YAHOO! RATIONALE
FOR AMENDMENTS TO DRAFT DATA PROTECTION REGULATION
AS RELATE TO PSEUDONYMOUS DATA[1]

I. Why pseudonymous data as a distinct class of personal data?

- In addition to traditional classes of data such as sensitive personal data and personally-identifiable data, information society services have increasingly looked to new forms of data to carry out useful activities that can leverage 'uniqueness' without having to know "identity."
- A legal regime that recognizes the added privacy benefits of this approach to data governance will seek to incentivise data controllers to use privacy-enhancing techniques to remove identification information or identifying attributes from data to yield a pseudonymous – though unique – piece of data that can be used in the place of identifying information.
- Privacy benefits of pseudonymous (or anonymised) data include specialised approaches to security, anonymisation, encryption, hashing, obfuscation, and segregation, which increase the protection for users and presents obstacles to misuse of data in the case of a breach or other unauthorised access.
- Reliance on pseudonymous data also facilitates effective and privacy-sensitive aggregate data analysis and scientific research (including analytics). This is particularly true where raw data may otherwise include sensitive personal information that is tied to an identified individual, thereby requiring consent of that data subject under the Data Protection Framework. Explicit consent is a well-recognised barrier to generating data sets that are large enough for effective aggregate data analysis.

  In the Proposed Data Protection Regulation:

  Draft Recitals (23) and (24) discuss identifiability and recognise that all data need not be considered identifiable at all times; specifically that "[t]he principles of data protection should not apply to data rendered anonymous in such a way that the data subject is no longer identifiable."

  Recital 23 reflects a clear incentive to render data anonymous (an affirmative action by a data controller), but suggests that *no* data protection principles need apply to such anonymised information. While this would provide a powerful corporate incentive to invest in anonymisation – it is unclear that the 'all or nothing' approach to data protection can be well supported in the information age. Instead, evolution of this concept likely points the way to *differentiated protections depending on the class of data implicated*.

  A sensible approach to incentivising use of anonymisation for data without forgoing all commitments to data protection would be to a) apply the highest class of protections for sensitive personal data, b) recognise an alternate set of protections for personally-

[1] For purposes of this paper only, 'pseudonymous' is being used as a general term, to capture policy discussions around pseudonymisation, anonymisation, de-identification, and similar techniques for removing identity. The specific criteria governing this practice will need to be scrutinized in order for Yahoo! to be supportive of any definition here. However, the existing reference to pseudonymisation in German law that is being used as the basis for proposed amendments suggests the use of this term may be most familiar to members of Parliament and the Council at this stage, and will aid general understanding of the principle by non-subject matter experts.

identifiable data that presume the opportunity for control by authenticated users, and c) adjust protections for pseudonymous data that rely on security and anonymisation technologies to remove identifying information, thereby also removing the necessity to authenticate an individual and make that individual responsible for his own privacy protection.

Recommendations:

- Modify Draft Recital 23 to state that *certain* principles of data protection should not apply if data is rendered anonymous, taking account of all the means likely reasonably to be used by the controller.
- Supplement the definition of personal data in the Draft Regulation by *adding* a distinct subset of data considered pseudonymous, which will trigger differential obligations under the Regulation (namely, that certain provisions will be inapplicable to this class of data – *See Section III of this Paper*).

II. Legal basis for processing of pseudonymous data

- Contract (Article 6, paragraph (1)(b)) and explicit consent (Article 6(1)(a)) are two legal bases in the Draft Regulation that are imperfectly suited to serve as the lawful processing of pseudonymous data by a controller.
- Both presume the opportunity for authentication of an identifiable individual to whom a record or piece of data may be assigned. Yet, as we have stated earlier, lack of identifiability to an individual --using means likely reasonable to be used by the controller to do so-- is also an attribute of pseudonymous data.
- Similarly, any retention or recording of consent presumes collection of additional data relating to identity, thereby extinguishing the data's status as *pseudonymous*.
- In light of these realities, it appears that there are two possible approaches to providing a legal basis in the Draft Regulation to support uses of pseudonymous data:
    - The first would be to develop the theory behind the balancing test inherent in the "legitimate interests of the data controller." (Article 6, paragraph (1)(f)).
    - The second would be the recognition of an entirely new legal basis for the processing of pseudonymous data within Article 6, paragraph (1)(g)(new): "*processing is undertaken using data that are not identifiable to a person, or are rendered anonymous in such a way that the data subject is no longer identifiable by the data controller, accounting for all means likely reasonably to be used by the controller to identify the individual*," or similar.

In the Proposed Data Protection Regulation

Article 10 states that "if the data processed by a controller do not permit the controller to identify a natural person, the controller shall not be obliged to acquire additional information in order to identify the data subject for the sole purpose of complying with any provisions of this Regulation." Although this language may be helpful in reinforcing our recognition of the pseudonymous class of data, it does not explicitly acknowledge that there nevertheless must have been a legal basis upon which such data is being processed – one that does not require identification of an individual. The only possible legislative approaches would appear to be use of the "legitimate interests" basis, or the creation of a distinct legal basis for pseudonymous data as described above.

Article 6, paragraph (1)(f) on "legitimate interests pursued by the controller" preserves a legal basis for other circumstances and describes a balancing test that must be conducted – weighing the proposed purpose of processing against interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child.

The areas of highest concern in terms of impact to rights and freedoms of the data subject that require protection of personal data should be those related to adverse uses of identifying information, but greater clarification of this point would be particularly useful in helping to establish objective criteria for application of this principle. It may well be that by removing identification from the equation, a presumption of lawfulness is gained in favour of the data controller's processing, unless certain prohibited uses that may be articulated in the regulation are proposed. For example, it might be that use of pseudonymous data for matters that substantially affect legal rights of an individual data subject ought to fail the balancing test as a matter of law. Additional guidance on what business uses are considered "legitimate" as opposed to those that are "illegitimate" can help bolster this legal basis.

