8-2010

Master Data Compliance: The Case of Sanction Lists

Jochen Kokemüller
Fraunhofer IAO, jochen@kokemueller.de

Master Data Compliance: The Case of Sanction Lists

Jochen Kokemüller
Fraunhofer IAO
Jochen.Kokemueller@iao.fraunhofer.de

ABSTRACT

Sanction lists as published by national and supranational organizations contain details on sanctioned entities. Those lists have to be obeyed in order to avoid legal implications. Yet, sanction lists are of very low information quality. Nevertheless, regulatory compliance demands, that organizations check their customer master data against sanction lists. We analyze sanction lists in this contribution with respect to their information quality and derive from this analysis requirements on a compliant system. We present a case study of a software vendor that equipped its information system with an extension that enables organizations to comply with sanction lists. We provide details on its implementation and evaluation.

Keywords

Master Data Management, Regulatory Compliance, Sanction Lists, Case Study, Embargo.

INTRODUCTION

In the past embargos limited the external trade with certain countries. Their main focus was the suppression of all kinds of actions and legal endorsements with a whole country. Depending on the restrictions an embargo imposed, it can be differentiated between total, partial and weapons embargo. After the September 11 attacks the situation changed. A new entity related embargo mechanism came into place. Subject to this kind of embargos are individual and legal entities independently of their location, such as natural persons, organizations and enterprises. The usage of this kind of embargo is increasing and it can be observed to substitute over time the classical country wide embargo.

While this entity specific embargo seems more desirable, as it does not oblige sanctions on uninvolved civilians, it is more difficult to implement. To enforce these embargos enterprises now have to check every transaction, whether it corresponds to a sanctioned entity or not. Essentially, organizations need to check, whether their customer master data contains sanctioned entities. Yet, it is in the nature of many sanctioned entities to hide reliable contact information from the public. Furthermore, intelligence agencies that might possess reliable information are usually not willing to publish them neither. This results in sanction lists, that are incomplete, ambiguously, contradictory, inaccurate thus error-prone. This is actually an information quality issue, as the provided information has a limited "fitness for use". At the end of the day, compliance demands from companies to take reliable decisions on unreliable information.

Regulatory compliance, as we discuss it here, refers to the goal that public agencies and organizations pursue in their efforts to ensure that their personnel are aware of and take steps to comply with relevant laws and regulations. While the involved steps may be taken in several areas, we focus here on the compliance related to master data (Kokemüller and Weisbecker, 2009). Master data is defined by the International Organization for Standardization (2009) as the data that "describes entities that are independent and fundamental for the organization; [It] needs to be referenced in order to perform transactions". This definition emphasizes the importance of master data compliance: As master data is referenced by transactions, non-compliant master data will result in non-compliant transactions. Furthermore, as master data is referenced often, its non-compliance is multiplied into various transactions of an organization. This makes master data an effective starting point for a compliance initiative.

Design science research contributions present novel Information Systems (IS) artifacts and suitable evaluation approaches that address the artifact's appropriateness to contribute to the problems' solution (Nunamaker Jr, Chen and Purdin, 1991). These two facets of rigorous design science-oriented research contribute to the foundations and the methodologies pool of Information Systems research, i.e. they contribute to its knowledge base (Hevner, March, Park and Ram, 2004). In our work we follow this research paradigm.

We therefore first analyze the problem domain, design and implement a suitable IS artifact and evaluate it then against the problem domain. The remainder of this paper is thus structured as follows: First, we continue with legal implications and information quality issues related to regulatory compliance with sanction lists. The next section then provides a brief review of relevant literature before we continue with the systems design. Afterwards, we provide details on its realization and evaluate it against the problem domain. We then discuss our approach and address some limitations before we conclude.

Legal implications

In Germany embargos are enforced by supervisory duties[1] whereto: The Management Board and their authorized representatives have to take reasonable organizational measures to ensure with a sufficient degree of probability that no infringements are committed due to negligence or intent. In case of infringements of the German Foreign Trade Act[2], the responsible executives in the company are personally liable for penal action. These persons can only be discharged through evidence of a working organization and the effective performance of the required and appropriate supervisory duties.

The German legislator formulated a severe threat of punishment: In the worst case, up to 15 years of imprisonment and a fine of up to 500,000 Euro, turnover levy, forfeiture and confiscation. Without the contribution of significant criminal energy, a violation of the EU antiterrorism regulation is likely to be considered as breach of embargo pursuant[3]. In this case, the threat of punishment is not less than two years of imprisonment, which – if convicted – typically excludes a suspension on probation. In minor cases the threat of punishment ranges from three months to five years of imprisonment (Gensior-Mages and Drabe, 2009).

