Innovative use of data in an industrial context:

The Need for a Balanced Approach to the Draft General Data Protection Regulation

In "Pushing the Boundaries of Minds and Machines", GE explains why the "Industrial Internet" holds the potential of significantly transforming the economy and of bringing substantial economic benefits in terms of faster economic growth and higher living standards. According to estimates, the Industrial Internet could add EUR 2 trillion to Europe's GDP by 2030, by achieving efficiency gains as a result of connected machines and real time data analytics. [1]

This paper discusses the role that personal data plays in the Industrial Internet. In particular, it explains the need – particularly in light of ongoing discussions concerning the draft General Data Protection Regulation -- to consider both the *context* in which personal data is used, and to design proportional responses to the associated *risk* when regulating the use and transfer of such data.

The role of personal data in the Industrial Internet

The Industrial Internet is a growing economic reality that consists in connecting machines-to-machines, and machines-to-people-to-machines, to improve safety, efficiency and the performance of machines. For this purpose, the collection of personal data is limited when compared to the personal data collected in the context of the Consumer Internet. The core function of the Industrial Internet is **not** to collect vast amounts of personal data to create individual profiles for marketing or data brokerage. Rather, at the heart of the Industrial Interest are the vast quantities of data generated by machines, and the advanced algorithms that convert these data into actionable information for use by equipment operators. By remotely aggregating these data and applying unique software-based algorithms, equipment anomalies can be rapidly identified and repaired, and safety, unit performance, efficiency and reliability can be enhanced. In some cases, this may include the collection of personal data for the purpose of a functional machine-to-machine communication or to correct errors in machines that affect the home, transportation, energy, or the provision of healthcare services. For example, limited information about mechanic services or passenger weight may be needed for flight diagnostics, and information concerning patient weight or age may be needed to support proper healthcare diagnostic results and equipment support.

Examples from the transportation, energy and healthcare sectors help to illustrate the use of machine data in an Industrial Internet setting.

In both **transportation** and **energy**, data generated by machines, such as aircraft or locomotive engines, and turbines for power generation, can be used to improve infrastructure efficiency and safety over time. The data concerning machine performance is communicated over the internet to computing facilities, where it is analysed in real-time. A single turbine can generate up to 50 gigabytes of data per day, while an aircraft or locomotive engine can generate thousands of pieces

[1] *Industrial Internet: A European Perspective. Pushing the Boundaries of Minds and Machines", June 2013* http://www.ge.com/europe/downloads/IndustrialInternet_AEuropeanPerspective.pdf

of data instantaneously concerning fuel flow, engine temperature and the effect of environmental conditions. In both of these scenarios, a limited amount of personal information concerning machine operation may be analysed, and it is likely – particularly in transportation – that data will cross national borders along with the underlying equipment. Using these data in an Industrial Internet setting generates important efficiency savings. For instance, an efficiency increase of only 1% in gas-fired power generation installed base alone represents EUR 11 billion in fuel savings in Europe over the next 15 years. The same holds true in aviation. The European commercial airline sector spends about EUR 35 billion per year on jet fuel. Industrial Internet efficiency savings of 1% would represent nearly EUR 6.5 billion in fuel cost savings over 15 years.

The **healthcare** sector also stands to gain in efficiency and cost reduction through the innovative use of data. Analysing healthcare data allows hospitals and other providers to move from episodic care to a more integrated patient-centric approach in healthcare delivery. At the same time, analytics empower research and innovation in Europe to drive better health outcomes, increase productivity and contribute to economic growth. Over EUR 1.2 trillion was spent on healthcare in the EU in 2012. It is estimated that 10% of this expenditure is wasted due to system inefficiency, of which 59% is in clinical and operations inefficiency. This is where the Industrial Internet could yield the greatest benefits: a 1% efficiency increase in these clinical and operations would translate into EUR 11 billion of savings over the next 15 years. These are important savings in light of Europe's ageing population particularly given that public healthcare budgets are under severe pressure.

A patient's hospital stay – even for relatively straightforward care – involves hundreds of people, processes and assets. The medical outcome and cost of that care, as well as the quality of the patient's experience, depend on how efficiently and effectively all of these resources are integrated and managed. The effective use of data optimises the operations of hospitals by improving asset management and optimising patient flows.

The infograph below provides an example where available data is used real-time to improve operations that can have a significant impact on Hospital Acquired Infections (HAI), thereby reducing the number of infections with patients. HAIs are an unintended consequence of care delivered by healthcare organisations. As shown here, a hospital may use a variety of data, including sensor readings when a care provider enters a room, to determine whether care protocols – in this case related to hand washing - are being followed, as well as to research whether care protocols or hospital design should be modified to improve overall care in a hospital or across one or more medical systems. Though the data is initially used to manage current healthcare operations, some data may be needed to perform longer-term research into equipment fixes and improvements and into care protocols.

A hand washing monitoring system uses real-time data via sensors to improve compliance of hygiene protocols. The results: a 30% sustained improvement was realised in the first eight weeks after implementing the monitoring system, and an 80% increase of care protocol compliance: from 700 observed hand washes to 1.8 million tracked hand washes on an annual basis, thereby significantly reducing the number of infections of patients with Hospital Acquired Infections.

