Legal Restraints and Security Requirements on Personal Data and Their Technical Implementation in Clouds

Benno Barnitzke[1], Wolfgang Ziegler[2], George Vafiadis[3], Srijith Nair[4], George Kousiouris[3], Marcelo Corrales[1], Oliver Wäldrich[2], Nikolaus Forgó[1] and Theodora Varvarigou[3]
[1] *Institut für Rechtsinformatik, Leibniz Universität Hannover, Königsworther Platz , 30167 Hannover, Germany*
Tel: (+49) 511 762-8242, Fax: (+49) 511 762-8290, Email: { barnitzke, corrales, forgo}@iri.uni-hannover.de
[2]*Fraunhofer-Institute for Algorithms and Scientific Computing (SCAI), Schloss Birlinghoven, D-53754 Sankt Augustin, Germany*
Tel: +49 2241 14 2258; Fax: +49 2241 14 42258, Email: {Wolfgang.Ziegler, oliver.waeldrich}@scai.fraunhofer.de
[3]*National Technical University of Athens, 9 Iroon Polytechniou Str., Athens, 15773, Greece*
Tel:+302107722546, Email: gvaf@iccs.gr, gkousiou@mail.ntua.gr, dora@telecom.ntua.gr
[4]*BT Innovate & Design, Martlesham Heath, Ipswich IP5 3RE, United Kingdom,*
Tel : +44 1473 606243, Email : srijith.nair@bt.com

Abstract: Cloud computing has emerged as the new trend in the IT industry. However, in order for the adopters of this technology to be legally compliant with regard to the handling of personal data, a series of actions must be undertaken. In this paper, we present the legal requirements that exist inside the EU with regard to transnational data transfer and storage. Based on these legal requirements, we then used the WS-Agreement standard to create corresponding SLAs. Subsequently, we extended existing and well known technologies in order to create a data management framework that is necessary from an Infrastructure Provider point of view, so that the latter can be considered as a trusted entity with regard to data management. As a result, customers and Cloud providers using the presented SLA and data management framework are able to be compliant with the legal requirements stipulated by the Data Protection Directive with regard to transnational data transfers.

1. Introduction

Cloud Computing involves location-independent processing and leads to the global transfer of personal data. On the one hand, the scaling of resources in the cloud enables users to save considerable costs. On the other hand, data protection laws might prevent cloud providers from transferring the data to other countries and additionally requires them to implement data security measures. Cloud providers therefore need smart tools to ensure compliance both with data export and data security requirements and thus extend their business to applications dealing with personal data.

Due to dynamic resource allocations enabling cloud users to scale their applications across different cloud providers, cloud computing entails a highly distributed architecture. This implies more data in transit than in traditional infrastructures [1] and even leads to massive and worldwide transfers of personal data within a cloud. However, international data transfers to a country outside the European Economic Area (EEA) are subject to clear legal restrictions. Users and cloud providers therefore facethe problem of ensuring that personal data is not being transmitted to these countries without further legal safeguards.

We will describe these constraints in more detail and provide legal advice on a technical solution, enabling both users and providers to comply with these rules. This solution could be used by Service Providers (SPs) or Infrastructure Providers (IPs) who would like to offer data protection compliant cloud services within the EU. This way, SPs can offer cloud enabled applications (i.e. CRM software) running on a location aware cloud data management system without taking the risk of illegitimately transferring personal data to third countries.

2. Objectives

This paper aims to highlight the most relevant legal issues of cloud computing on a European level, regarding the transfer of personal data to third countries. It also provides a solution as to how to express these requirements on a technical level in the SLAs and the data management system. We show how these legal requirements are expressed in the SLA and what action is needed in order to include the information about the location of the data processing into the data management system. Furthermore, we address data security issues, particularly which technical measures in the cloud architecture should be implemented to protect personal data against unauthorised processing. A technology framework has been created in order to implement the management of these requirements and additions to existing toolkits. Certain approaches are discussed in order to ensure that the overall goal will be addressed.

The major objectives of the paper are the following:

1. Detailed analysis of the legal requirements based on the current European legal data protection framework. This will aid Cloud providers in understanding what is legally expected from them within the European Economic Area.
2. Contribution to existing specifications (e.g. regarding SLAs) in order to includeterms and conditions that are related to step 1.
3. Proposal of a technical framework that can be used by IPs in order to implement the requirements produced by step 1 and, as a result, be compliant with current data protection legislation.