For German address owners, there are three relevant entities, which define sanctioned entities: the UN[4], EU[5] and the German Government[6]. Whereas due to international treaties the UN lists are integrated into the EU lists and the EU and German lists are essentially synchronized. The German lists are subject to a fee and include, additionally to the EU lists the current US sanction lists. Some countries outside the EU define additional sanction lists, those countries include the US and Japan. Those sanction lists may be ignored by German address owners while they are not active in those countries, otherwise they must be respected.

Information Quality

Information quality is often defined as the information’s “fitness for use”. For a deeper understanding of this concept, Wang and Strong (1996) identified 15 relevant information quality dimensions. These 15 dimensions are grouped into 4 categories (Intrinsic, Contextual, Representational, and Accessibility). For our discussion we divide the dimension of completeness further. We differentiate between coverage, i.e. how many of all sanctioned entities are included in the list and density, i.e. how exhaustive are the entries of a sanctioned entity. In Table 1 the EU sanction list is analyzed in each of these 15 dimensions. For three exemplary sanction list entries, please refer to Appendix A. This analysis led to four different types of results. The information is in some quality dimensions normative (Believability, Objectivity, Reputation, Timeliness, and Coverage), the next type is of mere observational character (Relevancy, Value-added, Concise representation, consistent representation, understandability, and ease of manipulation). The dimensions of the third group are requirements on a compliant information system (Interpretability, Timeliness and Accessibility). Finally, there are real drawbacks (Free of error, Density, and Appropriate amount of data) in information quality that form challenges for a compliant information system.

For a compliant implementation of sanction lists, organizational processes need to be adjusted and documented to prove their effectiveness and reasonable coverage. It has to be assured, that every transaction that would lead to a proliferation of economic resources is correctly checked against the sanction list. Obviously, this not only requires the implementation of adequate processes but equally important the adjustment of information system to perform necessary checks. The SPH AG is an independent software vendor that develops a Customer Relationship Management system specially tailored for direct

[1] §130, §9 OWiG (Administrative Offence Act), §14 StGB (Penal Code)

[2] §70 AWV (German Foreign Trade Act) and AWG (German Foreign Trade Law)

[3] § 34 (4) AWG (German Foreign Trade Law)

[4] http://www.un.org/sc/committees/1267/consolist.shtml

[5] http://ec.europa.eu/external_relations/cfsp/sanctions/consol-list_en.htm

[6] http://www.awr-portal.de/

marketing. It decided to develop a specialized service as an enhancement to its existing solution. SPH AG essentially decided to offer to its customers a service to automatically detect sanctioned entities in their customer data base. To this end an algorithm to precisely identify sanctioned entities was developed. This contribution provides details on this case.

Table 1: Analysis of information quality dimension with respect to sanction list compliance

Information Quality Dimension	**Analysis of sanction lists in specific dimension**	**Classification of result**
Category: Intrinsic		
Believability	The information source is normative with respect to these information quality dimensions. Although the score in these dimensions might not always be very high, other sources may only be worse. This is mainly due to the political decision involved in formulating an entity specific embargo.	Normative
Objectivity		
Reputation		
Free of error	The content of the sanction lists are contradictory and are characterized by low accuracy.	Challenge
Category: Contextual		
Timeliness	With respect to timeliness the information source is normative. Therefore, only the most recent sanction list should be employed.	Normative / Requirement
Relevancy	The information is for compliance very relevant.	Observation
Value-added	The information adds value as it reduces the risk for fines.	Observation
Completeness	Here it must be distinguished between coverage and density. As it is normative, the sanction list is complete in coverage. Nevertheless, the density is very low, thus provides only very limited attributes for comparison.	Normative / Challenge
Appropriate amount of data	The amount of data is unsatisfactory, as an unambiguous decision whether a contact is a sanctioned entity or not is in most cases not possible.	Challenge
Category: Representational		
Interpretability	Compliance demands, that the user of an information system has very low interpretability, whether a contact is sanctioned or not. This requires that the decision is taken automatically by reproducible and well defined rules.	Requirement
Concise representation	Sanction lists are usually available as XML Data.	Observation
Consistent representation	To the EU sanction list a DTD schema is present.	Observation
Understandability	Contained data is usually understandable independently, yet the usefulness of the address “By the Shrine Next to the Gas Station” (Figure 2) is questionable.	Observation
Category: Accessibility		
Accessibility	Likewise, compliance demands, that sanctions are enforced. Thus, the information has to be easily accessible. Note, that this does not require that a user is informed that a contact is a suspected terrorist but only that transactions have to be stopped.	Requirement
Ease of manipulation	Sanction lists are and should not be editable by companies	Observation

