Author: Andrew Cormack (JANET)

Date: June 2012

FEDERATED ACCESS AND THE LAW

1. Introduction

Federated Access Management involves a complex exchange of information between people and organisations who may not have any direct relationship, often taking place across international borders (or even involving entities whose geographical location is unclear), and often including types of identifiers that were not contemplated in 1995. It is therefore inevitable that it will present significant challenges of interpretation and application of data protection laws that have their origin in the European Directive (95/46/EC) of that year.

Having educational organisations act as identity and service providers for their members, as will be required for e-research, raises new legal issues because, unlike other sectors, the individual's relationship to the organisation is not just that of customer of an access management service. Any legal framework for access management in research and education must take into account the existing relationships and contracts associated with employment, education and the provision of services.

This chapter looks at the legal issues that have been identified by research and education federations in Europe when trying to implement federated access management, nationally, internationally and across continents. The final section then considers which of these problems might be resolved by the proposed *European Data Regulation* (based on the draft published in January 2012) and which will need to be addressed by other means.

2. Universities/Colleges as Identity Providers

Federated access management in the research and education sector raises new issues because of the relationship between the individual and the organisation that provides their digital identity. In the e-commerce and e-government sectors the individual will generally be a customer of their identity provider – for example, banks, telephone providers or social network services might all act as

Author: Andrew Cormack (JANET)

Date: June 2012

identity providers. The UK Government's draft Identity and Privacy Principles[1] emphasise this by their Principle of Multiplicity: "I can use and choose as many different identifiers or identity providers as I want to". In research and education, by contrast, each individual will normally have a stronger legal and personal relationship – that of pupil, student or employee – with one particular identity provider.

Service providers benefit from this stronger relationship because attributes are likely to be derived from internal systems that the identity provider itself relies on; indeed the law may require the provider to verify some attributes to a high level of assurance. Universities and colleges also have more power to deal with any misuse of services by their members than an identity provider who has a purely commercial relationship with the individual. However the relationship also means that in some cases the individual will not be free to make decisions on the processing of their attribute information: if use of a particular service is required for employment or study then the individual has little choice but to release the required information to that service.

There are also likely to be more consequences if an individual wishes to change identity provider than in an open commercial market. This issue is particularly clear when a service is provided to individuals under a licence negotiated and paid for by their educational institution, as is normal for most on-line publications or where the institution has outsourced a service. In this case only one identity provider (functioning as an attribute provider) can assert that the individual is authorised to access the service; an individual who moves to a different identity provider will lose access to that service.

Content licences also highlight the range of pre-existing contractual arrangements that federated access management may need to support. Where an organisation has paid for a site licence, it will be the organisation that has a contract with the service and there is unlikely to be any legal relationship between the individual and the service they use. Conversely if the individual wishes to use their organisation as identity provider to access a professional collaboration service then (as for a commercial identity provider) the contract is likely to be between the individual and the service and there will be no legal relationship between the identity provider and the service provider.

3. Legal Issues

Most of the information passed through federated access management system comprises assertions about the current user of a particular web browser or other client program. Assertions may range from "the user is subject to our Acceptable Use Policy" (as in the eduroam federation[2]) or "the user is a student" to "the user's name, e-mail address and student number are as follows". At least some of these assertions will constitute personal data, so the main law to be considered is the European *Data Protection Directive* (95/46/EC) and its various national transpositions. In a later section the likely effect of the draft *Data Protection Regulation* (published in January 2012) will be considered.

[1] http://digital.cabinetoffice.gov.uk/2012/04/24/identityand-privacy-principles/

[2] http://www.eduroam.org/

Author: Andrew Cormack (JANET)

Date: June 2012

International federated access management raises issues of how to interpret and apply data protection law in the following areas, which are considered in turn in the following sections:

- Legal status of different identifiers and other information exchanged within the access management system;
- Relationships between the parties and the legal basis for each of their processing of personal data;
- International exchanges of access management information with identity and service providers in other continents;
- Appropriate legal framework for group-based authorisation (e.g. within research projects).

In many of these areas there are also significant additional problems resulting from the lack of harmonisation between the implementations of the Data Protection Directive by different Member States.

