



R&I

Webinar Series

Data Cultures in Higher Education:
Emergent practices, professionalism and
the challenge of social justice.

Organizers



United Nations
Educational, Scientific and
Cultural Organization



UNESCO Chair
Education and Technology for Social Change



Universitat
Oberta
de Catalunya

EDUL@B



UOC-EPCE



Universitat
Oberta
de Catalunya

Faculty
of Psychology
and Education
Sciences

Opening Data in Technology Enhanced Learning Research

The challenge, the practices



Dr. Davinia Hernández Leo

Full Professor, Department of
Information and
Communication Technologies
University Pompeu Fabra

*Head of TIDE Research Group on Interactive and Distributed
Technologies for Education.*

13 January 2020 – 17hs GMT+1

Series Coordinator
Juliana E. Raffaghelli

Transcripción

“Datos abiertos en la investigación sobre aprendizaje mediado por tecnologías: El desafío y las prácticas”

J: Muy bien, vamos a dar comienzo a este webinar en un momento bastante crítico e inusual. Nos vamos a focalizar sobre el tema de la apertura de datos en la investigación sobre tecnologías educativas y educación en línea, considerándolo un reto y una práctica que se ha de llevar a cabo. Para tratar este tema hemos invitado a la profesora Davinia Hernandez Leo de la Universidad Pompeu Fabra del TIDE Research Group on Interactive and Distributed Technologies for Education.

Este webinar aparece en un contexto de trabajo propuesto por el grupo de investigación Edul@b que es una serie de encuentros sobre las culturas de datos en la educación superior, considerando esta temática como una cuestión emergente que afecta el profesionalismo docente y que se convierte en un reto conectado al equilibrio y la justicia social dentro de lo que son las IES, las instituciones de educación superior. Nos hemos hecho toda una serie de preguntas respecto a si tenemos que considerar la riqueza de los datos con entusiasmo o bien, ante las varias cuestiones críticas que han aparecido, si la presencia de la recogida sistemática de datos por parte de los sistemas digitales es un problema o es una oportunidad. Entonces, ante esta serie de preguntas, la cuestión de recoger o no recoger los datos durante los procesos de investigación en tecnologías educativas y en ambientes mediados por tecnologías es un tema central, pues tiene una connotación positiva pero también tenemos que descubrir sea las dificultades técnicas, sea las dificultades éticas que pueden generarse en este sentido.

Vamos a hacer este encuentro, lo vamos a llevar a cabo a partir de una serie de preguntas y vamos a comenzar con la primera pregunta, Davinia: ¿Por qué has dicho en alguna ocasión pasada que “la ciencia en abierto es una faceta transversal dentro lo que es una investigación en procesos de aprendizaje y educación mediados por tecnologías responsable”?

D: Gracias por toda la presentación ya, Juliana, y encantada de participar en este webinar de esta temática tan importante y, como digo yo, transversal a las diferentes líneas de investigación, ámbitos de trabajo que unen el tema de la educación y la tecnología. Veamos, yo digo que el tema de las prácticas de ciencia abierta, de las prácticas de Open Science, es un tema transversal a todos estos ámbitos de investigación y es un tema de prácticas de investigación responsable. Lo digo porque conllevan un mayor beneficio de la investigación para la comunidad investigadora y para la sociedad, no, el poder llevar estas prácticas de Open Science llevadas a cabo. ¿Por qué esto es así? ¿Por qué estas son prácticas de investigación responsable? Porque las prácticas de ciencia abierta tienen que ver con que los investigadores hagamos nuestras actualizaciones, nuestros temas de investigación, más transparentes, más frecuentemente publicados y con un acceso abierto a toda la sociedad, a otros investigadores en un proceso de investigación. Por tanto, una compartición más transparente de nuestros avances en investigación y una compartición abierta de nuestros avances en investigación. Además, hace que esto que compartimos de forma más frecuente y abierta lo hagamos de una forma más documentada que si no siguiéramos estas prácticas de ciencia abierta. Más documentada para que sea posible en procesos de reproducción de la generación de conocimiento que nosotros hemos llevado a cabo y por tanto también de más verificación, más posibilidades de verificación de esta investigación que llevamos a cabo. Todo ello, el hacer nuestros resultados más abiertos, más frecuentemente compartidos, más documentados

