIAL ASPECTS (BUSINESS CASE)****5.a.**

Please comment on the business case made in the proposal (please notice that even in case no funding is requested for, applicants must fill in subsections 5.1 and 5.2 of the business case)

The budget appears reasonable, though there is little detail. e.g., are costs for training users and for the "brain gain" activities included under "Dissemination and outreach"? 5.2 makes clear the extent of other activity which the project will draw on, which makes it more feasible that CLARIAH can achieve its ambitious goals within the proposed budget.

**5.b. What is your assessment of the business case?**

Please choose: poor / fair / good / very good / excellent\*

very good

**6. TECHNICAL FEASIBILITY/TECHNICAL CHALLENGES (TECHNICAL CASE)****6.a. Please comment on the technical case made in the proposal.**

The proposal envisages detailed workplans being constructed as early deliverables. It is sensible for there to be considerable flexibility in the project, since the technology is changing so rapidly. However, the proposal was less specific about technical aims and methodology than I would have liked and this makes assessment somewhat difficult.

I am concerned about the possible expectation of development in search technology and NLP beyond the current state of the art (e.g., "tools for analysis of discrepancies between stories" p10) - it is possible that the proposers assume something relatively limited which is currently feasible, but it's unclear exactly what. (Obviously there are space limitations on the proposal, but citations would help understanding at minimal cost in space.) The existence of CLARIAH should help stimulate research into these areas, especially if it provided a way for PhD students and researchers to quickly embed their research into a functioning system, but it would be important to avoid any "mission creep" on CLARIAH itself.

I believe that the risk analysis on p25 underplays the likelihood and impact of the 'one size fits none' problem, given the experiences of projects which have attempted to deploy related technologies for use in the sciences. Providing solutions that work for all humanities research could well prove over-ambitious: collecting information about user requirements would in itself be a huge task. In particular, the problems of semantic interoperability have proved acute even with research on subjects such as astronomy. There is a danger that work on this becomes a huge timesink - as indeed has already happened with many of the attempts at standardization for linguistics.

One issue that is not mentioned is that of finding suitable personnel to develop and maintain the services. It appears that considerable software engineering skills would be required, plus good knowledge of research in the various academic subjects, but that the work itself would not primarily be research-oriented. Presumably, the personnel would be university employees (though I note some outsourcing is envisaged).

Despite these issues, I am confident that the proposers could deliver a resource that would be of great use to humanities researchers.

**6.b. What is your assessment of the technical case?**

Please choose: poor / fair / good / very good / excellent\*

very good

**7. POSSIBLE FOCUS FOR THE NETHERLANDS****7.a. Please comment on the possible focus for the Netherlands.**

The coverage of this proposal is very broad: the Netherlands is an international leader in some of the areas covered (I do not have the expertise to judge standing in all fields mentioned in the proposal). The case for

leadership in the study of Dutch history and the Dutch language is undeniable, of course, and the aspects of the project which involve data conversion, enrichment and annotation will naturally focus on such resources. It is also clear that the partners are playing a leadership role in the EU projects with which they are currently involved.

**7.b. What is your assessment of this criterion?**

Please choose: poor / fair / good / very good / excellent\*

excellent

**8. CRITICAL MASS**

**8.a. Please give your opinion on the criterion "Critical mass".**

The Netherlands has an important concentration of research in some areas of the humanities which are very directly related to this proposal. In particular, the environment for linguistics (using the term broadly) is exceptional, both because of the number of researchers with an international reputation and because of the dialogue between computational linguistics and other areas of linguistics which is now almost non-existent in many countries (including the UK and the US). Hence the Netherlands is in a unique position to develop a resource such as CLARIAH.

**8.b. What is your assessment of this criterion?**

Please choose: poor / fair / good / very good / excellent\*

excellent

**9. EMBEDDING**

**9.a. Please give your opinion on the criterion "Embedding".**

The previous projects on which CLARIAH builds and the engagement of the partners in EU research make this very clear.

**9.b. What is your assessment of this criterion?**

Please choose: poor / fair / good / very good / excellent\*

excellent

**10. PROVEN WILLINGNESS TO COLLABORATE**

**10.a. Please give your opinion on the criterion "Proven willingness to collaborate".**

The management structures look practical. The partners have extensive experience of collaboration and there are good ties to existing projects.