LITERATURE REVIEW

Challenges coming from error-prone and low density data are related to fuzzy searches. The information quality literature discussed these algorithms under a multitude of names. Most common are reference reconciliation, merge/purge problem, record linkage, duplicate (record) detection, or duplicate elimination. While some of these may contain slightly different connotations, at the end of the day they pursue the same goal: Identifying duplicate references to identical real world entities in noisy data. Elmagarmid, Ipeirotis and Verykios (2007); Gu, Baxter, Vickers and Rainsford (2003) and Rahm and Do (2000) give a concise resume of the current practice.

In general the problem is twofold. First, in comparing all *n* customers with all *m* sanctioned entities, the complexity of the algorithm is of the order $\mathbb{O}(n \times m)$. Especially *n* may grow large, which renders a search over the complete Cartesian product unfeasible. On this behalf Hernández and Stolfo (1998) proposed the Sorted-Neighborhood Method (SNM). Here records are ordered using a key build of the attributes that contain the most information. This results in a sequence where similar records are situated close to each other. Next, only records within a window of a certain width are compared. A different approach is chosen by Monge and Elkan (1997), they assume that similar objects are always found close to each other and compare records only with the 4 previously encountered duplicate clusters. Monge and Elkan (1997) show, that their algorithm runs efficiently, yet the precision[7] and recall[8] are not as high as in the SNM approach. Both algorithms though are designed for comparing one dataset internally, thus are efficient in selecting only those records of one dataset that have a high probability of being duplicates. In comparing customers with a sanction list, records from two different data sets need to be compared.

Once decided which records to compare, the second problem domain needs to be addressed: Defining how two records are compared. Cohen, Ravikumar and Fienberg (2003) analyze several string distance metrics. They conclude that the tf.idf algorithm performs best in their testbed. This token-based similarity measure assigns higher importance to tokens (token frequency: tf) that are less common in the document (inverse document frequency: idf). Bilenko, Mooney, Cohen, Ravikumar and Fienberg (2003) discuss a trainable algorithm that may perform well if trained well. Nevertheless, regulatory compliance to sanction lists demands a reproducible outcome of an algorithm, something a trainable algorithm cannot provide.

SYSTEM DESIGN

As we had observed, that both low *interpretability* and high *accessibility* have to be ensured, we now design an IS artifact that extends an information system providing accessible and unambiguous information to the user. The first three use cases describe the scenarios in which addresses need to be checked against the sanction list:

- *Cleansing of entire database:* In case of a newly released sanction list, the whole customer data set is checked against the new sanction list. Likewise, this is done in initial cleansing.
- *Verification of single entity:* Whenever an address record is created, the software checks whether the entity is included in the sanction list.
- *Cleansing of external lists:* Externally purchased or leased lists are checked against the sanction list prior to using them. Suspicious records should be excluded from further processes (advertising measures, inclusion in the system)

The final two use cases describe how suspicious addresses are handled:

- *Automatic exclusion of identified addresses:* If an address can be marked unambiguously as sanctioned, it is blocked automatically.
- *Manual batch processing:* Decisions in ambiguous cases are decided by authorized personnel manually.

Requirements

We now continue with requirements a system needs to fulfill to provide above presented use cases. Observing the sanction lists we recognize that most attributes are left empty. Usually no address, occasionally a town, sometimes only a first or last name is given. The demands arising out of the information quality dimensions of *density (completeness)* and *appropriate amount of data* lead us to the first requirement:

Requirement I The comparison needs to function with sparse data.

[7] Precision increases, if detected duplicates are real duplicates. High precision may be achieved by comparing entities very strictly. Thus this optimization leads to many undetected duplicates.

[8] Recall increases, with the number of detected real duplicates. High recall may be achieved by detecting as many duplicates as possible. Yet, this also increases the number of entities falsely identified as duplicates.