Legal basis is at the heart of legal certainty in application of the draft Regulation, so the proposal to delegate the act to the Commission (Article 6, paragraph 7) to define such conditions seems, at the very least, to be a missed opportunity to provide that legal certainty.

Recommendations:

- Consider a Recital clarifying the spirit of Article 10: that in the context of pseudonymous data, legitimate interests – rather than contract or explicit consent – is the best-adapted legal basis for processing of this class of data.
- Alternately, consider introducing sub Article 6(1)(g), establishing *per se* lawfulness where 'processing is undertaken using data that are not identifiable to a person, or are rendered anonymous in such a way that the data subject is no longer identifiable by the data controller, accounting for all means likely reasonably to be used by the controller to identify the individual.'
- Develop a Recital that creates a presumption of legitimacy in favour of the data controller where data has been properly pseudonymised, while articulating what effects on a data subject's legal interests would override that presumption. These could include significant legal effects on an individual.

III. Data Controller Obligations Flowing from Authentication Waived; Other Protections Preserved

- Recognising the virtue of investments in privacy-enhancing pseudonymisation practices to remove identifying information, and/or to rely on non-identifying unique data, Articles that create certain data subject rights should be waived, as listed below in this Section.
- Such obligations assume that a data controller can authenticate a unique user, so as to provide access to a specific set of account-level data that a data subject could review about himself/herself. However, as has been shown – by definition a pseudonymous identifier would not likely be capable of providing the same level of access and control.
- Such an approach is consistent with and would rationalise the indication already present in Article 10.
- Articles relating to data controller obligations of accountability and security should remain applicable, as listed below at the end of this Section.

In the Proposed Data Protection Regulation

Article 10 tells us that "if the data processed by a controller do not permit the controller to identify a natural person, the controller shall not be obliged to acquire additional information in order to identify the data subject for the sole purpose of complying with any provisions of this Regulation." This language is helpful in reinforcing our recognition of the pseudonymous class of data. It implicitly acknowledges that certain (though not all) proposed Articles could require a data controller of pseudonymous data to acquire identifying information, thereby rendering the data personally-identifiable and forgoing the privacy benefits of pseudonymisation. We suggest that the following Articles, falling principally in Chapter III, and consistent with the spirit of Article 10 that immediately precedes many of them, would only make sense when a user has been authenticated, and as such should be waived from the list of affirmative data controller obligations when pseudonymous – rather than identifying data – is used for data processing[2]:

1. *Article 8: Processing the Personal Data of a Child.* Parental consent requirements presume the opportunity for a data controller to authenticate not only a specific age, the identity of a specific child, but also of its parents.
2. *Article 15: Access.* Authentication would be required for exercise of right.
3. *Article 16: Right to Rectification.* Authentication would be required for exercise of right.
4. *Article 17: Right to be forgotten and Erasure.*[3] Authentication would be required for exercise of right.
5. *Article 18: Portability.*[4] Authentication would be required for exercise of right.
6. *Article 19: Right to Object.* Authentication would be required for exercise of right.
7. *Article 20: Measures based on Profiling.*[5] This section is applicable to measures that either produce legal effects concerning a natural person, or which "significantly affect" a natural person – both of which imply knowledge of "identity," which is incompatible with the status of information as pseudonymous. Presumes data controller knowledge of the identity and status of a child (in Art 20(3)); ability to obtain express consent, which would require authentication as has been discussed (Art 20(2)(c)). Much online activity in this field can be conducted with pseudonymous data, and incentives should be preserved to recognise the value of this approach.
8. *Article 32: Communication of Personal Data Breach to the Data Subject.* Authentication would be required for individualised communication. This approach is already consistent with the statement in Article 32(3) that contemplates a state of pseudonymisation of a sort: "The communication…shall not be required if the controller demonstrates … that it has implemented appropriate technological protection measures, and that those measures were applied to the data concerned by the personal data breach. Such technological protection measures shall render the data unintelligible to any person who is not authorised to access it."
9. *Article 74: Right to Lodge a Complaint with a Supervisory Authority.* Authentication would be required for exercise of right.
10. *Article 76: Right to a Judicial Remedy Against A Controller or Processor.* Authentication would be required for exercise of right.
11. *Article 78: Right to Compensation and Liability.* Authentication would be required for exercise of right.

[2] Note that Yahoo! has distinct positions on certain of these Articles even in the authenticated context, but those concerns are reserved for separate discussions.
[3] See note 2.
[4] *Id.* Note that Article 18 of the proposal already applies only to data processed under contract or consent.
[5] *Id.*

In contrast, the Regulation contains several provisions that do not require an individual to have first authenticated his or her identity. As such, they may also be applicable to the pseudonymous data class. These include:

1. *Article 22: Responsibility of the Controller*
2. *Article 23: Data Protection by Design and by Default*[6]
3. *Article 28: Documentation*[7]
4. *Article 30: Security of Processing*
5. *Article 33: Data Protection Impact Assessment*
6. *Article 35: Designation of a Data Protection Officer*
7. *Articles 40-46: Transfers of Data*

Recommendations:

- Amend Article 10 to enumerate its applicability to the following Articles: 8, 15-20, 32, 74, 76, 78.
- Consider a Recital clarifying the justification for the above exclusions, as flowing from the Principle in Article 10 that data controllers should not seek to re-identify pseudonymous data in order to comply with those above Articles.

[6] *Id.*
[7] *Id.*