**10.b. What is your assessment of this criterion?**

Please choose: poor / fair / good / very good / excellent\*

excellent

**11. REFLECTION OF SOCIAL TRENDS**

**11.a. Please give your opinion on the criterion "Reflection of social trends".**

I have limited knowledge of the Netherlands policy frameworks and priorities, but the project addresses many issues which are recognised as important within the European context, with respect to all aspects of information management. However, the comments in section 11 of the proposal about "Safety" are a considerable stretch and not convincing.

**11.b. What is your assessment of this criterion?**

Please choose: poor / fair / good / very good / excellent\*

very good

**SUMMARY AND OVERVIEW BY REVIEWER**

.a.

**What is your opinion on the entire application? Please justify your overall assessment by summarizing or briefly commenting on the strengths and weaknesses of the proposal.**

I rate this proposal as excellent overall. The possible weaknesses I have identified above are consequences of its generality and scope. It is clearly on the ambitious/risky end of the research spectrum, but I think the possible benefits are so large that this is fully acceptable. However, I would suggest a greater focus on a small, well-defined group of humanities topics, where the proposers have most expertise, would be beneficial, at least as a starting point. I think it very important that practical utility to researchers is emphasized from the early stages and that this takes precedence over more abstract/long-term issues.

**.b. What is your overall assessment of the entire proposal?**  
**Please choose: poor / fair / good / very good / excellent\***

excellent

**Adviseur: 3****1. THE LIKELIHOOD OF SCIENTIFIC BREAKTHROUGHS (SCIENCE CASE)****1.a. Please comment on the science case made in the application.**

The CLARIAH project is decidedly ambitious and does not conceal its vast dimension - 'it includes all humanities researchers in the Netherlands'. The massive facility for ehumanities research that is envisaged is very welcome, although the proponents are not pathfinders in the field, as the many references to other projects and organizations makes clear. For instance CLARIAH is described as 'building on' the well-reputed CLARIN and DARIAH initiatives.

The proposers are right in suggesting that the facility will be a very useful tool for new researchers but neglect to mention the very possible passive resistance on the part of older researchers in the humanities. The planned user-friendliness and even gamification will probably have little effect on such subjects. However, young researchers familiar with annotation tools, pattern recognition software, etc. will be well served by this facility. The possibility to use, re-use and, particularly, pass on to others treated material (e.g. insertion of automatic metadata) is a very welcome innovation, especially the provision that authors must make any underlying data used from the facility available to the wider community.

Fig. 1 provides a useful overview of the proposed infrastructure and highlights the tripartite approach reflecting the most common data types used by scholars (text, structured data, audiovisual) called, for some reason, 'kits'. The vast amount of analogue text is identified as the most used resource in humanities research and the one most needing innovative digital processing, especially when dealing, for example, with ancient volumes. Rather grandiose schemes are envisaged herein - 'trace the development of ideas through time', 'clarify the properties of successful literature'. If these are achieved then the term 'scientific breakthrough' can be applied.

Structural data in the form of databases, spreadsheets, etc. are described as essential for modern research purposes and linguistics is used as an example. As a linguist I can only applaud the projected developments outlined in the proposal. A number of other examples from other fields (migration, health, etc.) show the dedication to data intensive science, and again ground-breaking developments are predicted - 'enhancing our understanding of the origins, causes and character of the process of global inequality', and 'a synthesis of Dutch society in the last two hundred years'. On this point it is not entirely clear how Dutch-oriented the facility will be in terms of the languages made available in the vast gamut of material envisaged.

As regards audiovisual material, much work has been done over the last twenty years and the project does not seem to add anything spectacularly new to what is presently around, though the size of the samples envisaged and their interoperability with the other kits are positive features.

The size of the facility and the variety of material it will contain would seem to be one of the strong points of the proposal, bolstered by the promise that all these resources will 'interoperate seamlessly', and by the fact that all entries must comply with CLARIAH standards and procedures, though these are not spelt out. The expertise required to carry out the proposal is described as 'abundantly present' and the array of partners who are involved is indeed impressive. The only doubts to emerge from the proposal lie in the quite frequent use of modal verbs (should, may, will, can) and expressions such as 'as much as possible', 'whenever possible', 'preferably' and 'looks promising' which indicate, on occasion, a lack of certainty. Similarly it is admitted that a number of the innovations projected are not currently realizable and will not be until the relevant software is created or perfected.