3.1 Status of Identifiers

Access management systems can use a wide range of information when deciding whether or not to grant access to a service. Among others, a typical educational identity provider will normally be able to provide the status of the currently logged in user (staff/student/etc.); an opaque identifier that only the identity provider can associate with the real user (normally unique to that service); or a unique name for the user (often constructed from the user's login name and their organisation's DNS name). To protect privacy, service providers should only ask for only the information that is necessary for the provision of the service and identity providers should only release that information. For example a static resource licensed to any current member of an educational organisation should only need confirmation of the current user's membership status.

Since the identity provider knows the real identity of the person who is logged in, all these facts about that person will constitute personal data while held by the identity provider. However it is very unclear, and member state laws and interpretations may well differ, whether the first two examples constitute personal data when processed by the service provider. Membership status does not allow the service even to distinguish one user from another; opaque identifiers allow users to be distinguished one from another but should be assigned in ways that make it impossible for the service provider to associate them with a real-world individual or a user of another service. Many federations have additional contractual provisions prohibiting service providers from attempting to make links to identified individuals. Despite these precautions, and the analogous example of clinical trial identifiers in Example 13 of the Article 29 Working Party Opinion 4/2007 on Personal Data,[3] many member states seem to consider that services using only these opaque identifiers are nonetheless subject to personal data regulation.

This lack of clarity and the resulting variation in Member State interpretations make it difficult to exchange attributes between countries. If the same attribute is considered personal data in one

[3] http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp136_en.pdf

Author: Andrew Cormack (JANET)

Date: June 2012

country but not in another it is unclear whether the attribute can lawfully be transferred between them. Even within a country there may be significant problems: opaque identifiers are specifically designed to make it impossible to identify the real-world individual, yet classifying them as personal data appears to impose on data controllers a duty to do just that in order to satisfy subject access requests and breach notification requirements.

For service providers in particular this legal uncertainty creates an unknown regulatory risk, which can discourage use of these attributes. Although federated access management could offer novel attributes that are both privacy-protecting for the user and allow the service provider to accurately express their access rules, this legal uncertainty around them may well result in both service providers and identity providers preferring to use familiar attributes, such as name and e-mail address, that offer legal certainty even though these are neither privacy-protecting nor suitable as unique identifiers (personal names are not unique and a single individual may well have multiple e-mail addresses). Paradoxically a law that ought to promote privacy-enhancing identifiers in fact makes them less likely to be used.

3.2 Relationships/Basis for Processing

The relationships between the individual, their identity provider and the services they wish to access have implications both for the status of the identity and service provider under data protection law, and for the legal justification under which any processing of personal data takes place.

3.2.1 Relationships between the parties

Much of the discussion of federated access management appears to assume a model where the individual has separate relationships with the identity provider and service provider (each acting as an independent Data Controller in terms of the Data Protection Directive) and is free to instruct them to exchange information based on the individual's consent. In this model there may be no direct agreement between the Identity Provider and Service Provider, or only a temporary one established at the user's request for the duration of their transaction. Such an access management system may therefore be described as mediated by the individual.

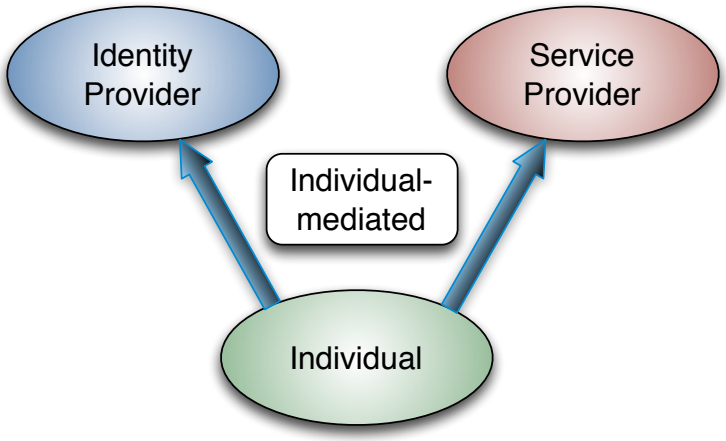

Figure 1 Individual-mediated relationship between Identity Provider and Service Provider

Author: Andrew Cormack (JANET)

Date: June 2012

Federated access management in research and education may occasionally follow this model (even though it would go against the usual legal presumption that an employee cannot give valid consent to an action by their employer) but other arrangements are likely to be more common. These will often involve a direct relationship between the service provider and identity provider (the educational organisation), with a corresponding reduction in the importance of the relationship between the service provider and the individual. Although it will still be the individual that initiates the exchange of access management information, the relationships on which that exchange depends are mediated by the organisation.