para que puedan ser reproducidos y verificados, hace que el propio trabajo nuestro como investigadores sea más creíble y más facilitante también del progreso de la generación colectiva del conocimiento. Por tanto, son prácticas de investigación responsable y conllevan mayor beneficio de nuestra investigación para la sociedad y para el resto de la comunidad científica. Esto no quiere decir tampoco que no lleven beneficios de cara al propio investigador, es verdad que seguir prácticas de investigación abierta, prácticas de Open Science, es muy costoso para un investigador individual, para un grupo de investigación, conlleva tiempo, conlleva conocimiento de estas prácticas de Open Science, pero también hace que sea más beneficioso el trabajo que está realizando el propio investigador porque hace que este trabajo de investigador pueda tener un mayor impacto en cuanto a sus resultados de cara hacia la sociedad y hacia la comunidad científica. Claro, yo, como vicepresidenta de la Asociación Europea en Tecnologías para la Educación, tecnologías que pueden mejorar la educación, *Technology Enhanced Learning*, hemos visto que, al ser un aspecto transversal de trabajo responsable, nuestra posición como investigadores lo tendríamos que fomentar, al menos la conciencia sobre la necesidad de aplicar prácticas de ciencia abierta en nuestros ámbitos de investigación. Es por eso que, si vais a la web de la Asociación Europea de *Technology Enhanced Learning*, de EATEL, podéis encontrar un webinar donde hemos conseguido que personas representativas de nuestro, del ámbito de *Technology Enhanced Learning* puedan deciros cuáles son los beneficios, los desafíos y las recomendaciones, y las mejores prácticas en ciencia abierta en nuestro ámbito que podemos conocer y aplicar también en nuestros propios grupos de investigación, en los diferentes centros y universidades.

J: Bueno, eso es extremadamente interesante porque justamente lo que probablemente está faltando, y aquí vamos a la siguiente pregunta, es ver más prácticas relacionadas con una toma de conciencia de este enfoque que tú nos comentabas inicialmente. Pero, en este sentido, creo que el TIDE, habiendo conocido vuestro trabajo y vuestras publicaciones, está muy avanzado. Creo que es uno de los grupos seguramente a nivel nacional y probablemente a nivel europeo, considerando tu rol a nivel europeo, Davinia, está muy avanzado en este proceso de apertura de la ciencia y de nuestra materia, que es el *Technology Enhanced Learning*, y por eso también nos vas a poder comentar qué es lo enriquecedor y cuán difícil, o sea cuáles son los aspectos de dificultad. Así que bueno, ¿qué tal si nos comentas?

D: Las prácticas que con nuestro grupo estamos llevando a cabo para poder seguir el movimiento de ciencia abierta son diversas. Por un lado son las prácticas de acceso abierto, de Open Science, donde intentamos cuando posible publicar en revistas de acceso abierto y sino, pues en el repositorio de la universidad siguiendo una aproximación de Green Open Access. Después, a nivel del código de las herramientas que desarrollamos, de las herramientas de tecnología que desarrollamos, intentamos publicarlo en Github como código abierto, y también los datos los publicamos como datos abiertos también, en la medida de lo posible, en concreto en la plataforma Zenodo, con metadatos, incluso no hablamos solamente de datos cuantitativos pero también de datos cualitativos, descripciones del contexto, material complementario que permite documentar mejor, con más detalle, los trabajos de investigación que reportamos en los artículos de investigación que publicamos. Entonces, muchos de nuestros artículos de investigación tienen asociada una entrada de datos abiertos en Zenodo, donde hacemos disponible todo esto que mencionaba antes. De hecho, dentro de Zenodo, pues nos encargamos de gestionar una comunicación concreta que se llama *Educational Data Analytics Community*, y lo que hacemos es que investigadores del ámbito que comparten datos abiertos en esta plataforma de Zenodo puedan asociar estos datos a esta comunidad de *Educational Data Analytics*, y por tanto pueda ser más accesible de cara a nuestra comunidad de investigación porque pueda ahí encontrar una colección de datos, de datasets abiertos. También intentamos hacer metainvestigación de alguna manera, o lo que decimos "investigación sobre procesos de investigación en nuestro ámbito". En concreto, en el caso de las prácticas de ciencia abierta, hemos intentado reproducir estudios ya realizados y ver