All in all, however, the project is very ambitious and the proposers seem to be very clear in their aims and at the same time honest about potential limits. The proposal is convincing.

**1.b. What is your assessment of the science case?**

Please choose: poor / fair / good / very good / excellent\*

very good

**2. THE POTENTIAL FOR 'BRAIN GAIN' (TALENT CASE)****2.a. Please comment on the talent case presented in the application.**

The proposers claim that the instruments and data offered by the CLARIAH infrastructure will attract top level researchers from the Netherlands and from abroad, pointing out that the Netherlands already has an excellent reputation in the world of research. I can confirm this from my position in another part of Europe. It is known that Dutch projects are often well respected and numbers of researchers are keen to join them. Dissemination through training and education, involving most humanities faculties in the country, can be expected to create 'a whole new generation of researchers', harking back to the youngsters mentioned in the previous section. I think this is no idle boast. A part of the budget is dedicated to this 'brain gain' and this project, together with all the other initiatives with which CLARIAH intends to be involved and which are constantly referred to in the proposal, definitely has the potential to attract foreign researchers. The talent question should be no problem.

**2.b.**

**What is your assessment of the talent case?**

**Please choose: poor / fair / good / very good / excellent\***

excellent

### 3. SOCIAL AND COMMERCIAL RELEVANCE (INNOVATION CASE)

**3.a. Please give your opinion on the innovation case presented in the application.**

The proposal points out correctly that industry is not indifferent to the findings of humanities research. Many businesses recruit humanities trained staff. The desire for collaboration on the part of high tech companies is described as 'remarkable'. Innovation is seen in terms of a cross-fertilisation of scientists and humanists, with commercial products being the spin-off. Audiovisual products and entertainment-based products are particularly fertile areas for creative synergy, also at international level, as the constant demand for new media products increases. The provision of open access and open data will lead to international exchanges of private sector users. On the other hand software tools that reveal information can be of great interest, for example, to the police or to stock exchange agents. Thus, notwithstanding the general impression that industry stays clear of humanities activities, the proposal makes a good case for the social and commercial relevance of its infrastructure programme.

**3.b. What is your assessment of the innovation case?**

**Please choose: poor / fair / good / very good / excellent\***

very good

### 4. COLLABORATION AND COMPETITION (PARTNERSHIP CASE)

**4.a. Please give your opinion on the partnership case.**

The proposal claims (see Science case) that all humanities faculties in the Netherlands are involved in the project, which is a good start. Mention is constantly made throughout the proposal to the many other major and minor initiatives and projects with which CLARIAH will collaborate, particularly CLARIN, DARIAH, DARIAH-EU, DASISH, Flarenet, Clio-infra, HASN, EHRI, etc., with meetings between the various directors already planned. The pan European ventures such as DARIAH-ERIC and ESFRI are also to the fore, though some of the ventures proposed are qualified by the phrase 'it is expected that'. In any case it is easy to see that a critical mass can be reached leading to the sharing of resources and technologies. 'It is natural to join forces' as the proposal states. A general assembly envisaged for the project will also supervise subprojects within the main CLARIAH infrastructure based on already tried and tested models, for example those established in CLARIN-NL. Work packages on the EU model are planned. The creation of 'users panels' in the dissemination and outreach stage show that efforts will be made to establish fruitful contacts with public organisations and private enterprise. Finally an International Advisory Panel will advise the general assembly. Fig. 2 outlines the governance structure. As long as CLARIAH does not simply replicate the actions of some of the other collaborating bodies, the partnership structure would seem to be very sound indeed.