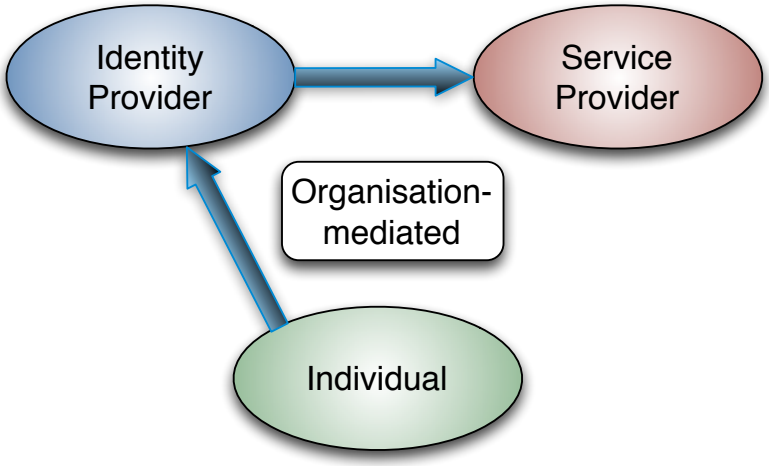

Figure 2 Organisation-mediated relationship between Identity Provider and Service Provider

The most extreme example of an organisation-mediated system is where the organisation has outsourced a core service such as e-mail or a research repository, which needs to be incorporated into the organisation's own authentication and administrative systems. This is likely to involve exchanging detailed information about individuals, for example the departments or groups of which they are members, and is likely to be covered by a detailed outsourcing contract designating the identity provider as Data Controller with the outsourced service as Data Processor on its behalf. Individual users will have an agreement with the identity provider – as either employees or students of the university or college – but will not have an individual agreement with the service provider.

Where an education organisation has obtained a site licence for its members to have access to a content service, the licence agreement is unlikely to contain such detailed instructions on processing of information, and the identity provider and service provider are more likely to act as joint Data Controllers. Services offering content of particularly high value, or that is subject to ethical controls, may wish to also have a direct agreement with individual users but this is unusual.

Under European law both the Data Controller/Data Processor and Joint Data Controller relationships normally require a written agreement between the parties. For outsourced services and site licenses this should not be a problem since these arrangements will already be expressed in the form of a long-term contract between the organisations. Adding clauses to this contract to cover

Author: Andrew Cormack (JANET)

Date: June 2012

data protection issues is unlikely to require significant additional effort. Furthermore outsourcing and content publishing services are generally provided by a few large organisations, and each education organisation is likely to choose only a small number of them to provide to all its members. The number of contracts each education organisation has to negotiate and sign is therefore likely to be small.

By contrast, services supporting e-Research are often more specialised and there are many more of them. Many are specific to a single subject so have a sparse distribution of users: researchers on a particular topic come from a large number of different educational organisations, but there are only a small number of researchers from each one (for example the CLARIN linguistics group expects to have around 25 service providers and 30000 users at 176 different universities). Individual researchers or research groups, rather than organisational management, are likely to choose which service(s) they use. For convenience, security and privacy these researchers should still use their organisation's central identity provider system.

At technical, organisational and legal levels this poses a challenge for the organisation-mediated model presented above. Organisations are unlikely to want to sign contracts with hundreds of e-Research services each of which benefits only a few of their members, while small services will find it impossible to have discussions on either legal agreements or technical configurations with all the organisations from which their users come. The existing relationships between individuals and services seem to have more in common with the user-mediated model, so this may be a more suitable approach, though it is questionable whether a researcher can give free consent to release and processing of personal information in a situation where her job may depend on having access to a particular service. If it is not possible to find a way to satisfy the Data Protection Directive by building on the existing relationships between the parties (in particular those between individuals and the services they wish to access) then an approach based on common declarations or agreements, perhaps with a trusted third party, might be investigated. Any system that involves creating relationships between parties that do not currently have them is likely to take a long time to establish.

2.2.2 Justification for Processing

The relationship between the parties is closely linked to another question: which justification for the processing of personal data does each use? The available justifications are contained within Article 7 of the Data Protection Directive (95/46/EC). As suggested above, it seems that most discussion of federated access management assumes that the individual gives their free and informed consent to all processing by both the identity provider and service provider (Article 7(a)), but where the identity provider may be the individual's employer and accessing the service a requirement of their job this seems questionable at least.