cuáles son las dificultades en poder conseguir reproducirlos, a la vez que intentamos que nuestros propios procedimientos puedan ser reproducidos por otros. Y en este sentido tenemos un artículo publicado en el *Journal Computer Assisted Learning* en 2018 que se intitula “Seeking reproducibility: Assessing a multimodal study of the testing effect”, la parte interesante aquí es que reportamos como hacemos esto de la búsqueda de la reproducibilidad mediante la reproducción de un estudio previo y las dificultades que nos encontramos para realmente reproducirlo, y cuáles son los pasos que nosotros seguimos para permitir una reproducción futura, incluyendo aspectos de práctica de ciencia abierta como es el registro previo del diseño experimental. Después, una cosa que también me gustaría remarcar es en nuestro grupo decía que las herramientas de tecnología educativa que desarrollamos intentamos distribuirlas con código abierto cuando es posible, y además también las hacemos disponibles en la nube, las hacemos disponibles para uso abierto, digamos, con ‘código abierto’ digo uso abierto, siendo una aproximación de SAS, la nube de Software Assist Service, el software para uso abierto para que los usuarios finales pero también otros investigadores tengan acceso a las herramientas sobre las que nosotros reportamos evaluaciones, no, o procesos de diseño de las propias herramientas en nuestros artículos para que ellos mismos puedan seguir construyendo sobre ello de forma más fácil o reproducir nuestros estudios o estudios similares que hagan uso de estas herramientas. Es decir, nuestras herramientas no solamente están descritas en los artículos sino que están disponibles para uso a través de Internet. El caso paradigmático de estas herramientas que digo es el *Integrated Learning Design Environment* que es una comunidad, es una plataforma para profesores a modo de comunidad, donde se pueden compartir diseños de aprendizaje y allí lo que hacemos es integrar herramientas de autoría que hemos ido desarrollando en diferentes proyectos en diferentes líneas de investigación, donde también mostramos las diferentes aproximaciones que tenemos para los datos de aprendizaje, datos de diseño y de comunidad, y como ellos pueden apoyar procesos de diseño y de aprendizaje por profesores, etc. Si cualquier persona que esté viendo ese webinar que quiera ver - a que me refiero, no como es el *Integrated Learning Design Environment*, como allí vamos haciendo, integrando y haciendo disponible de forma abierta nuestros resultados de investigación desde el punto de vista de herramientas o características que apoyan comunidades de profesores en entornos tecnológicos - pueden ir allí y verlo. Y verán cómo las diferentes facetas de la tecnología y las diferentes herramientas tienen conexión con los artículos de investigación que vamos publicando y están disponibles para uso abierto.

J: Genial.

D: En cuanto a las dificultades que me preguntabas, ¿no? En cuanto a las dificultades, veamos, la mayor dificultad para nosotros y que nos lleva mucho tiempo para poder hacerlo bien son los aspectos de ética. Claro, en investigación en educación, en investigación en tecnología y educación, involucramos humanos, diferentes ramos de raza, de edad, muchas veces menores de edad, y recogemos datos personales, no. Decías, al comienzo en tu introducción, ¿es esto positivo? Sí, no, porque nos permite investigar el impacto de lo que queremos estudiar en nuestras investigaciones, pero, por otro lado también, conlleva riesgo para los participantes, no, de los que estamos recogiendo datos. Especialmente en este caso cuando hablamos de ciencia abierta, especialmente si encima los vamos a hacer públicos, los vamos a compartir y no solo analizar, y luego presentar de forma agregada, no, en nuestros artículos. Por tanto los aspectos éticos presentan una dificultad en nuestro ámbito de investigación, y por otro lado nuestro ámbito de investigación tiene otra característica que hace difícil también la filosofía de compartir datos y facilitar los procesos de reproducibilidad es el tema de la dependencia del contexto. En educación sabemos, cuando hacemos estudios en escenarios reales, que el contexto tiene un buen impacto en nuestros resultados y esto hace que si queremos que los datos compartidos se puedan analizar de una forma correcta tienen que ser analizados considerando dicho contexto. Por tanto debemos documentar correctamente y documentar bien el contexto en el cual se han recogido dichos datos.

J: Justamente la dificultad que conlleva muchas veces de identificar un contexto, o sea tratar de mantener aspectos clave anonimizandos, porque es verdad que seguramente una de las cuestiones es la anonimización. Aquí está Davinia lo que comentabas del functional environment, del ambiente ILDE, donde hay una integración entre las herramientas de diseño que utiliza el docente en su práctica en conexión con la investigación y bueno, y las publicaciones, creo que, algunas de las más recientes, hay todavía más recientes, por lo que he ido siguiendo del grupo, y es extremadamente interesante observar cómo se enfoca también la enseñanza, la docencia, como una tarea que al utilizar las tecnologías, siendo una tarea de diseño, también tiene un enfoque sistemático que permite una recogida sistemática de datos y en ese sentido también es muy interesante como pone en relación al ámbito educativo de práctica con el ámbito de la investigación educativa. Creo que el ILDE es un ambiente digamos donde hay varias herramientas de gran utilidad, si bien al docente medio le cuesta pensar que es más que un artesano en el sentido de cómo trabaja dentro del aula y con las dificultades con las que se trabaja.