**4.b. What is your assessment of the partnership case?**

**Please choose: poor / fair / good / very good / excellent\***

excellent

### 5. FINANCIAL ASPECTS (BUSINESS CASE)

**5.a.**

**Please comment on the business case made in the proposal (please notice that even in case no funding is requested for, applicants must fill in subsections 5.1 and 5.2 of the business case)**

The total budget costs, according to NWO categories, are outlined in tables. The sums involved are large though presumably consonant with the National Roadmap for Large-scale Research Facilities funding arrangements. 87% of the budget will go on technical work packages, though the number of organisations involved, directly or indirectly makes it 'difficult to disentangle the national component from EC grants and investments'. Office and workspace costs and administration will be covered by the participating organizations. The financial feasibility of the project is not really addressed except in the statement that 'the requested NWO funding is equal to the total budget'.

**5.b. What is your assessment of the business case?**

**Please choose: poor / fair / good / very good / excellent\***

very good

### 6. TECHNICAL FEASIBILITY/TECHNICAL CHALLENGES (TECHNICAL CASE)

**6.a. Please comment on the technical case made in the proposal.**

Integration with already established major initiatives such as CLARIN and DARIAH is a guarantee of a certain level of technical know-how. We are dealing with 'professional and robust common infrastructure services'. The risk of technical challenge is thus reduced. However, some of the technical advances discussed are still 'under development'. Similarly, while 'interoperability' is rightly considered a major strength within the project's goals, it is admitted that 'achieving semantic interoperability is not easy', and the management and operation of web services 'requires significant attention in the project'.

CLARIAH will provide an open system which can work in different research environments, another risk reducing feature. A chart shows risk analysis in terms of likelihood, impact, preventive action and contingency plans. For example the risk of overestimation of effort can be prevented by early checking of deliverables and, if necessary, the contingency plan of reallocating tasks and priorities can be brought into play. If national ICT infrastructures for storage do not emerge, storage can be shifted elsewhere or more funds will need to be sought.

Given this attention to risk management, which is promised to be 'continuous', the possible technical shortcomings inherent in the various uses of 'may be', 'is very likely' or 'needs to be addressed', are not likely to lead to unsurmountable challenges. The technical feasibility of the project, while not 100%, is reasonably convincing.

**6.b. What is your assessment of the technical case?**

Please choose: poor / fair / good / very good / excellent\*

very good

**7. POSSIBLE FOCUS FOR THE NETHERLANDS**

**7.a. Please comment on the possible focus for the Netherlands.**

The proposal reiterates at frequent intervals that the Netherlands are a leading partner in the kinds of projects under discussion. For example, Utrecht University has a leading role in CLARIN-EU and the Netherlands will host CLARIN ERIC, all associated with CLARIAH. It is known that the Netherlands are in the forefront in many fields including those relevant to the project such as repository developments and a number of historical research domains. The HSN is recognized at world level. This leads the proposers to assert that CLARIAH 'will strengthen the position of the Netherlands' and 'will enable our country to take on a new driving role in the transformation process that the arts and humanities research is currently going through'.

**7.b. What is your assessment of this criterion?**

Please choose: poor / fair / good / very good / excellent\*

excellent

**8. CRITICAL MASS**

**8.a. Please give your opinion on the criterion "Critical mass".**

Again by referring to the myriad organisations and schemes connected to CLARIAH it can be verified that critical mass is being achieved. The entire humanities community in the Netherlands (a grandiose claim) is said to be involved as well as tens of thousands of foreign researchers who will be brought in through a process of continuous networking.. Figures are provided to support these claims. Internationally renowned programmes and organisations are involved; my personal experience supports the claim that linguistics in the Netherlands enjoys international acclaim. The HSN (Historical Sample of the Netherlands) is cited as 'a model of multidisciplinary collaboration and outreach'. In terms of critical mass in the humanities sector the Netherlands is well placed.

**8.b. What is your assessment of this criterion?**

Please choose: poor / fair / good / very good / excellent\*

excellent

**9. EMBEDDING**

**9.a. Please give your opinion on the criterion "Embedding".**

Again the question of the contribution of all interested parties in Humanities research is recalled, even at the risk of repetition, the Europe-wide CLARIN and DARIAH infrastructures again being mentioned. In effect it is difficult at times to provide responses that go beyond questions that have already been addressed. However the importance of international networks is stressed with reference to comparable initiatives in the USA, Japan, Korea, etc., and mention is made of several related research projects running parallel (IMIX, STEVIN, ERC, the Multilingual Project, etc.)