Addresses in the sanction list are of different types. The XML Representation of the EU list distinguishes “P” for Individuals and “E” for organizations as companies or institutions:

Requirement II The comparison needs to cope with different entity types.

Regulatory compliance demands from an organization to enforce the embargo by all reasonable measure. Therefore it is required:

Requirement III The comparison must produce as many correct hits as possible. At the same time it has to be very strict otherwise a significant amount of post-processing is necessary or lost sales opportunities would occur.

Especially transcribed names are included in the sanction list in a variety of spelling alternatives or mistakes. To cope with the information quality dimension of *freeness of error* we thus require:

Requirement IV High variety in spelling alternatives should be resolved. Including the recognition of phonetic differences and similarities.

Additionally, entities are represented with several different names, addresses, birthdates, and passports. We are thus left with a sparse but complex comparison space. Based on the information quality dimension of *freeness of error* we additionally require:

Requirement V Attributes with cardinality higher than one should correctly be assigned to a single entity.

REALIZATION

To reduce the number of necessary comparisons, a Multipass-SNM approach was chosen. Effectivly, to every entry in the sanction list several keys are computed. A key is configurable but it is usually assembled of the first few letters of an attribute with high information density as firstname, lastname, or wholename. Different keys of one entity are often permutations of other keys. This is useful, as keys are used for lexicographic ordering and therefore permutations define different sequences. Figure 1 shows exemplary keys for the sanctioned entity shown in Figure 4 (Appendix A).

For a contact, the same keys are computed. Afterwards, for every key the index in the sanction list is searched, where it would belong following the lexicographic order. Then all sanctioned entities within a specified window surrounding the computed virtual insert location are identified. Those entities are added to the initially empty comparison list. The procedure is repeated for the next key configuration and the identified sanctioned entities are merged into the comparison list, thus duplicate entries are only included once. This procedure is repeated until all configured keys are processed. In this manner the group of sanctioned entities for pair wise comparison is identified.

The field `wholename` contains often permutations of the fields `firstname` and `lastname`. One could argue that the field `wholename` should then be discarded. Yet, it has a much higher density than the other two fields. Moreover, the name of an organization is always included as a `wholename`. Therefore, only the field `wholename` is used for comparison. The fields `firstname` and `lastname` are used to fill the field `wholename` if it is encountered empty.

The comparison is then done using a token-based similarity measure. Strings are split into n overlapping tokens of length q called q-gramms. To every q-gramm it is then counted how many times it occurs in the contact n_q^c respective the sanctioned entity n_q^s. The similarity of two strings is then the normalized Euclidean product of the q-gramm counts. The normalization is achieved using the number of all q-gramms of the contact n^c and the sanctioned entity n^s. The similarity S_i of the attribute i is thus given by

$$S_i = \frac{1}{n^c + n^s} \sum_q n_q^c \cdot n_q^s .$$

Finally, we compute the weighted similarity S for the entire entity using the weights w_i

$$S = \sum_i w_i \cdot S_i .$$

The result of the comparison is limited to the most similar sanctioned entity, its similarity measure, and a similarity classification. Based on configurable threshold contacts are classified as either certainly, probably, probably not, and certainly not being a sanctioned entity.

The algorithm is implemented to be configurable with respect to the sorting keys, the q-gramm length and class boundaries. Some additional measures have been taken, to improve the run-time behavior. Especially all computations, which can be done during initialization, are performed then. This increases the need for memory while it decreases the processing time needed for a single check. Most important, all computations on the sanction lists, as the parsing, key generation, and sorting are done during initialization.

To prove the compliance with regulations an organization has to provide evidence of the effective performance of the required and appropriate supervisory duties. In other words, it has to provide evidence, that a certain contact was checked against the sanction lists using documented parameters. To this end a log is written that documents every compared contact and its result. Additionally, every change in parameters is logged.

EVALUATION

The evaluation of the approach was done in a focus group interview with five customers of the SPH AG. During the interview it was recognized by the customers, that a comparison based on the identity of attributes of the customer data with the sanction list is not sufficient. This is owed to the poor information quality, especially in the dimensions *free of error*, *completeness* and *appropriate amount of data*. In contrast, the above described approach and algorithm was considered to provide a reasonable solution. Especially, due to the configurability a broad acceptance could be achieved. This is understandable as it is a business decision, whether to take more risk or to loose more sales opportunities. To this decision, only guidelines for the configuration but no fixed solution may be provided.