Bueno, esto ya lo comentabas, y era una pregunta específica, un poco lo habías comentado cuáles son los aspectos éticos de recoger datos de personas en situación de enseñanza y aprendizaje, por qué se recogen datos sea de los estudiantes sea de los educadores. Digamos, sé también por haber seguido vuestro trabajo que tenéis, digamos, líneas de investigación también sobre los aspectos éticos del consentimiento informado a través del cual se les pide a las personas de sus datos. Pero bueno, si quisieras agregar algo en relación a esto, aquí está digamos la referencia al trabajo.

D: Muy bien. Sí, veamos, en nuestro empeño de tratar de hacer nuestra investigación lo más éticamente posible, pues al final decidimos incorporar la ética como línea transversal al igual que lo hacemos con la ciencia abierta. Una línea de trabajo sobre la que también investigar, no, como objeto de investigación, de una manera digamos humilde, no, porque hay muchas cosas que se pueden investigar en ética en nuestro entorno y hay personas que se dedican más al ámbito de trabajo de la ética, pero viendo cómo eso puede contribuir de alguna manera a investigadores como nosotros, que tratan de recoger datos y hacerlos disponibles incluso de manera abierta. Entonces una de las cosas que miramos en nuestros procedimientos de recoger consentimiento informado que son unos procedimientos globalmente aceptados y que seguimos en nuestra comunidad de investigación de forma abierta y lo hacemos porque sabemos que la compartición de datos conlleva riesgos y lo dicen no solamente los procedimientos éticos sino que también lo dice la ley. Lo que veíamos es que estos procedimientos de consentimiento informado o con estos procedimientos nos podríamos estar relajando sobre todo ahora que los avances en las tecnologías del aprendizaje están aumentando los riesgos sobre los que exponemos a nuestros participantes en nuestras investigaciones. Por ejemplo en investigaciones que tienen que ver con la recogida de datos fisiológicos, pues nos acercamos a investigación casi médica; además cada vez estamos recogiendo más datos provenientes de múltiples fuentes por lo que permite la tecnología, pero además la tecnología, los avances en tecnología nos permiten agregar mejor en la análisis de estos datos y con técnicas de aprendizaje automático somos capaces de extraer más conocimiento que antes de estos datos. Esto es bueno porque nos permite avanzar más con nuestros resultados de investigación pero a la vez también hace que los riegos potenciales sobre lo que podemos inserir de los datos de nuestros participantes sean más altos. También otra cuestión es que con tecnologías de, con el uso de tecnologías en la educación también podemos recoger datos que van más allá de los datos del aula con el aprendizaje informal, etc. Por tanto, como los datos, no solamente hay cada vez más datos disponibles sino que además los podemos entender mejor. Con el uso de los ordenadores veíamos que era importante trabajar con más detalle, por otro lado veíamos que la investigación en otros ámbitos más sensibles, decía, a los que nos vamos acercando cada vez más, habían estado explorando hasta que punto los participantes en la

investigación tienen el conocimiento previo suficiente sobre los riesgos de compartir datos como para entender de forma adecuada los formularios de consentimiento, porque a veces no tienen el conocimiento previo suficiente sobre estos riesgos o tienen unos conocimientos sesgados o han hecho malas interpretaciones, eh, sobre ello. Entonces, lo que veíamos es que si lo que dicen otros estudios en ámbitos por ejemplo clínicos es que esto tiene un impacto en que ellos sean suficientemente libres o que estén suficientemente informados para tomar una decisión sobre si van a participar o no en la compartición de datos, y esto es lo que hemos hecho en un artículo, lo que hemos hecho en una investigación para entender cuál es el conocimiento previo de los participantes en una investigación y lo que veíamos es que hay déficits en la conceptualización de cuáles son los riesgos en la compartición de datos en profesores y en estudiantes eh, en ambos colectivos. Entonces claro, esto tiene implicaciones en cómo diseñamos los formularios de consentimiento, cómo comunicamos la investigación que vamos a realizar a los participantes para que ellos puedan tener, estar suficientemente informados para tomar una decisión en cuanto a su participación.

J: Excelente. ¿Y sobre esto estáis generando también material, por ejemplo divulgativo, que ayude a los mismos docentes si tienen que, si quieren realizar un proceso de trabajo, de estudio, digamos, de investigación didáctica, o bien otros investigadores o doctorandos, lo estáis pensando como una posible orientación futura también de formación, por lo menos de formación de investigadores?

