1. This is JANET(UK)'s response to the European Commission's Consultation on the legal framework for the fundamental right to protection of personal data[1]. JANET(UK) (www.ja.net) is the operator of the JANET computer network that connects all universities, colleges and research organisations in the UK to each other and to the Internet. We are also the operator of the UK Access Management Federation for Education and Research[2], an Authentication and Authorisation federation that was recently recognised by Educause as an internationally significant technological and organisational advance in providing access to protected services while preserving the user's privacy.

2. In developing and operating the UK Access Management Federation and its international peers we have discovered a number of areas where data protection law fails to encourage privacy-protecting behaviour, or where its implications for new ways of working are unclear and therefore detrimental to their adoption. We believe that clarifications and improvements in three areas of law could make it significantly easier to protect the privacy of Internet users. This response first discusses how federated access management relates to privacy and data protection law, then highlights these three particular areas of difficulty.

Federated Access Management and Data Protection

3. Federated Access Management, which is recognised by the UK's Information Commissioner as a privacy-enhancing technology, [3] allows one organisation, known as the Identity Provider, to authenticate a user's claimed identity and then make trusted statements about that user to other organisations, known as Service Providers. These statements may include information that is clearly personal data, such as the user's name or e-mail address, but can also be statements about what the user is – e.g. staff, student, member of a particular course – that are not personal data. In many cases these statements of the user's attributes are what the Service Provider needs to know whether to grant the user access and, if so, to which services or materials. The Service Provider can therefore make these access control decisions without processing any personal data.

4. However for many services it will also be useful or essential for the user to be able to store information (recent searches, progress through a module etc.) from one visit to the next. In conventional access management this is done by each user having a personal account in their own name with the service provider; in Federated Access Management it is possible instead for the Identity Provider to give the service provider a unique reference number for the user without disclosing the user's identity. In this way the user can ask the service provider to store their information while still protecting their privacy, confident that the Service Provider will be able to recognise them on future visits (and therefore release stored information back to them) without being able to identify them and harm their privacy.

1.

[1] http://ec.europa.eu/justice_home/news/consulting_public/news_consulting_0003_en.htm

[2] http://www.ukfederation.ac.uk/

[3] http://www.ico.gov.uk/upload/documents/library/data_protection/detailed_specialist_guides/privacy_enhancing_technologies_v2.pdf

5. Unfortunately the legal status of these unique reference numbers, known as **pseudonymous identifiers** as they allow a user to be identified while preserving anonymity, is very unclear. This gives both Identity Providers and Service Providers little incentive to use them. This is a particular problem when, as is common in education, federations in Europe wish to work together and with colleagues in other continents to provide students and staff with access to services and journals published in other countries.

6. We believe that three particular areas of the current *Directive 95/46/EC*[4] could be improved to promote the use of privacy protecting technologies: clarifying whether the identity of the person holding information is significant in whether it constitutes personal data; clarifying whether recognition alone is sufficient to constitute personal data or whether identification is needed; and basing regulation on risk assessment rather than a binary compliance test. Each of these will be considered in turn.

7. We note that under the present regime there is considerable variation on these questions between individual Member States' laws and regulatory advice, making services for an international audience particularly hard to establish.

Who can identify?

8. Article 2(a) of the Directive defines personal data as

   "any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity".

   This suggests that a given piece of information will either be personal data or not, and that this status can never change.

9. However some member states, in either legislation or advice, explicitly allow the status of information to change depending on what other information the person holding the information may have access to. For example section 1(1) of the UK *Data Protection Act 1998*[5] defines personal data as

   "data which relate to a living individual who can be identified (a) from those data, or (b) from those data and other information which is in the possession of, or is likely to come into the possession of, the data controller".

   Thus the same information that is personal data in the hands of one data controller under part (b) of this definition may not be personal data in the hands of another data controller who does not have, and is not likely to gain, access to the "other information" necessary to make the identification.

1. _____

[4] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:EN:HTML

[5] http://www.opsi.gov.uk/acts/acts1998/ukpga_19980029_en_1

10. This is particularly relevant to pseudonymous identifiers. For the Identity Provider such an identifier is clearly personal data, since they have the original authentication information to link it to the individual user. For a Service Provider who receives a pseudonymous identifier without the authentication information, and who does not attempt to obtain any other linking information, it appears that the pseudonymous identifier may not be personal data under UK law. Discussions with federations in other member states have suggested that at least some of their regulators take the same view. This matches the reality that using pseudonymous identifiers can indeed preserve users' privacy.

11. Indeed the Article 29 Work Party's Opinion 2/2007[6] on the nature of personal data explicitly mentions the use of pseudonymous identifiers for patients in Example 13, concluding that data associated with such identifiers may be non-personal in the hands of pharmaceutical companies, since the doctors who know the identities are bound by confidentiality rules not to disclose them. Elsewhere in the same Opinion, however, the working party seem to take the view that if anyone can make the link between data and a person then everyone must treat that data as personal.

12. The present unclear situation at best gives little incentive to use pseudonymous identifiers, since the Service Provider may turn out to be subject to exactly the same regulatory requirements if, indeed, their regulator concludes that identifiers are personal data. At worst, this uncertainty may actually discourage use of these identifiers: service providers may prefer to use privacy-invading personal identifiers whose regulatory status is clear.

Recognition versus identification

13. The Directive does not define "identify" however the examples of information that might allow someone to be identified all involve linking information to a particular physical person. It therefore appears from the Directive that **identification** in this sense is intended, as opposed to **recognition**, where a series of visits by the same user can be connected, but not linked to the user's physical person.

14. However it is far from clear in advice from regulators whether identification or recognition is intended to be the requirement for information to be classed as personal data. The Article 29 Working Party's Opinion 4/2007 suggests on page 20 that it is significant whether or not "identification is embedded in the purposes and the means of the processing" of pseudonymous identifiers (i.e. that recognition alone is not sufficient) but on page 14 there is an implication that simply collating activity by a single IP address, which would involve recognition but not, in general, identification, does constitute processing of personal data.

15. As in paragraph 12 above, failure to clarify this distinction may well be discouraging the use of privacy-protecting technologies.

Risk reduction

1.

[6] http://ec.europa.eu/justice_home/fsj/privacy/docs/wpdocs/2007/wp136_en.pdf

16. Finally, the Article 29 Working Party Opinion 4/2007 suggests on page 18 that where appropriate measures are taken to reduce the risk of identification "although data protection rules apply ... the application of these rule will justifiably be more flexible". We would strongly support such a flexible approach, however it is not provided by existing legislation. This contains only binary choices – information is either personal data or not, countries outside the EEA either provide equivalent protection or they do not.

17. The report on the Directive commissioned by the UK Information Commissioner from RAND[7] also concluded that:

    "It has unclear objectives and insufficient focus on detriment, risk and practical enforcement; it is seen as bureaucratic, burdensome and too prescriptive; it focuses on "how" organisations should do things, rather than on "what" they should be achieving."

18. We would therefore wish to see a revised Directive concentrating on effective assessment and mitigation of risk to the individual, rather than approval of pre-defined organisations, countries or practices. This would encourage the adoption of new privacy-protecting technologies that clearly reduce risk but raise compliance questions of the types we have highlighted above. It would also allow data protection to remain relevant to a world where large international data flows are commonplace and, because of load-balancing and resilience measures, the geographical location of processing may not even be determined until the user clicks their mouse.

1. ---

[7] http://www.ico.gov.uk/upload/documents/library/data_protection/detailed_specialist_guides/review_of_eu_dp_directive.pdf