# 1. General Comments

Overall, the report shows that the consortium is doing a good job. At last year’s meeting, the CS expressed some reservations about an apparent lack of transparency concerning its organization and structure; this year’s report is admirably explicit and thorough in its coverage, and such concerns are addressed well also by the consortium's website, which should perhaps be given more prominence in the report.

The statement of new objectives (1.1.) begins by talking about the creation of a network of researchers in the digital humanities. But there are many such networks. Why not specify the particular thematic interest (authorial texts) which distinguishes CAHIER from other research communities in the SHS? The same section talks of the need to address the “potentiel de recherche” and the “création de nouveaux savoirs”, which is indeed a laudable objective. It would be good to see some specific examples of the kinds of research and knowledge creation envisaged, even better to see some evidence of their effect within the traditional disciplines. We applaud the explicit statement that the development of the Consortium is essentially user-driven (“*suit les tendences et les besoins de sa communauté*”). The phrase “*humanités numériques des textes*” appears here for the first time as a means of implicitly defining the “community” envisaged; maybe that definition should be more explicitly addressed, if not in this report, then on the website.

As a general comment, the CS notes that the report enumerates the activities of the Consortium bery well, but does not attempt any assessment of their impact. This will require the development of metrics by which the Consortium would measure its performance.

# 2. Constitution of the Consortium

The criteria for accepting new members are fairly general: it seems a little surprising that there were only four requests to join the Consortium; certainly the pool of projects interested in CAHIER, even simply at the French level, goes beyond these 4 projects. Why was the fourth applicant rejected ? It might be useful for people discovering the consortium for the first time on the website to know not just where projects are based geographically, but also what state of maturity each project has achieved: adding the date of joining the Consortium to the summary pop up might do that; distinguishing visually those projects which already make materials available online would also be useful. It would also be good to see the criteria by which projects apply.

The CS notes that any acceleration in the development of CAHIER would require a significant increase in shared resources to maintain the same quality of shared activity. However, the future of the consortium must clearly be considered in conjunction with the question of its enlargement: the Consortium needs to decide whether it wishes to continue to grow indefinitely, or whether it would be better to consolidate its achievements so far. In this context the CS suggests that a consideration should be given to a possible COST Action on the central theme of CAHIER : this might be a good mechanism for European expansion as it comes with significant funding for Networking, and would also serve as a good basis for further EU funding.

Nevertheless the CS encourages the Consortium to investigate the conditions necessary for a gradual opening up of the Consortium beyond its current constituency. The idea of creating a category of “*Membre associé international*’ is a step in the right direction. The COPIL should work out criteria and an appropriate workflow, seeking advice from the CS once a draft of these is available. Within Europe, it seems that we can distinguish two types of potential targets: projects based in communities that are equally competent and well funded in their use of DH, and projects in less advanced communities with which there may be strategic advantages in cooperating.

# 3. Activities of the Consortium

The funding made available by the Consortium to support member projects is another activity for which take-up seems surprisingly low. This is basically a marketing and positioning issue. Would the Consortium consider funding specific groups of researchers (e.g. early-career researchers) who could be more precisely targeted? Or specific types of activity (e.g. training in relevant technologies)? Short reports from these bursaries could be made more prominent on the website, when for example conference attendance at a conference has been supported. The Consortium should of course be credited in such presentations. In the report the “Objet de la demande” should be given in full. At present, the text seems sketchy. The same goes for the paragraph which talks about the call for proposals. Perhaps it could also be explained why there were “problems” with the call. If mistakes were made they should be explained and taken as incentives to do better in the future (see also 5.2.4).

Support for the participation of Cahier members or affiliates in training events (2.6) or international conferences (2.7, 2.8) is a useful benefit if it leads to greater visibility for the Consortium and if the expertise gained can be shared with other Consortium members. It’s not clear what benefit flows from the “labelisation” of major non-CAHIER events such as the ANF or EDEEN (2.6 -2) unless the event concerned actually has a specific CAHIER focussed session (e.g. a session presenting CAHIER projects). In any case, the Annual Report should perhaps include links to background information about such events (not all readers of the report know what the ANF is), as well as a brief report from the beneficiary of such support. These considerations apply equally to the “*journées d’etude*” mentioned in section 2.9. The CS would be happy to participate in the proposed discussion of the “*modalités de labelisation”*, which, we suggest, should be consistent across all types of event (*journées d’étude*, foreign conferences, training sessions, etc.) and should be clearly stated on the website.