D: Sí, bueno, lo que hemos hecho, que también explicamos en el artículo, hicimos un marco de comunicación con los elementos que tenemos que tener en cuenta cuando se explican los beneficios y riesgos de una investigación no, porque, cuáles son los riesgos en la compartición de esos datos y por qué estamos, por qué se propone que se compartan, no, cuáles son los beneficios, eh. No solamente es la mirada negativa de que no se conocen los riesgos suficientes entonces no se toma la decisión informada, pero, a ver, a veces también lo que pasa es que los beneficios no los explicamos bien y entonces tampoco estamos facilitando que esta es una decisión informada hacia lo positivo no, eso es en ambas direcciones. Entonces lo que proponemos en el artículo como conclusión y también es un elemento que nosotros incluimos en el propio diseño de investigación de este artículo es un marco de comunicación sobre los riesgos y los beneficios de participar en una investigación que lleva la compartición de datos y entonces hemos generado unos materiales, unos vídeos, que son los que utilizamos nosotros en esa investigación, y entonces estos vídeos pueden inspirar a otros investigadores en nuestro ámbito, por supuesto el marco de comunicación para crear sus propios materiales a utilizar, en sesiones donde se va a recoger el consentimiento informado.

J: Genial. Excelente. Bueno, esta es una pregunta que probablemente ya no tiene la base sólida que tenía en vuestro trabajo de investigación en el TIDE pero que tiene que ver con una percepción de escenarios futuros. Un poco tus respuestas iban en esa dirección porque al generar recursos, al generar ambientes y al generar digamos meta investigación, hay digamos una tendencia a generar información, sea experta sea de tipo divulgativo, para orientar las prácticas. La percepción es que todavía el movimiento hacia la ciencia en abierto en el ámbito de la educación y de las tecnologías educativas es lento, y que todavía hay necesidades. ¿Cuál es digamos tu percepción de dificultad y como de oportunidad en este sentido?

D: Ehm, a ver, yo lo que veo es que tenemos algunas dificultades en nuestro campo, decía anteriormente, hay complejidades por la propia naturaleza del dominio, hay muchos factores, no, que se involucran, muchos relacionados con el contexto. Y después hay todos los temas éticos, no, y eso hace que quizás estemos un poco más lejos de poder seguir las prácticas de ciencia abierta que otras disciplinas, aunque diría no tan lejos, también las otras disciplinas tienen problemas. Pero,

quizás la naturaleza de nuestro dominio lo hace más difícil. Yo creo que para poder avanzar y conseguir que todos trabajemos de forma más responsable en nuestros procesos de investigación y por tanto involucremos los procesos, las prácticas de ciencia abierta, yo creo que para poder conseguirlo, creo que lo tendríamos que hacer es difundir más la importancia, como estamos haciendo ahora por ejemplo en este webinar, y necesitamos más ejemplos eh, como los ejemplos que veíamos en el webinar de la EATEL, no, donde los investigadores explicaban los beneficios de sus propias prácticas. Espero que las que yo he explicado, no, en este webinar sobre las prácticas que llevamos a cabo en mi propio grupo de investigación también puedan ayudar a otros investigadores y a que se varíe no, expandiendo la adopción del Open Science en nuestra disciplina y después que pueda haber incentivos, más incentivos por el sistema como reconocimiento, como mérito, el hecho de que los investigadores se involucren en prácticas de ciencia abierta, exigencia en proyectos, exigencia en publicaciones, y que haya mas formación eh, como decías.

J: *Bien, bien. Probablemente también desde el nivel inicial de la formación a la investigación en los doctorados e introduciéndolo lentamente también a nivel de máster para que se vaya convirtiendo en una práctica difundida entre la comunidad educativa. Bueno me quedo con esta imagen que realmente me parece potente porque el grupo, se ve que es un grupo, digamos, de trabajo amplio, joven, y que por la diversidad de líneas justamente la interdisciplinariedad se ve con mucha claridad desde el ámbito de la computer science, formación informática detrás de los procesos de los ambientes digitales y de interacción junto a los aspectos pedagógicos, humanísticos y éticos de la educación, que es la riqueza y la complejidad de nuestro dominio.*

D: Muchas gracias por la oportunidad.

J: *Bueno, gracias.*

Translation Spanish-English

J: Ok, so let's start with this webinar even though it's a quite critical and unusual moment. We're going to focus on the question of Open Data within the research on educative technologies and online education, considering it a challenge and a practice that we need to carry out. To tackle this question we have invited professor Davinia Hernandez Leo from the University of Pompeu Fabra and member of the TIDE Research Group on Interactive and Distributed Technologies for Education.