The Need to Consider *Context* and *Proportional Responses to Risk* when Regulating Data

Throughout the drafting process, the General Data Protection Regulation has been promoted as a means to lower barriers to innovation and to productive uses of data and promoting a risk-based approach to data protection governance. As EU Justice Commissioner Reding explained when issuing the first draft, the GDPR is intended to "*help build trust in online services . . . while making life easier and less costly for businesses. A strong, clear and uniform legal framework at EU level will help to unleash the potential of the Digital Single Market and foster economic growth, innovation and job creation*."

The hand washing case described above is just one example of the innovative use of data in a healthcare setting that can provide opportunities for growth of innovative technologies across Europe while also improving patient care and hospital efficiency. It is, however, also an example where important and innovative uses of data will only be fostered by a regulatory environment that focuses on the ways in which data will be used and on how organizations address risks of loss or misuse rather than creating more administrative burdens and undue limitations on its use.

While a streamlined regulatory environment holds out the prospect of such benefits to industry and to consumers, the current discussions of the GDPR give rise to concern. Some of the most notable issues with the current drafts under consideration are discussed below.

Consent: Article 6 (Lawfulness of Processing) and Article 7 (Conditions for Consent) of the LIBE draft require that individuals be notified of, and give "explicit consent" to, each specific purpose for which data is processed, and consent would lose its validity once the initial purpose given for the collection of data is no longer necessary. Moreover, Article 7 of the LIBE draft provides that Member States may enact further specifications in these areas, including purpose specification, processing and storage. These proposals contrast with existing data protection law, and with the Council's current proposal, which provide for "unambiguous consent" and protects uses that are compatible with the initial purpose of collection.

These requirements, as well as the prospect of further limitations from each Member State, would be particularly difficult for organisations whose products are used in a complex supply chain, since it would be difficult to ascertain whether proper consent exists to use data for important purposes such as equipment support and avoiding "regression" of problems as new versions of healthcare solutions and other products are released. It would be difficult, at best, for individuals to consent to all potential legitimate uses of data for machine analytics and long-term care given the complexity of the relationships involved with infrastructure, transportation and healthcare systems. However, without such consent, organisations may not be able to use data to address safety and efficiency issues in products or in people's interaction with those products. Thus, in the example above, consent for collection and use of patient data for care may be deemed insufficient to read and analyse hand washing data either for care or for measuring improvement in clinical outcomes.

Privacy Impact Assessments: An important goal of the GDPR was to remove the varying and sometimes inconsistent registration requirements across the Member States, which are burdensome and generally have not proven to materially improve protection of data. However, under the current

drafts, organizations would be required to perform detailed privacy impact assessments and make those assessments available for review by individual Member State DPAs in light of different factors, including in cases where the processing involved new technology. Thus, the most innovative uses of technology may be subject to the highest level of scrutiny. Moreover, since the requirement to perform a privacy impact assessment ("PIA") extends to a range of undefined circumstances beyond those already defined in the law (such as where processing would involve "sensitive data"), there is significant risk that the PIA requirement would be interpreted differently by the Member States, leading to both greater administrative burden and splintered or conflicting regulation of new technologies that otherwise would serve a single market across the EU. In order to avoid the PIA requirement becoming a complex administrative burden and a significant disincentive to innovation, the obligation to perform PIAs should be limited to circumstances involving sensitive data, as defined, and controllers should be provided explicit protection against the disclosure of PIAs that could reveal trade secret information and other intellectual property.

Cross Border Data Flows: In recent months there has been a significant amount of discussion in the context of the Regulation about a perceived need to restrict cross-border transfers in order to better protect the privacy of EU citizens. By way of example, Article 43a of the LIBE draft, which would require DPA approval before records could be disclosed to law enforcement and others outside the EU, would create direct conflicts of law for many organizations operating -- including those headquartered – in the EU. Some involved in the negotiations have suggested going farther in limiting cross-border data flows.

Industrial Internet services are by definition without boundaries, and cross border data flows should be allowed as long as the data is bound to be protected in a manner consistent with EU data protection standards. A focus on managing risk would permit significantly greater innovative use and transfer of data without adding materially to the associated risk. From the perspective of the Industrial Internet, cross border data flows should remain lawful while maintaining adequate protection of personal data.

Damages: The LIBE draft includes the potential for damages up to 5% of a company's global turn-over, and some involved in the negotiations have called for that number to be revised up to 10% (in the case of GE, this could theoretically amount to as high as $15 billion for a single data protection incident). There currently is no provision limiting damages in light of the organization's overall data protection program, the nature of the harm suffered or the intent of the organization. The current proposed framework, and particularly the uncertainty concerning regulatory enforcement for a single incident, would provide a significant disincentive to organizations hoping to conduct business involving innovative technologies in the EU. By clarifying the damages available for incidents, as well as mitigating factors designed to encourage good data practices, the GDPR would promote strong data protection practices rather than discouraging investment in local innovation.

Conclusion

The use of Big Data for the Industrial Internet can provide significant economic and social benefits. However, regulating the use of data for industrial purposes purely in terms of narrowly granted one-to-one consent and permissions for point-to-point transfers would dramatically impede the value and benefit of Big Data in the industrial context, without providing significantly greater protections

for individuals. Similarly, re-asserting burdensome administrative requirements and imposing open-ended damages for privacy harms would create significant disincentives to investment in new technologies in the EU.

Instead, in the context of industrial use of machine data, it would be both more efficient and ultimately more protective to focus on c*ontext* (how data is used and whether the use is reasonably anticipated) and *risk* (protecting data appropriately through security, monitoring and de-identification where possible). This would permit data to be used to improve and ensure the safety of machines, while also requiring that data be kept securely and not be used for unanticipated purposes.