Particular emphasis is given to DANS (part of the DARIAH organization) which receives funding from NWO. As DANS is responsible for access to Dutch arts, humanities and social science data, there is 'a sound embedding of CLARIAH activities in the Dutch research community'.

CLARIAH is represented in a number of important research institutions such as the APARSEN Network of Excellence.

The general impression is that CLARIAH would be embedded in an already established and well-reputed network of projects, initiatives, associations and organizations, including the private sector. This is particularly true of the Dutch situation; the foreign connections are less clear but undoubtedly exist. All this of course requires careful

planning and sound governance but the impression given is that the main promoters are known to one another, have worked together in the past, and are moving in the same direction (see next section).

**9.b. What is your assessment of this criterion?**

Please choose: poor / fair / good / very good / excellent\*

excellent

**10. PROVEN WILLINGNESS TO COLLABORATE**

**10.a. Please give your opinion on the criterion "Proven willingness to collaborate".**

As mentioned above there is proven collaboration among the many bodies involved in a whole series of connected research ventures, projects, sub-projects and so on. Flanders has been brought into the picture for Dutch language ventures. One example quoted is that of Clio-infra within which scholars in sister institutes cooperate, building on national and international research programming. Foreign partners mentioned include the University of Michigan at Ann Arbor and Umea University.

As a particular example of collaboration, five KNAW institutions worked together successfully on the Alfalab project which will lead straight into the Tex Kit element of CLARIAH.

With CLARIN and DARIAH becoming ERICs, the existing cooperation, which seems durable and functional, can be expected to continue at a European level.

CLARIAH thus promotes itself as a maximiser of potential for synergy. This remains to be seen but there is no doubt that a proven willingness to collaborate exists.

**10.b. What is your assessment of this criterion?**

Please choose: poor / fair / good / very good / excellent\*

excellent

**11. REFLECTION OF SOCIAL TRENDS**

**11.a. Please give your opinion on the criterion "Reflection of social trends".**

The data gathering and information extraction envisaged by CLARIAH will definitely have some effect on social and political questions. An example provided is that of migration and mobility, suggesting that data analysis can help discover changes and trends and thus positively affect government policy. The barely concealed claim that CLARIAH could have helped prevent the 9-11 attacks and the murder of Theo van Gogh is, however, pure speculation.

It is undeniable that digital databases are essential in today's society, though size alone is a double-edged weapon: there is too much information out there! CLARIAH claims that it will be able to cut through and filter this mass. The favourite term used to describe how textual and audio-visual tools will operate on data is 'seamlessly' though the technical problems involved may provide some obstinate seams.

One of the pillars of CLARIAH is listed as ICT for social challenges. The social, economic and political challenges of the Netherlands are similar to those elsewhere in the western world, some of which can only be addressed, it is asserted, by analyzing data on long-term trends. The claim made by CLARIAH that problems such as the negative effect of parental divorce or the question of below-replacement fertility can be better understood through longitudinal analysis of large quantities of data, is certainly of great interest but to be totally convincing we will need some concrete proof. Many other, though less ambitious, databases exist and have been used in social policy research, though without as yet yielding any really ground-breaking results. Nonetheless, the CLARIAH project, given its embedding in a serious and dynamic national framework, perhaps promises more than most.

**11.b. What is your assessment of this criterion?**

Please choose: poor / fair / good / very good / excellent\*

very good

**SUMMARY AND OVERVIEW BY REVIEWER**

.a.

**What is your opinion on the entire application? Please justify your overall assessment by summarizing or briefly commenting on the strengths and weaknesses of the proposal.**

The first strength of the CLARIAH project lies in the extremely competent environment in which it wishes to be embedded. The proposers seem to have an excellent working knowledge of the projects they wish to fit CLARIAH into and thus should have little difficulty in moulding their facility into the overall national framework. For the same reasons the reputation earned by many of the other associated projects and institutions should attract non-Dutch (and non-Dutch speaking) researchers.

The second main strength is the proposed harvesting (and subsequent filtering) of large amounts of material from the humanities sector. This could well be one of the largest and most useful collections of such data yet attempted. Certainly the arts and humanities are in need of some judicious modernization, also in order to justify equal status with scientific research.