3. Methodology

Firstly, this paper outlines the legal conditions for data transfers and storage in clouds outside the EEA. We consider legal requirements as part of the Quality of Service (QoS) a Cloud user requires from its Cloud provider. The placement of data is fixed by a binding Service Level Agreement (SLA) with the cloud provider. A specification for describing these legally binding terms is included, in order to externalise resources used by an IP.

Finally, we describe the measures that are needed in order for the infrastructures to comply with the requirements by using existing tools and technologies from the OPTIMIS project (www.optimis-project.eu). In order to implement these, we extended existing toolkits that are widely used and adopted.

4. Analysis of Legal Requirements

a) Data Protection Requirements Regarding Transnational Data Transfers

Within the territorial scope of Directive 95/46/EC (the Data Protection Directive, hereinafter referred to as DPD), no restrictions exist for data transfers to countries within the EEA, since all Member States are deemed to provide an adequate level of protection [2]. The territorial scope of the DPD comprises all 27 EU Member States plus the European Economic Area (EEA), thus including Iceland, Liechtenstein and Norway. As a consequence, data transfers within the territorial scope of the Data Protection Directive are allowed. Furthermore, the European Commission has explicitly stated for specific countries outside the EU (after thorough examination of their national laws) that they also provide an adequate level of protection. These countries are Switzerland, Canada, Argentina, Guernsey, Isle of Man, Israel (since 31 January 2011) and US organisations that take part in the US safe harbour program [3]. While transfers to cloud providers in any of these countries are legitimate, transfers to any other country ('third countries') not mentioned here are prohibited if there are no further legal safeguards implemented by the data controller. The list of countries to which transfers are possible are therefore similar to a 'white list'. Thus, a cloud provider processing personal data must provide the necessary technical measures to prevent any transfer of data to third countries (= countries not listed on the 'white list') if no additional safeguards shall be taken (see Figure 3 for a graphical overview of admissible transfers of personal data). Although the transfer of personal data to these third countries is not allowed, there are additional legal bases (mostly on a contractual level) that provide an adequate level of protection for data exports even if the third country itself does not. According to Art. 26 (4) DPD, the European Commission may decide that certain standard contractual clauses offer sufficient safeguards for such kinds of data transfers. To this end, the Commission has set up a standardised set of clauses which can be used as a legal basis for transfers from each Member State to any third country (Standard Contractual Clauses) [4]. If a controller located within the EU or EEC enters into a contract which includes the EU Standard Contractual Clauses, the controller located outside the EU or EEC is considered to provide an adequate level of protection [5]. Consequently, if SLAs agreed between an end user and a cloud provider apply the (unmodified) EU Standard Contractual Clauses in a legally binding way, the transfer would be lawful. WS-Agreement should therefore offer cloud customers the possibility to include EU Standard Contractual Clauses. Otherwise, the data management system should restrict data placement to countries not included in the 'white list'. In view of these explanations, we derived three legal requirements for the distributed file system used in the OPTIMIS project (Table 1).

Provide information to end users about their data storage locations	**Support country-specific location for placement of personal data and separation from other data**	**Prohibit placement of data in third countries**
Cloud providers should disclose the fact of whether they are operating data centres in third countries or whether they are in a federation with other cloud providers that operate data centres in third countries. The user should be able to constantly track the location of the data processing by means of a monitoring tool.	The file system run in the cloud must be able to support locations, thus be "aware" where the data centres are located in order to enable users or cloud providers to decide where personal data may be transferred to. However, some users may process information not relating to an individual. Data of this kind does not fall under the scope of the DPD. Separation of the two categories of data would significantly reduce the impact on performance and management load for both users and IPs.	Where a user and cloud provider have not agreed upon additional safeguards for data transfers to third countries, the file system must not transfer personal data to data centres located in these countries. Requests to transfer data to locations in third countries must therefore be denied by the Data Management System. However, where users and/or cloud providers have agreed to use Standard Contractual Clauses, the Data Manager should accept the transfer.