This webinar is brought about within a working context suggested by the research group edu I@b and it is constituted of a series of meetings on data cultures in Higher Education, which can be considered an emerging question impacting on teacher professionalism. Data cultures in HE are also becoming a challenge connected to the social justice and stability within the so called HEIs, the Higher Education institutions.

We have wondered whether we might consider data richness with enthusiasm as an opportunity or, instead, faced with the various critical questions that have been raised, the systematic recollection of data by digital systems could become a problem. Therefore, faced with a series of questions, whether collecting or not the data during the research processes in educative technologies and in technology-mediated environments becomes a central theme, which has a positive connotation but we also need to discover both the technical and the ethical difficulties that might be generated in this sense.

We are going to carry this meeting through a series of questions and so, let's start with our first question, Davinia. Why did you argue that Open Science is a transversal facet within a responsible research on technology-mediated learning and educational processes?

D: Thank you for the introduction Juliana, and I am very glad to partaking in this webinar which deals with such an important thematic that is, as I said, transversal to the different lines of research and working environments that link the theme of education with that of technology. Actually, in my opinion, the question of the Open Science practices is transversal to all the research areas and it is also a question of responsible research practices. Being able to carry through with these Open Science practices benefits the research, hence all the research community and society. Why is this so? Why are these research practices responsible? Because Open Science practices have to do with the fact that we, researchers, make our updates and our research themes more transparent, more frequently published and with an open access to all society, and to other researchers involved in a research process. And thus, a more transparent sharing of our progress in research and an open sharing of our progress in research. So much so that what we share in a more frequent and open form becomes also more documented than without adopting these Open Science practices. More documented so that it is possible in the reproduction processes of knowledge generation which we have carried out, hence more opportunities to verify our research. Making our results more open, more frequently shared and more documented so that they can be reproduced and verified, leads also to a major credibility in our job as researchers and also facilitates the progress of the collective generation of knowledge. Therefore, these research practices are responsible and widely benefit our research for society and the rest of the scientific community. This doesn't mean that the researcher doesn't benefit from these practices. Indeed, it's true that Open Science practices are onerous for both an individual researcher and a research group, it takes time and knowledge on these Open Science practices, but it is also true that this kind of practices can make the researcher's work more impacting in terms of the results for society and the scientific community.

As the vice-president of the European Association for the Technology Enhanced Learning (EATEL), but also as a researcher, I, with my colleagues, have seen that we should raise consciousness about the necessity of using Open Science practices in our research areas since it is a transversal aspect

of responsible work. So much so that you can find in our website of EATEL a webinar in which representatives of the Technology Enhanced Learning area inform us on the vantages, the challenges, the recommendations, and the best Open Science practices in our field which we can familiarise with and also use in our own research groups in the different centres and universities.

J: Ok, so this is extremely interesting precisely because what is currently missing, and here we move to the next question, is the raising of awareness about these practices that you were commenting early on. However, in this sense, having familiarised with your work and your publications, I believe that TIDE is very much advanced. I think that it is one of the most advanced groups certainly on a national level and probably also on a European level (considering your role on a European level Davinia) in this process of science opening and in the Technology Enhanced Learning area. For this reason you'll probably be able to comment on the perks and flaws of the field. So, could you comment on that?

D: The practices that with my group we are carrying through with in order to continue the Open Science movement are varied. On the one hand, we have the Open Science practices involved in an open access through which we publish, whenever possible, in open access journals or in the university repository following a Green Open Access' approximation. On the other, on a code level of the the technologic tools that we develop, we try to publish on Github as open code, and also we publish our data as open data on Zenodo's platform with metadata, including not only quantitative data but also qualitative data, context descriptions, complementary material which allow a more detailed documentation within the research works that we report on the research articles we publish. Therefore, many of our research articles are associated with an open data access on Zenodo, where we make available everything mentioned before. In fact, within Zenodo we handle a concrete community which is called *Educational Data Analytics Community* through which the researchers of our area who share open data on Zenodo are also able to associate these data with the community. And thus, the platform becomes more accessible to our research community that is able to find there open datasets.

Moreover, we try to carry through with meta-research or what we call "research on research processes in our field". As a matter of fact, in the case of the Open Science practices, we have attempted to reproduce studies which have already been carried out in order to see what are the difficulties to reproduce them, and at the same time we figured out how others could reproduce our own procedures. In this regard, we have published an article in the *Journal Computer Assisted Learning* in 2018 which is entitled "Seeking reproducibility: Assessing a multimodal study of the testing effect". Here, the interesting fact is that we explain how we carry out the research on the reproducibility through the reproduction of a previous study, what are the difficulties that we face when reproducing it, and which are the steps to follow in order to allow a future reproduction, including aspects of Open Science practices such as the previous register of the experimental design.