The presented algorithm aggregates attribute values to cope with the special challenge of sparse data (Requirement I). By comparing entities based on the wholename, persons and organizations can be detected likewise (Requirement II). The matching algorithm and its thresholds are configurable. This enables organization to choose the configuration that best fits their needs (Requirement III). Currently, attributes with a cardinality higher one result in several entries of one entity. While this is possibly not the best solutions, it assigns every entry to an entity (Requirement V). It is questionable, whether other approaches would lead to better results. At the same time, the chosen approach guarantees, that to every entity all known spelling alternatives are correctly assigned (Requirement IV). We recognize that all requirements are met.

Those requirements were formulated to make best use of sanction lists in spite of their low information quality. The information quality dimension of *freeness of error* has been addressed by fuzzy search algorithms. Here we employed algorithms for duplicate detection. The dimension of *density* and *appropriate amount of data* are addressed by aggregation information into the field wholename. Finally, the information quality dimensions of low *interpretability* and high *accessibility* had to be dealt with. In this paper, this was done by proposing an IS artifact as an extension to an existing Customer Relationship Management solution. It was thus effectively integrated into already deployed business processes involving contact data.

Furthermore, the implementation was evaluated in its run-time behavior. To this end it was deployed in a virtual Windows

Key 1	**Key 2**	**Key 3**
BinUsa	UsaBin	Abu
BinUsa	UsaBin	Al
BinUsá	UsáBin	Ben
		Ben
		Ben
		Bin
		Bin
		Muh
		Ous
		Sha
		Usa
		Usa
		Usá

Figure 1: Exemplary Keys

XP machine on VMWare ESX running on an IBM Blade Center. The guest environment can therefore be compared to a standard PC. Certainly, this is not a very accurate description, yet there are many influences on the run-time behavior, including the .Net Framework, that an error free measurement seams unfeasible.

For the run-time behavior it has to be distinguished between the first call and all subsequent calls. As during the first call several start-up computations are performed it takes approx. 3s, thus non-negligible time. All subsequent calls profit from this trade off. They need an average comparison time per customer of 0.004s. This is of the same magnitude as algorithms for duplicate detection. We therefore consider it as a good result.

DISCUSSION AND LIMITATIONS

Most enterprises need to run a compliant business with respect to sanction lists. At the same time, the lists provide by the governmental authorities are only of very little information quality. A service provided by the government with a contact as its arguments that would give a concise answer whether a contact is a sanctioned entity or not could provide reliable answers. Yet, enterprises do not want to display all their business contacts to the authorities which renders this option void. Effectively, every enterprise needs to check its contacts on its own against sanction lists. Based on an analysis of sanction lists in 15 information quality dimensions, we formulated requirements for a compliant information system. While several algorithms have been presented to resolve duplicate entries in customer data bases, the problem domain of reliably identifying sanctioned entities in the customer data base still bears several challenges. In this contribution we showed how algorithms for duplicate detection may be applied to this problem domain.

Several enhancements to our algorithm might be thought of. First of all, phonetic string comparisons could be implemented. This is a challenging task, because names in the sanction lists are in various languages that are difficult to compare. Still, it may provide useful results we therefore plan this as future work. Additionally, we recognize that the string comparison algorithm that we are currently using could be enhanced by using a tf.idf algorithm.

Sanction lists are available in several formats (xls, pdf, xml, etc.). For automatic comparisons the xml versions are especially useful. The xml lists are furthermore normalized. To every entity potentially several names, addresses, birthdates, and passports are assigned. Currently, these cardinalities are not evaluated. Every entry is currently seen as essentially one independent entity. Nevertheless, several entries may represent the same entity resulting from denormalization. Thus, the extension of the xml lists do not suffer any information loss, only information based on the intention may get lost and could be used to further enhance the algorithm.

CONCLUSION

An increasing use of sanction lists for the enforcement of entity specific embargos can be observed. Legislators usually formulate substantial threats on those, who do not comply with sanction lists. It is therefore important to employ reliable mechanisms for master data compliance. The natural choice, to achieve compliance regarding the customer master data, is by extending a customer relationship management solution. Realizing that, the SPH AG developed a specialized service as an extension to its solution to automatically detect sanctioned entities.

Yet, this is not a trivial task, as sanction lists are of poor information quality. In this contribution we therefore analyzed sanction lists with respect to its information quality in 15 dimensions. Based on this analysis we then designed an information systems artifact, which provides a reliable comparison with low quality data. This works as an enabler to master data compliance with respect to sanctioned entities, being individual or legal entities.