Table 1: Legal requirements for the Data Management System in the OPTIMIS cloud.

b) Data Security Requirements Regarding Transfer and Storage of Personal Data

The DPD also contains constraints with regard to the security of processing. As the recent Amazon EC2 outage has clearly shown [6], data loss or even destruction of personal data is a major concern in clouds. According to Art. 17 DPD, the controller must implement appropriate technical and organisational measures to protect personal data against accidental or unlawful destruction or accidental loss.

While “destruction” of personal data represents the complete removal or serious corruption of physical data (i.e. on the hard disk or in main memory) in such a way that their recovery is impossible, “loss” refers to unplanned events such as natural disasters or hardware failures. Consequently, the data in virtual machines should be backed up by replication on different physical machines in different data centre locations on a regular basis. Compliance with this provision is not only a legal requirement, but also in the best business interests of the cloud provider, as it helps to preserve company reputation and customer trust [1]. However, it is not solely the cloud provider’s duty to protect personal data. Maintaining data integrity and availability is also an obligation of the cloud customer, which could mean running services across multiple providers [7]. To increase redundancy in cloud services, IaaS Providers should create a standard for sharing VM instances across clouds to simplify compliance with Art. 17 DPD. With the OPTIMIS toolkit being developed, this will – to some extent – come true, as it facilitates provisioning of services in Federated and Multi-cloud scenarios.

Due to considerable transfer and storage of personal data in clouds, as well as the aspect of multi-tenancy, confidentiality is also key for cloud computing. The DPD requires controllers and processors to protect personal data against unauthorised disclosure, in particular where the processing involves the transmission of data over a network. Hence, we recommend implementing strong encryption whenever an IP moves data within the cloud. Finally, we recommend that data should be encrypted at rest (stored data) with the same diligence as data determined to be transmitted, as they are exposed to the risk of disclosure in the same way. The DPD does not mention technical details of encryption type or strength. This depends on the type and quantity of data to be processed. A security analysis should be made in any case and a policy set in place. Compliance with that policy should be legally agreed on between all partners and should be constantly monitored.

5. Technology Description

As baseline technology for negotiating and creating the SLA selected, WS-Agreement, a standard of the Open Grid Forum (OGF), and WS-Agreement Negotiation are used. Besides the fact that WS-Agreement is a well-established standard, the rationale for using it is its flexibility. WS-Agreement, itself being completely domain agnostic, requires the use of a domain-specific term language to create SLAs for the domain. In OPTIMIS we develop term languages, e.g. for expressing Trust, Risk, Eco-efficiency and Costs (TREC parameters) related to infrastructure services of a provider. In this context, we also define a language to express the requirements with respect to data transfer and security. These terms are then part of a legally binding SLA with the IP ruling the transfer of the data within its data centres. Since it is not possible to figure out whether the Service Description Terms (SDT) and related guarantees were fulfilled or violated without the ability to monitor the state of the system, we also are developing a monitoring instrument for data location tracking. Suitable XML schemas are described in order to fully capture the legal requirements that the infrastructure must take into consideration when accepting new data or enabling federation between different cloud providers.

Given that the above requirements are adequately expressed, what is left for the Cloud providers is to create the framework that will enable them to abide by their commitments agreed to in the SLAs. To this end, we will describe the design of the OPTIMIS project data infrastructure and the tools that are used in order to ensure that the end user’s constraints are met, especially when two or more Cloud providers cooperate in the context of a federation.

This includes suitable RESTful [13] interfaces for the Data Manager to be informed of the constraints; to provide information regarding the current storage location of data; to have differentiated levels of security implementations as needed by the legal analysis; to provide information regarding the location of its datacentres etc. These interfaces must be coupled with the internal framework that keeps track of the current location of data and prevents their movement to domains that are restricted.

In the OPTIMIS project, the Hadoop Distributed File System (HDFS [14]) is used, in order to implement the storage system. On top of this implementation, the RESTful interfaces are applied, that directly manipulate the HDFS configuration, in order to guide the infrastructure. The metadata structures of HDFS are also used in order to notify and update the Cloud provider each time a data action is performed.

6. Developments – Inclusion of Legal Analysis in the Design

a) Creating SLAs by using WS-Agreement

As mentioned in section 3, we consider legal requirements as part of the QoS that can be requested for a Cloud infrastructure where services are going to be deployed. These QoS parameters are agreed upon between the SP and the IP after having negotiated an SLA. The dynamically created SLA, resulting from the negotiations, fixes the OPTIMIS TREC parameters and the legal requirements with respect to data location and encryption. Restraints and requirements discussed in the previous sections have been captured in a term language to be used in the electronic SLA.