It has been suggested that processing might be justified on the grounds that is necessary for a contract between the individual and the data controller (Article 7(b)). However, as discussed in the previous section, there will often be no contract between the individual and the service provider in situations where the service has been obtained by the educational organisation on its members' behalf. Even for the identity provider this justification seems open to challenge as it will be hard to

Author: Andrew Cormack (JANET)

Date: June 2012

tell which resources are "necessary" for a particular user's employment, research or study at a particular time.

Instead the most appropriate justification seems to be that the processing is necessary in the legitimate interests of each data controller (Article 7(f)). For the identity provider the legitimate interest is to support its members in obtaining secure access to services they wish to use; for the service provider the legitimate interest is to provide the service that has been requested of it by the individual. This enables both identity provider and service provider to process personal data without having to determine whether the particular user is in a position to grant consent or whether the access is necessary for their employment/study. This justification also provides an additional requirement for both service provider and identity provider to protect the individual's privacy – the service provider can only request from the identity provider those attributes that are actually **necessary** to provide the service, while the identity provider must also check that the release of the attributes is not over-ridden by the individual's fundamental rights. Where a service can make use of additional, optional, attributes then it can either ask the user to provide them directly, or else to grant the identity provider permission (now based on **consent**) to release the additional attributes. A number of interfaces have been developed that can present this information in user-friendly form, typically listing the necessary attributes that must be released for the service to function and allowing the user to approve additional, optional attributes that may enhance the service that can be provided.

2.2.3 Lack of Harmonisation

Unfortunately both justifications for processing and relationships between organisations exchanging data have been implemented very differently in different Member States. Some countries appear not to have transposed the "legitimate interests" justification at all, while others have applied additional restrictions[4] that prevent it being used in federated access management. It is not clear that all countries recognise Data Processor status, and the formal requirements for Data Controllers seem to vary, with some countries expecting that exchanges of personal data will be covered by a contract while others do not. This level of divergence makes it very hard to find a legal framework that will work in all Member States; without such a common framework it seems inevitable that the establishment of international federated access management will be hindered by the different laws and expectations in different countries. Consistent implementation and interpretation of the legal requirements is essential

3.3 International Transfers

Many applications of federated access management in education and research involve parties outside the European Economic Area (EEA). Educational resources or teaching may be provided by publishers or universities in other countries, while research collaborations often include both researchers and instruments in other continents. Since these overseas participants are likely to play the same roles in education and research as their European peers it is highly desirable to include them within a single legal framework and agreement, rather than have to maintain two (or more) different legal arrangements among partners who are otherwise treated alike.

[4] http://curia.europa.eu/juris/document/document.jsf?docid=115205&doclang=en&mode=&part=1

Author: Andrew Cormack (JANET)

Date: June 2012

Unfortunately the options provided by data protection law to achieve this are very limited. Under the current Directive the only justification that can be used for transfers of personal data both within the EEA and overseas is the consent of the individual. However as discussed above, it seems questionable whether consent is appropriate for choices that an individual may be compelled to make to continue their employment or study. Since many current research partners are in the USA, the Safe Harbor agreement might be an option for these, but it cannot be used US by universities as they are not covered by the relevant regulators.

Article 26(2) of the Directive permits transfers of personal data outside the EEA where the data controller "adduces adequate safeguards with respect to the protection of the privacy and fundamental rights and freedoms of individuals and as regards the exercise of the corresponding rights". Since the "legitimate interests" justification that is being proposed for use within the EEA already requires both identify provider and service provider to ensure protection of the individuals' fundamental rights and freedoms, it appears that the required safeguards may already be in place. Although "legitimate interests" is not currently listed in Article 26(1) as permitting exports, it seems that it might, under Article 26(2), offer a single legal framework that could cover transfers of personal data both within and outside the EEA while ensuring adequate protection of those personal data.

Unfortunately there is a very wide divergence between Member States in the implementation of Article 26(2) provision. The UK Information Commissioner encourages Data Controllers to base exports on their own assessment of risk, while other countries require that all such exports receive prior approval from the Data Protection Authority. Given the number of identity providers and service providers involved in e-Research, it seems unlikely that Data Protection Authorities would want to receive requests to authorise every such transfer. As with the relationship between identity providers and service providers above, a common approach that scales effectively to large numbers of relationships is essential.

3.4 Groups

Access management for teaching and learning generally involves three parties – the individual, their identity provider and their service provider – but research often adds a fourth party, the project, to further complicate the legal and technical arrangements.