The CS recommends some explicit consideration of such questions (Do we want to labelise? Why do we want to labelise? What is the ‘plus value’?) -- requests for labelisation are very frequent and the Consortium’s label has considerable value. The work done in CAHIER is at the forefront of what can be done in DH and its label should be a guarantee of quality. The CS suggests looking to models elsewhere, for example the MLA’s Seal of Approval for Scholary Editions (which it has expanded in recent years to include Digital Scholarly Editions: <https://www.mla.org/About-Us/Governance/Committees/Committee-Listings/Publications/Committee-on-Scholarly-Editions>) in making such an assessment.

Cahier’s Work Groups (GT) are one of its more distinctive and valuable features. The CS has suggested that perhaps the constitution of these groups might be more explicitly formalised: certainly there is some variation in the way they are presented on the website. It might be useful to have a single page listing all work groups, giving for each a brief “mission statement” , its constitution, its chronology, and where appropriate links to its outputs/reports. The drop-down list of “activities” on the website doesn’t (for example) mention the EVENT or dataCahier work groups at all, though the former at least has a very interesting draft page on the website, if you look for it. Equally, the current Annual Report makes no mention at all of the Correspondence WG, even though it has produced an excellent report this year, which is listed on the website, along with the equally excellent reports from the GT on legal issues (questions juridiques). These reports are full of useful practical advice and deserve wide circulation; the Consortium should however consider how best to ensure they are maintained, distributed, and kept up to date.

The GT “Questions juridiques” (3.1) should seek to establish links with other groups which focus on legal questions, above all on the European level (see for example CLARIN EU Legal Issues Committee <https://www.clarin.eu/governance/legal-issues-committee>). The RGPD is a European law with major implications for research. Concerted actions might be needed at some point to guarantee continued access to open data in support of Open Science. Legal matters are of course a huge territory (the report hints itself at this: “*ampleur de la tâche*”) in which collaboration is essential: one GT cannot do everything, as noted in section 3.7. It seems correspondingly important to have a clear picture of the GT’s individual perspective. But although the text mentions a “guide” it does not make explicit its contents.

The activities of the Data\_Cahier WG in particular need to be better publicised and more transparent. Provisioning of the OAI “entrepôt” is an important activity for the future of all the Cahier projects, but the website doesn’t explain how this should be done, presumably by means of a properly constituted TEI Header. (Is there a CAHIER-specific minimal Header? Where is it documented and how was it developed?) Similarly tantalising is the page on the curiously-named “R2Cahier” work group : this mentions (but doesn’t link to) a number of documents and publications which ought perhaps to be made more accessible; it also does not appear to have been updated since 2017, and so doesn’t provide any detail of the work mentioned rather sketchily in this year’s annual report. The CS notes that testing of tools based on models other than the tree model is really very interesting in the treatment of co-reference. It seems that several international groups are working on the issue, for example the W3C (https://www.w3.org/TR/WD-xml-link-970406.html).

The Typologie WG has an interesting wiki page, and is exemplary in its listing of meetings, but the wiki has not apparently been updated since 2016: and there is no link in the Report to the current prototype of the “OPEN-Theso” based thesaurus. Finally, if as noted in the report, the currently dormant Crowd Sourcing WG aims to prepare a new report on best practice, it is to be hoped this will be informed by the experience of other French projects of this type, notably that of the Archives Nationales.

Cooperation with CORLI and other Consortia should be sought not only with respect to legal matters (see 3.7) but also for the discussion of conceptual matters: Resources like “oeuvres” become a corpus and the corpus becomes data which can be analysed from many different perspectives and surely not only the ones we foresee from the point of our individual disciplines and interests. This will have to be taken into account also when making the editions available.