For negotiating and creating SLAs, we use the WSAG4J framework developed at the *Fraunhofer Institute SCAI* [9]. WSAG4J is a full implementation of WS-Agreement [10]. Moreover, WSAG4J also provides an implementation of the current draft of the standard for negotiating SLAs based on WS-Agreement: WS-Agreement Negotiation [11].

For the restraints and requirements regarding data placement and data protection in the Cloud we use a schema based on OVF (the Open Virtualisation Format standard of the DMTF) [12]. The following XML-code fragments show a part of the schema (Figure 1) and a section of the corresponding template (Figure 2) where restrictions of data placement and the requirements with respect to encryption of the data are described.

```
<xs:simpleType name="DataProtectionLevelType">
        <xs:restriction base="xs:string">
                <xs:enumeration value="DPA"/>
                <xs:enumeration value="None"/>
        </xs:restriction>
</xs:simpleType>
[…]
<xs:complexType name="EncryptionLevelType">
        <xs:choice>
                <xs:sequence>
                        <xs:element name="EncryptionAlgoritm" type="opt:EncryptionAlgoritmType"/>
                        <xs:element name="EncryptionKeySize" type="xs:int" default="128" minOccurs="0"/>
                </xs:sequence>
                <xs:sequence>
                        <xs:element name="CustomEncryptionLevel" type="xs:anyType"/>
                </xs:sequence>
        </xs:choice>
</xs:complexType>
```

Figure 1: Schema definition of the data protection requirements

```
<ws:ServiceDescriptionTerm ws:Name="DataConstraints" ws:ServiceName="MultipleImages">
     <opt:DataProtectionSection>
       <opt:DataProtectionLevel>DPA</opt:DataProtectionLevel>
       <opt:DataEncryptionLevel>
         <opt:EncryptionAlgoritm>AES</opt:EncryptionAlgoritm>
       </opt:DataEncryptionLevel>
     </opt:DataProtectionSection>
    </ws:ServiceDescriptionTerm>
```

Figure 2: SLA template for the legal data protection requirements

While the dynamic SLA created between SP and IPs in conjunction with a written ("paper") framework contract can provide binding guarantees for the two parties, additional verification of whether the terms of the SLA are fulfilled or violated is required. In the first implementation of the OPTIMIS toolkit, monitoring is already available for the TREC parameters. Monitoring of fulfilment of legal requirements will be implemented in the next iteration. Clearly, the most challenging aspect already identified is the automatic verification of the location of the datacentre storing the data.

b) Design of a location-aware Data Management System

The Data Manager Storage System is the key element to achieve compliance with the already mentioned legal requirements. By using the Management Storage System, cloud-enabled applications such as CRM will be able to handle data in a data protection compliant way, even if the application itself does not support such a feature.

The challenges we face with regard to data management and data security in a federation have to be resolved in a practical and secure way. The simple case of a single IP having total control of all data nodes does not apply to a federated cloud. In this extended case we have multiple IPs sharing data among each other. The initial IP uses federated resources that may be placed in different geographical locations. Given that the legal requirements are correctly expressed in the SLA, the OPTIMIS Data Manager (DM) is able to receive them and act upon them accordingly, by mediating the HDFS placement policy and demanding that the blocks of personal data are kept within the geographical domain specified by the end user. To this end, we take advantage of HDFS's rack awareness characteristics in order to ensure that during runtime the relevant policies are followed. Furthermore, replication aspects are handled through HDFS's interfaces and can be manipulated so that the necessary reliability can be achieved as demanded by the legal requirements. For example, replicas may be placed in different racks or even in different datacentres so that no single point of failure exists, neither at the rack nor at the geographical location level.

However, in order to be proactive, the DM of the IP also exposes a RESTful interface that provides an XML description of the location of its datacentres. This is necessary during the selection process performed by the SP. Thus the SP can filter out the IPs that cannot meet the legal requirements. Furthermore, during the initial account creation by the SP on the IP Data Manager, a flag is used in order to indicate whether the specific service contains personal data as defined in the Data Protection Directive and thus needs the advanced data location management offered by the DM.