In research it is common for access to resources such as computers, experiments and instruments to be granted to a particular research project, with the project then determining how that access will be allocated among individual project members. Authentication and Authorisation are thereby split: project members are authenticated by their home institutions but authorisation is provided by the project (often generalised to the term "virtual organisation" or VO). Virtual organisations are generally "hosted" by a university or research organisation, for example the home organisation of the principal investigator or the organisation providing the virtual organisation infrastructure.

Although there are a number of different technical approaches to VO authorisation, the processes and relationship involved in access management are generally similar:

1) When he joins the project, an individual authenticates to his home organisation and uses this authentication to register with the virtual organisation (this is likely to involve an out-of-

Author: Andrew Cormack (JANET)

Date: June 2012

band verification that the on-line identity corresponds to the real-world individual known to be part of the virtual organisation). Note that each virtual organisation is likely to include members from multiple home organisations.

2) When he wishes to access a resource, the individual authenticates to his own home organisation and attempts to access the service; since the service does not know which individual users have been authorised to access it, the service must request confirmation from the virtual organisation that the individual is currently authorised to access the service. In technical terms, the virtual organisation is sometimes considered to be an attribute authority, providing at least the attribute "current member of this project". Membership of virtual organisations, and allocation of resources within them, may change frequently so the service should not attempt to maintain this information locally.

It is therefore likely that the home organisations (as identity provider) and the virtual organisation (as attribute authority) will both be processing personal data about an individual known to them; depending on the information disclosed and whether the Member States concerned classify that as personal data, the service provider may also be a data controller.

The existing relationships between these different organisations and the individual are likely to be a hybrid of the user-mediated and organisation-mediated models considered above. The individual is likely to have strong relationships with their home organisation and the virtual organisation; the virtual organisation is likely to have a strong relationship with the service. Thus the relationship between the home organisation and the virtual organisation is likely to be mediated through the individual, but the relationship between the individual and the service is likely to be mediated through the virtual organisation.

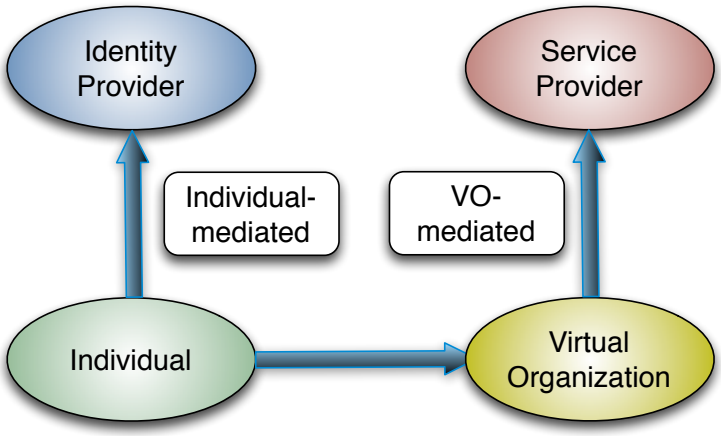

Figure 3 Relationships for VO-authorised Service Provider

Although some virtual organisations are long-term projects with significant technical and legal resources, there will also be many that are small groups of researchers who may, for example, decide at a conference to collaborate on a single research paper. Such *ad hoc* groups could obtain great benefit from a general-purpose e-Research infrastructure, however any requirement for

Author: Andrew Cormack (JANET)

Date: June 2012

heavyweight technical or legal arrangements will be impossible for them to satisfy. The legal framework for a successful e-Research infrastructure must therefore be based almost entirely on the existing relationships between individuals, organisations and services to avoid creating barriers to entry that these groups will find impossible to meet.

4. Summary of Issues and Recommendations

The discussion above highlights four areas where the law currently creates barriers to the use of federated access management for international e-Research. These areas are summarised, with recommendations for reducing the barriers (including any effect from the proposed Data Protection Regulation), in the following sections

4.1 Status of Identifiers

The current unclear status of opaque identifiers, with different national interpretations and apparently contradictory legal obligations, creates a significant barrier to their use by service providers.