The CS wishes to underline the importance of this and other international activities to create common reflection, exchange, visibility of the Consortium. It is disappointing that the Report says so little about such cooperation: what progress has been made? What kind of co-operation is envisaged? The Consortium is well-placed to contribute to the goal of making DH truly international beyond the anglophone sphere, facilitating the development of different ways of doing research in the global DH community.

# 4. Impact of the Consortium

The most important aspect of Cahier’s work is likely to concern the development of new research practices profiting from its interdisciplinary focus and its use of information technologies in disciplines where both these aspects have traditionally been under-used. Section 5 of the report could therefore be enhanced by citing some specific cases where participation in Consortium has had demonstrable scientific effect, in terms of facilitating otherwise impossible research for example. The section currently reiterates the objectives stated initially, which is useful and appropriate, but does not go beyond that to consider the scientific effects of all this activity. For example, it is good that the Consortium’s resources should be catalogued and accessible in a uniform way via Isidore. But the specific benefits consequent on that accessibility need to be stated too. The importance of these for long term support of the Consortium should not be under-estimated. Many DH projects (to generalise wildly) have a tendency to speak only to and about other DH projects: it’s important to ensure that Cahier’s activities are seen to be contributing to the traditional, non digital, disciplines as well. It could also cite some cases which show that questions coming from the Humanities require an interdisciplinary focus, and can be better answered by taking into account factors, phenomena and the like which “traditional” and often highly specialised Humanities disciplines have tended to ignore.

Overall, the CS is concerned that in section 5.1 the *plus-value apportée* is very much reduced to tools and resources and making things available. Research -- the creation of new knowledge – seems to have been forgotten.

Consortium members now have at their disposal a significant amount of data between all its projects. What might it do to leverage this beyond creating editions? Another way of putting this: is the sum greater than the individual parts of the Consortium’s outputs? Such considerations could be addressed during a phase of consolidation before current funding ends. They would thus drive an agenda that is relevant, forward looking, and research-oriented.

5.2.1 could also be phrased in a different way. It sounds quite weak. Perhaps growing all the time in quantitative terms is not the most important. Perhaps CAHIER has reached a steady state, beyond which further expansion is unnecessary. Perhaps the absence of further growth is a good thing. In any case, as noted above, the Consortium needs to plan its future composition carefully.

The statement in 5.2.2 that “*peu de formations ont été nécessaires*” is perhaps questionable, especially since it is followed by discussion of many topics in which training is both necessary and under consideration! Perhaps the intention was to contrast “basic level” training concerning the creation of digital resources with more “advanced” training on their exploitation, but this is a false dichotomy, since what you can do with a digital resource is intimately bound up with the way in which you have created it. Doing research on this aspect in groups might be interesting.

Section 5.3 on “difficulties encountered” apparently considers only the challenge of devising an exit strategy or future development plan for the Consortium’s activity. The CS does not underestimate the difficulties associated with this, and applauds the Consortium’s willingness to confront them now, but finds it seems a little implausible that this should be the only challenge CAHIER has had to address during the past year.

As regards the possible “*pistes*” mentioned for the future, the CS doubts whether scholarly associations will be of much use, since these are representative bodies only, with no remit to create resources or to do research. Instead, it suggests, Cahier must engage in a close dialogue with the academic community. In which context, it seems that one activity missing from the plans for 2019 might be called evangelisation [“promoting”]: some efforts to promote a better understanding of the skills and methods characterising the network to the broader SHS research community. This is a difficult thing to achieve, but efforts made in this direction can only contribute positively to future support of the network beyond current funding. Partnerships could be considered with the “Centre pour l'édition électronique ouverte” (CLEO). The CS is of course very happy to contribute to this effort.

We note also that some further simplification of ways of participating might be advantageous. The distinction between “projects” and “associated projects” remains a bit unclear. The original definition suggests that an associated project is simply one which is ineligible for funding from CAHIER, but the reasons for that ineligibility might be that the project has become inactive, or has had its fair share of funding already – or that its methods and praxis are inconsistent with CAHIER principles e.g. open access. These differences have implications beyond the goal of growing the network: an inactive project can still contribute to the overall effect, can indeed continue to be useful to the community, if only as a counter-example.

Finally, the CS warmly welcomes and endorses this report, which is full of material demonstrating the vitality and scientific value of the Consortium.