One weakness lies in the conditional approach to some of the innovations proposed, though this shows a desire

for honesty. An explanation as to how 'seamless interoperability' is to be achieved would have been useful. As regards the international angle, it is not sufficiently clear how Dutch-centric the facility will be. While the Netherlands maintains a high reputation in research circles, including humanities research, much of the work is carried out in and presented in languages other than Dutch.

**.b. What is your overall assessment of the entire proposal?  
Please choose: poor / fair / good / very good / excellent\***

excellent

**Adviseur: 4****1. THE LIKELIHOOD OF SCIENTIFIC BREAKTHROUGHS (SCIENCE CASE)****1.a. Please comment on the science case made in the application.**

The general nature of the science case is well made, coherent and credible. The general aspiration of the proposal is impressive and appears to be in the forefront of eHumanities thinking in the world. The proposal however could more fully have developed the integration of its argument by providing more and clearer examples of individual research fields and inter-and trans-disciplinary collaborations that could be conceived from the CLARIAH.

**1.b. What is your assessment of the science case?**

Please choose: poor / fair / good / very good / excellent\*

Very good

**2. THE POTENTIAL FOR 'BRAIN GAIN' (TALENT CASE)****2.a. Please comment on the talent case presented in the application.**

I consider this to be the most promising dimension of the proposal. It is highly probable that, if successfully implemented, CLARIAH would attract innovative scholars in the various humanities disciplines confident that what would be available to them would be a world class facility for integrated humanities research. I imagine that the talent case will include not merely the research that can be stimulated and deepened by CLARIAH but also the collaborative and communicative innovations, the conceptual developments and the discursive shifts that are afforded and produced by so much integration. Seminars to disseminate knowledge of outcomes will in themselves become promotional practices for Dutch innovation and have a strong reputational effect. Research infrastructure allows wider, deeper questions to be posed and tested and the meta-reflective understandings these will stimulate will surely be very attractive because of their promise of overcoming sterile discipline centred limitations so prevalent even today.

**2.b. What is your assessment of the talent case?**

Please choose: poor / fair / good / very good / excellent\*

Excellent

**3. SOCIAL AND COMMERCIAL RELEVANCE (INNOVATION CASE)****3.a. Please give your opinion on the innovation case presented in the application.**

The innovation case also appears strong, though precise connections between 'industry' and 'business' worlds and humanities researchers could have been conceptualized better. The overall case for improved science literacy, public communication and efficient marshaling of information sources available in dispersed locations, and in non- or pre-structured forms appears very strong, but this was not really drawn out in the application. Nevertheless I am convinced the innovation potential is immense. Some of the commercializable prospects of tools, eg the Clarin-NP WIP subproject tools are suggested to afford application to parliamentary discourse monitoring, corporate product perception. I have no doubt that spin-off applications would flow to many communication based and persuasion dependent practices in public and commercial life.

**3.b. What is your assessment of the innovation case?**

Please choose: poor / fair / good / very good / excellent\*

Excellent

**4. COLLABORATION AND COMPETITION (PARTNERSHIP CASE)****4.a. Please give your opinion on the partnership case.**

The case for partnership and deep collaboration in an ambitious, multi-sector and multi-modal innovation such as CLARIAH is palpable. I am unfamiliar with what I might call the collaborative-competitive disposition of Dutch institutions, and their collective or individual relations with European and non-European counterparts. I am however very familiar with the collaborative-competitive dispositions of individual Dutch humanities scholars, which in my experience is positive and deep. I can only surmise from the letters of support, the extent of discussion that would have gone into eliciting this level of support and the manifest benefit that will derive from this project that the partnership case is extremely strong. The governance structure appears well judged, a reasonable balance between monitoring and accountability without becoming overly bureaucratic or intrusive.

**4.b. What is your assessment of the partnership case?**  
Please choose: poor / fair / good / very good / excellent\*

Excellent.

**5. FINANCIAL ASPECTS (BUSINESS CASE)**

**5.a.**

**Please comment on the business case made in the proposal (please notice that even in case no funding is requested for, applicants must fill in subsections 5.1 and 5.2 of the business case)**

I am least able to assess this criterion in precise detail however I have extensive experience with budgeting for large scale proposals in several settings. Therefore my methodology for assessing the financial aspects of the case was to contrast and compare with a Swiss precedent and two Australian precedents of large, though in all cases smaller than CLARIAH, collaborative mulch-institutional Humanities projects, by selecting cost items in the budget for comparison. I also studied the justifications for items of expenditure, their duration and level. Using this approach the business case appears to me to be strong.