The proposal that in future data protection law should take the form of a Regulation rather than a Directive ought to increase harmonisation, however the draft text of Recital 24 ("online identifiers ... need not necessarily be considered as personal data in all circumstances") seems certain to perpetuate the current unclear and un-harmonised position. The provision of small service providers, particularly common in e-research, would be made much simpler by an interpretation that followed the Article 29 Working Party's analysis (in example 13 of Opinion 4/2007[5]) that "serial numbers attributed randomly" with "all other [legal, technical and organisational] measures ... taken to prevent the data subjects being identified" might be treated as non-personal data, thus significantly reducing the legal burden on services designed only to use such identifiers.

Recommendation: A clear statement on the legal status of processing opaque identifiers, implemented consistently across Member States, is essential to support the use of these privacy-protecting identifiers in federated access management. This statement should offer the possibility for service providers to treat suitably protected opaque identifiers as non-personal data or, at least of representing a very low risk to privacy with correspondingly light regulatory requirements.

4.2 Relationships Between Parties

The parties involved in federated access management for education and research have a complex series of pre-existing relationships, mediated in some cases by the individual and in others by their organisation. A successful federated access management system must build on these existing relationships rather than requiring new ones to be created. This is particularly important for e-Research, where the link between service provider and identity provider may involve both user-mediated and (virtual) organisation-mediated relationships.

[5] http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp136_en.pdf

Author: Andrew Cormack (JANET)

Date: June 2012

The proposed Data Protection Regulation makes few changes to the relationship between the parties, though there is a concern that increasing the formal requirements on either the parties or the relationship between them may make it impractical for small service or attribute providers to participate.

Recommendation: Guidance, perhaps in the form of a study, is needed on how adequate protection of personal data can be achieved by incorporating lightweight agreements into existing relationships between researchers, projects, services and home organisations, whether these are user-mediated, organisation-mediated or (for example in VO-based authorisation) both.

4.3 Justification for Processing

The existing relationships (employment, site licences, etc.) between individuals, identity providers and services in education and research cast doubt on whether Consent (Article 7(a)) is the appropriate justification for processing in federated access management (the draft Regulation would make the use of Consent within an employment relationship even more questionable). Instead both identity providers and service providers appear to have a legitimate interest in providing access to the services their members seek to use, which justifies them exchanging information necessary to do so. Consent can then be reserved for information that is not necessary to provide the service, but where the user wishes to enhance the service by providing it. Page 8 of the Article 29 Working Party's Opinion 15/2011 on Consent[6] has an example of a similar (though more complex) hybrid justification. Unfortunately current laws prevent the adoption of this approach across Europe, since Member States have not implemented Article 7(f) of the Data Protection Directive in a consistent way.

The draft Regulation should ensure a harmonised implementation of the Legitimate Interests justification, and also allows it to be used for exporting personal data from the EEA. The Regulation would therefore allow a common legal framework to be used for federated access management covering e-research participants both in Europe and overseas.

Article 7(f) contains an additional test, that the processing not be overridden by the fundamental rights and freedoms of the data subject. Work is already taking place to classify types of service and the attributes those services might reasonably request. Such a classification could provide an additional assurance to identity providers and individuals that the release of these attributes to these service providers did not harm their fundamental rights.

Recommendation: harmonisation of the Legitimate Interests justification, as provided by the draft Data Protection Regulation, is essential for the adoption of a consistent legal framework for federated access management in e-Research.

[6] http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2011/wp187_en.pdf

Author: Andrew Cormack (JANET)

Date: June 2012

Recommendation: the current work on service provider classifications should be considered as providing an assurance that fundamental rights will not be over-ridden by the exchange of attributes necessary for federated access management.

4.4 International Research

Many e-research activities include both researchers and services located outside the European Economic Area (EEA), however the export provisions in the current Data Protection Directive (and their divergent national implementations) do not allow a common legal framework to be used to cover those inside and outside the EEA. Other approaches to exporting personal data may also be unsuitable, for example the Safe Harbor agreement with the USA does not cover universities and public research organisations where research partners are likely to be located.

The draft Data Protection Regulation adds a "Legitimate Interests" justification for exporting personal data from the EEA, so offers the possibility that the hybrid interests/consent justification proposed for federated access management within Europe could also be extended to overseas researchers and services. It has also been suggested that the Regulation might be accompanied by a review of the current arrangements for export of personal data: including provisions suitable for use by overseas universities and public research organisations would further assist collaboration between European and overseas researchers.

Recommendation: the extension of the Legitimate Interests justification to cover international transfers, as proposed by the draft Data Protection Regulation, would permit the use of a common legal framework for all e-research involving European researchers or services.

Recommendation: any future review of the provisions for export of personal data should support the requirements of international e-Research.