Furthermore, the DM will expose a GUI to end users, through which the latter will be able to monitor at any time the location of their data and the history of their transfers to other locations (e.g. in the case of federation or load balancing decisions). More information on the details of OPTIMIS data management can be found in [15]. In Figure 3 we provide a graphical overview of the Data Management System.

c) Security measures regarding the confidentiality of personal data during storage and transfer phase

OPTIMIS IPs provide the customers with a secure storage device that can be used to store sensitive information. This storage device is encrypted at block device level and in real time, providing seamless transparent secure storage that can be used by all applications running in the customer's virtualised environment. The keys to decrypt the storage device can be stored outside the virtualisation infrastructure of the IP, with even the option of hosting it within the SP's infrastructure. The key management server that is responsible for releasing the keys can be instructed to do so based on policy rules associated with the environment of the VM, like the OS, the CPU architecture, IP address etc.

Data moved between the VM and the storage is protected by using the SSHFS protocol which allows for application level encryption on top of the secure storage provided. This protects the data from snooping and other eavesdropping attacks during transfer.

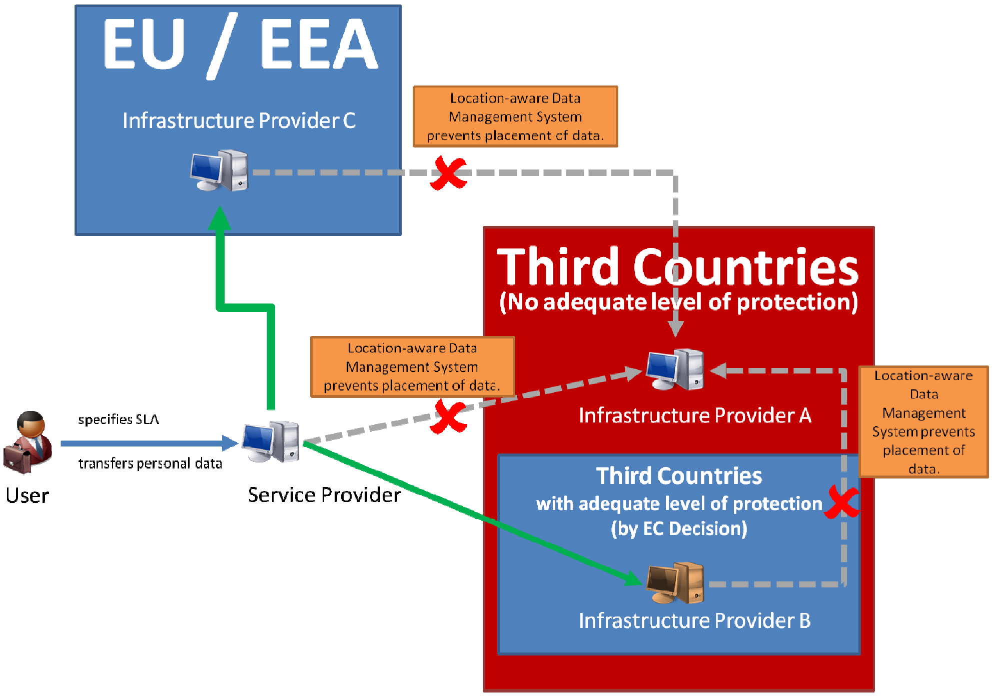
EU / EEA
Infrastructure Provider C
Location-aware Data Management System prevents placement of data.
Third Countries
(No adequate level of protection)
Location-aware Data Management System prevents placement of data.
Infrastructure Provider A
Location-aware Data Management System prevents placement of data.
specifies SLA
transfers personal data
User
Service Provider
Third Countries
with adequate level of protection
(by EC Decision)
Infrastructure Provider B

Figure 3: Graphical overview of Data Management System in OPTIMIS according to legal requirements of the Data Protection Directive.

7. Conclusions

Data processing involving personal data within a cloud is subject to clear restrictions with regard to countries not providing an adequate level of protection. For this purpose, we created an SLA framework by using WS-Agreement to let customers express where personal data may be transferred to, according to the legal requirements pursuant to the DPD. The Data Manager System is able to receive these requirements and distribute the data correspondingly.