Another thing I would like to point at is the fact that in our group we try to share our educational technology tools with open code, whenever possible, and also we make them available in the Cloud, so we make them available for open use. When I say 'open code' I mean open use, being an approximation of the SAS, the Software Assist Service Cloud, that is to say the software for open use so that the final users, but also other researchers, have access to the tools which we make evaluations on or to design processes of our own tools in our articles so that they themselves are able to work on it in an easier way or to reproduce our studies or similar studies which make use of these tools. This means that our tools not only are described in the articles but also are available for use through Internet. The most significant example of these tools is the *Integrated Learning Design Environment* which is a community, or a platform for professors in the form of a community, where learning designs could be shared. In this platform we integrate authoring tools which we have been developing in different projects with different lines of research, and where we also show the different

approximations that we have for learning data, design data and community data, and how professors can support design and learning processes of other professors. If anyone among those of you who are watching this webinar is interested in knowing more about what we do, that is to say the integration and sharing in an open form of our research results from the point of view of the tools or features that support communities and professors within technologic environments, you can go to the *Integrated Learning Design Environment* platform and see everything there. You'll see how the different aspects of technology and how the different tools have a link with the research articles that we publish and they are available for open use.

J: Fantastic.

D: You were asking me before about the difficulties, right? So, in terms of the difficulties, one of the main ones for us and that requires a lot of time for doing it well is the ethics. Of course, in the research on education and technology we involve human beings, different branches of race and age, we often deal with minors, and we collect personal data. You were asking at the beginning whether this is positive or not. Well, yes, because it allows us to investigate on the impact of our studies on our research. On the other hand, it brings about risks for the participants whose data we are collecting. Especially in the case of Open Science, and especially because we are making these data public, we are going to share them, we are going to analyse them and present them on an aggregated level in our articles. Therefore, the ethical aspects bring about this difficulty in our research area which also has another feature that makes the philosophy of sharing data and the reproducibility processes difficult, that is to say the question of the context dependency. In education, when we carry out studies within real scenarios, we know that the context has a great impact on our results. So, if we want that the shared data be analysed in a correct way, we have to consider the above-mentioned context. And thus, we have to document properly the context in which these data have been collected.

J: The difficulty that often has to do with the identification of the context is precisely the question of anonymisation, that is to say keeping key-aspects anonymised. And here we have what you, Davinia, were commenting before on the functional environment, on the ILDE, where there is an integration between the design tools, which are used by the professor in the practice connected to the research, and the publications. In the latest publications, well there are even more recent ones as I could see from your group, it's extremely interesting to observe how the learning and teaching practices are approached as a task that, using technologies and being a design task, has also a systematic focus which allows a systematic recollection of data. In this sense, it's also very much interesting how the educational practices and the educational research are put in relation. I believe that ILDE is an environment where, let's say, there are various extremely useful tools, even though the average professor is somewhat reluctant to think about himself as something more than an artisan with regard to the way in which he/she works in the classroom and the difficulties which he works with.

Well, you were already elaborating on the ethical aspects of collecting data of users who find themselves in a teaching or learning situation, and on the reasons behind the data collection of both students and professors. Since I've been following your work, I know that you have also lines of research on the ethical aspects of the informed consent through which users are asked for their data. If you want to add something else about this, here you can do it in reference with your work.

D: Well, in order to carry out our research in the most ethical way possible, eventually we decided to incorporate ethics as a transversal line as it happened with Open Science. It became a working line on which we could investigate as an object of research but from a humble viewpoint, because