**5.b. What is your assessment of the business case?**  
Please choose: poor / fair / good / very good / excellent\*

Very good.

**6. TECHNICAL FEASIBILITY/TECHNICAL CHALLENGES (TECHNICAL CASE)**

**6.a. Please comment on the technical case made in the proposal.**

I defer to other assessors for most of the technical case involved.

**6.b. What is your assessment of the technical case?**  
Please choose: poor / fair / good / very good / excellent\*

I prefer not to offer an assessment of the technical case as it is outside my field of expertise.

**7. POSSIBLE FOCUS FOR THE NETHERLANDS**

**7.a. Please comment on the possible focus for the Netherlands.**

As I have implied in comments above I believe that this is a worthwhile focus of Netherlands research investment. CLARIAH is poised to ask large research questions which will benefit by facilitating the interrogation of knowledge stocks conserved in diverse ways, some structured and shaped and others unstructured and diffuse. Individual Dutch institutions will have available to them, and the Dutch research community likewise, a potentially unique aggregation of resources, what might be called a 'capstone' perspective. If the collaborative and partnership dimensions involve regular seminar based discussion of emerging questions it is highly likely that the Netherlands will be at the centre of questions of intersection between humanities disciplines and public policy problems. This dimension was not brought out in the application but it seems to me an important 'lurking' possibility.

**7.b. What is your assessment of this criterion?**  
Please choose: poor / fair / good / very good / excellent\*

Very good

**8. CRITICAL MASS**

**8.a. Please give your opinion on the criterion "Critical mass".**

My impression is that Dutch researchers and Netherlands based scholars in general are overrepresented in academic prestige in particular fields. I know this to be true of language acquisition research, historical studies and some other fields. The proposal is assured of achieving critical mass by its very design specifications.

**8.b. What is your assessment of this criterion?**  
Please choose: poor / fair / good / very good / excellent\*

Excellent.

**9. EMBEDDING**

**9.a. Please give your opinion on the criterion "Embedding".**

As a non-European I am not able to judge this criterion except to offer a probabilistic assessment. Prior to reading this criterion I had reached a conclusion that the project showed high levels of institutional 'embeddedness'. This is impressionistic, but it is based on extensive experience with applications for similar kinds of multi-organisation and multi-modal collaborations in the humanities.

**9.b. What is your assessment of this criterion?**  
Please choose: poor / fair / good / very good / excellent\*

Very good.

**10. PROVEN WILLINGNESS TO COLLABORATE**

**10.a. Please give your opinion on the criterion "Proven willingness to collaborate".**

This appears credible and well founded.

**10.b. What is your assessment of this criterion?**

**Please choose: poor / fair / good / very good / excellent\***

Very good.

**11. REFLECTION OF SOCIAL TRENDS**

**11.a. Please give your opinion on the criterion "Reflection of social trends".**

I was disappointed in this dimension of the proposal and seemed to me to reflect a failure of imagination. When the innovation case is defended by reference to a variety of social benefits, such as monitoring the attitudes and positions of Dutch politicians with regard to important national and international problems it seems odd that national social developments and trends cannot be reported, even prospectively.

**11.b. What is your assessment of this criterion?**

**Please choose: poor / fair / good / very good / excellent\***

Poor

**SUMMARY AND OVERVIEW BY REVIEWER**

**.a.**

**What is your opinion on the entire application? Please justify your overall assessment by summarizing or briefly commenting on the strengths and weaknesses of the proposal.**

Overall the proposal is excellent, but it is slightly marred, however, and unfortunately, by little attention being paid to a solid consideration of social trends, national public policy and even international or global practices. In these latter areas the likely impact of networked, highly integrated 'common lab' infrastructure for the humanities could be substantial and would repay more effort of imagination and discussion. Nevertheless the talent, commercial, innovation and partnership dimensions are quite outstanding.

**.b. What is your overall assessment of the entire proposal?**

**Please choose: poor / fair / good / very good / excellent\***

Very good to excellent.