The user should be able to constantly track the location of the data processing location by means of a monitoring tool. The distributed file system used in the cloud must support country-specific location for placement of personal data and prohibit placement of data in countries other than within the territorial scope of the DPD in case additional legal safeguards (Standard Contractual Clauses, Binding Corporate Rules) are not agreed in the SLA. Users processing data not relating to an individual should be able to separate their data to profit from improved performance and lower costs.

As regards legal data security obligations, data in clouds should always be replicated and stored redundantly elsewhere to avoid a Single Point of Failure. In order to protect the confidentiality and integrity of the data as required by the DPD, both transmission and storage of personal data demand efficient encryption.

In the OPTIMIS project, the aforementioned requirements are actively taken into consideration, in order to create a toolkit that will enable Cloud providers to abide by legal requirements in all modern scenarios of Cloud infrastructure usage.

Finally, it should be noted that globalisation in information management can have significant benefits with regard to cost or reliability of the infrastructures, but the fact remains that the current legal requirements within the EU prevent these benefits from being fully utilised. By using a technical framework that corresponds to these requirements, European Cloud customers and Cloud providers will be able to exploit cloud infrastructures that reside in different continents, while remaining compliant with European data protection legislation. Furthermore, they can extend their business pool to cases or applications that process personal data.

8. Acknowledgement

This research is partially funded by the European Commission as part of the European IST 7th Framework Programme through the project OPTIMIS under contract number 25715.

References

[1] European Network and Information Security Agency (ENISA) 2009. Cloud Computing – Benefits, risks and recommendations for information security [online]. Available at: http://www.enisa.europa.eu/act/rm/files/deliverables/cloud-computing-risk-assessment/at_download/fullReport [Accessed 20 April 2011], p. 38.

[2] Kuner, C., 2007. European Data Protection Law: Corporate Compliance and Regulation. 2nd ed. Oxford: Oxford University Press, p. 153.

[3] For further information see Commission decisions on the adequacy of the protection of personal data in third countries [online]. Available at: http://ec.europa.eu/justice/policies/privacy/thridcountries/index_en.htm [Accessed 20 April 2011].

[4]Commission Decision 2010/87/EU of 5 February 2010 on standard contractual clauses for the transfer of personal data to processors established in third countries under Directive 95/46/EC of the European Parliament and of the Council.

[5] Helbing, T. How the New EU Rules on Data Export Affect Companies in and outside the EU [online]. Available at: http://www.thomashelbing.com/en/how-new-eu-rules-data-export-affect-companies-and-outside-eu [Accessed 20 April 2011].

[6] Von Eicken, T. Amazon EC2 outage: summary and lessons learned, RightScale Blog [online]. Available at http://blog.rightscale.com/2011/04/25/amazon-ec2-outage-summary-and-lessons-learned/ [Accessed 26 April 2011].

[7] Evans, C. So Your AWS-based Application is Down? Don't Blame Amazon [online]. Available at http://www.thestoragearchitect.com/2011/04/22/so-your-aws-based-application-is-down-dont-blame-amazon/ [Accessed 26 April 2011].

[8] Finley, K. Stop Blaming the Customers - the Fault is on Amazon Web Services [online]. Available at http://www.readwriteweb.com/cloud/2011/04/almost-as-galling-as-the.php [Accessed 26 April 2011].

[9] WSAG4J – WS-Agreement for Java [online]. Available at http://packcs-e0.scai.fraunhofer.de/wsag4j

[10] WS-Agreement – Web Services Agreement [online]. Available at http://www.ogf.org/documents/GFD.107.pdf.

[11] WS-Agreement Negotiation – Web Services Agreement Negotiation [online]. Available at https://forge.gridforum.org/sf/docman/do/downloadDocument/projects.graap-wg/docman.root.current_drafts.ws_agreement_negotiation_specifi/doc16194.

[12] OVF - Open Virtualization Format [online]. Available at http://www.dmtf.org/sites/default/files/standards/documents/DSP0243_1.0.0.pdf.

[13] Roy T. Fielding and Richard N. Taylor. 2002. Principled design of the modern Web architecture. ACM Trans. Internet Technol. 2, 2 (May 2002), 115-150. DOI=10.1145/514183.514185 http://doi.acm.org/10.1145/514183.514185.

[14] http://hadoop.apache.org/common/docs/current/hdfs_design.html.

[15] OPTIMIS Project D1.2.2.1 (MD1) OPTIMIS Detailed Design, UMEA and other partners, November 2010.