in our field there are many subjects on which one could make research in ethical terms and there are people that devote their studies to the ethical field only. In this way they could help researchers like us, who attempt to collect data and make them available for open use. So, one of the things that we look at are the procedures of collecting informed consent, which are procedures globally accepted and which we follow in our research community through an open form, and we do that because we know that data sharing can also have risks, something that is made evident not only by the ethical procedures but also by the law. What we have noticed is that these procedures of informed consent or with these procedures we could be led to a relaxed attitude, especially right now that the advances in the learning technologies are increasing the risks which our participants in our research are exposed to. For instance, the researches that have to do with the recollection of physiological data are very close to almost a medical research, and we are collecting increasing amounts of data coming from multiple sources thanks to technology. Also, the advances in technology allow us to make a better aggregation in the analysis of these data and with automatic learning techniques we are able to extract more knowledge than before these data. This is positive because it allows us to make progresses with our research results, but, at the same time, the potential risks with respect to our participant's data become higher. Also another issue is that with the use of technologies in education we can collect data which go beyond the classroom data or the informal learning data, etc. Therefore, not only there are more and more data available but also we can understand them in a better way. With the use of computers we noticed that it was important to work more in detail, on the other hand we observed that research in other more sensible fields to which, as I said, we are getting closer and closer, has been exploring to what extent the participants in the research have previous sufficient knowledge on the risks of sharing data, so that they could understand the informed consent forms in a proper way. Sometimes, in fact, they do not have the previous sufficient knowledge on these risks or they have biased knowledge or they have misinterpreted it. So, what we have noticed is the fact that other studies, in the clinical field for instance, claim that the knowledge on the risks have an actual impact on the users' freedom and that they have to be sufficiently informed to take a decision on whether they are going to participate or not on the data sharing. This is what we have done in one of our articles, in a research, in order to understand what is the previous knowledge of the participants in a research. What we have observed is that there are deficits in the conceptualisation of what are the risks in the data sharing on the part of both professors and students. So, of course, this has implications on how we design the consent forms, how we communicate the research, which we are going to carry out, to the participants so that they can be sufficiently informed to take a decision on their participation.

J: Excellent. In this regard, are you also producing for instance informative material which could help for example the same professors if they want to carry out a work or study process of teaching research, or which could help other researchers or PhD students? Are you thinking about it as a possible future orientation for the formation, at least for the researchers' formation?

D: Yes, well, what we have done, which we have also explained in the article, is a framework of communication with the elements that we have to bear in mind when the vantages and risks of a research are explained. So, what are the risks of sharing these data, why are they shared and what are the vantages. It is not only the negative view of the fact that the risks are still partly unknown so it is not completely possible to take an informed decision, but also, sometimes, what occurs is that we don't explain the benefits properly and thus we don't facilitate the informed decision towards a positive attitude, so this is in both directions. Therefore, what we suggest in this article as conclusion and it is also an element that we include in our own research design of this article is a framework of communication on the risks and the vantages of participating in a research that has to do with data sharing. So, we have produced some materials, some videos, which we use ourselves in this research, and so these videos could inspire other researchers in our field, the framework of

communication could help them in the creation of their own materials to use, within sessions where the informed consent is gathered.

J: Fantastic. Excellent. Well, this is a question which probably doesn't have the solid basis that had in your research work in the TIDE, but which has to do with a perception of future scenarios. Your answers were actually going in that direction, because in producing resources, environments and meta-research, there is a tendency to generate material, either expert or informative, to orient the practices. The perception is that the movement towards Open Science in the education field and in that of educational technologies is still slow, and that there are still many urges. What is, let's say, your perception of this difficulty and at the same time as an opportunity in this sense?

D: Ehm, what I see is that we are having some difficulties in our field, I was saying early on that there are complexities due to the very nature of the domain, there are many factors that are involved, many of them are related to the context. Then, there are all the ethical themes, you know, and it is likely that, due to this fact, we are further from being able to follow the Open Science practices than other disciplines, although I wouldn't say that far, because also the other disciplines have their own problems. But, perhaps, the nature of our domain makes it more difficult. In order to be able to move forward and achieve the goal of working in a more responsible way in our research processes, hence to involve the Open Science practices, I believe that we should spread more the importance, as we are doing right now with this webinar for instance, and we need more examples, I mean, like the examples that we were seeing in the EATEL webinar, where researchers explained the vantages of their own practices. Hopefully, what I explained in this webinar on the practices that we carry through with in our own research group could help other researchers and could help to diversify the practices, spreading the use of Open Science in our discipline. Then, also, I hope that more incentives could be provided from the system as acknowledgement, as merit, the fact that the researchers be involved in Open Science practices, as well as more demand in projects, more demand in publications, and that more formation be implemented, as you said.

J: Well, probably from the initial level of the formation to the research in the PhDs, introducing the Open Science practices slowly also on a master level so that it could become a spread practice throughout the educative community. So I'd like to keep this image in my mind which sounds very powerful to me because the group, it is evident that it is a group which is young and works broadly through the diversity of lines. Precisely, it is very clear that you work interdisciplinary from the field of computer science and computer training behind the processes of the digital environments in interaction between pedagogical, humanistic and ethical aspects of education, which is the asset and complexity of our domain.

D: Many thanks for the opportunity.

J: Thank you to you too.