Following the design science approach we first analyzed the problem domain of master data compliance in the special case of sanction lists. We then presented the concept and architecture of an IS artifact that enables organizations to comply with sanction lists. We than evaluated it against the formulated requirements, which is according to Hevner et al. (2004) a suitable descriptive method for the evaluation of an IS artifact. In this contribution we have discussed the need for compliance with sanction lists. We then outlined the organizational requirements on a compliant information system. On the implementation of the IS artifact we showed how those requirements may be met.

REFERENCES

1. Bilenko, M., Mooney, R., Cohen, W., Ravikumar, P. and Fienberg, S. (2003) Adaptive name matching in information integration, *Intelligent Systems, IEEE*, 18, 5, 16-23.

2. Cohen, W. W., Ravikumar, P. and Fienberg, S. E. (2003) A comparison of string distance metrics for name-matching tasks, *Proceedings of the IJCAI-2003 Workshop on Information Integration on the Web (IIWeb-03).*

3. Elmagarmid, A. K., Ipeirotis, P. G. and Verykios, V. S. (2007) Duplicate Record Detection: A Survey, *IEEE Transactions on Knowledge and Data Engineering*, 19, 1, 1-16.

4. Gensior-Mages, S. and Drabe, H. (2009) EG-Anti-Terrorism Resolution and sanction lists: Possible technical implementation of legal regulations, *7th German Information Quality Management Conference & Workshop*, November, 6-7, Potsdam.

5. Gu, L., Baxter, R., Vickers, D. and Rainsford, C. (2003) Record linkage: Current practice and future directions, *CSIRO Mathematical and Information Sciences Technical Report*, 3, 83.

6. Hernández, M. A. and Stolfo, S. J. (1998) Real-world Data is Dirty: Data Cleansing and The Merge/Purge Problem, *Data Mining and Knowledge Discovery*, 2, 9-37.

7. Hevner, A. R., March, S. T., Park, J. and Ram, S. (2004) Design Science in Information Systems Research, *MIS Quarterly*, 28, 1, 75-105.

8. International Organization for Standardization (2009) Technical Specification ISO/TS 8000-100:2009(E), Geneva, Switzerland.

9. Kokemüller, J. and Weisbecker, A. (2009) Master Data Management: Products and Research, *Fourteenth International Conference on Information Quality*, November 7-8, 2009, Potsdam, 8-18.

10. Monge, A. E. and Elkan, C. (1997) An efficient domain-independent algorithm for detecting approximately duplicate database records, *Research Issues on Data Mining and Knowledge Discovery*, 23-29.

11. Nunamaker Jr, J. F., Chen, M. and Purdin, T. D. M. (1991) Systems development in information systems research, *Journal of Management Information Systems*, 7, 3, 89-106.

12. Rahm, E. and Do, H. H. (2000) Data Cleaning: Problems and Current Approaches, *IEEE Data Engineering Bulletin*, 23, 4, 3-13.

13. Wang, R. Y. and Strong, D. M. (1996) Beyond accuracy: what data quality means to data consumers, *Journal of Management Information Systems*, 12, 4, 5-33.

APPENDIX A: EXEMPLARY SANCTION LIST ENTRIES

NAME				
LASTNAME	FIRSTNAME	MIDDLENAME	WHOLENAME	GENDER
			Al-Shifa Honey Press For Industry And Commerce	
ADDRESS				
NUMBER	STREET	ZIPCODE	CITY	COUNTRY
	Al-Hasabah	PO Box 8089	Sanaa	YEM
	By the Shrine Next to the Gas Station, Jamal Street		Taiz	YEM
	Al-Arudh Square, Khur Maksar		Aden	YEM
	Al-Nasr Street		Doha	QAT

Figure 2: Exemplary sanction list entry

NAME				
LASTNAME	FIRSTNAME	MIDDLENAME	WHOLENAME	GENDER
Blé Goudé	Charles			M
			Gbapé Zadi	
ADDRESS				
NUMBER	STREET	ZIPCODE	CITY	COUNTRY
	Bloc P 170		Yopougon Selmer	CIV
	Hotel Ivoire		Abidjan, Cocody	CIV
BIRTH				
DATE	PLACE	COUNTRY		
1972-01-01	Guibéroua (Gagnoa)	CIV		
1972-01-01	Niagbrahio/Guiberoua	CIV		
1972-01-01	Guiberoua	CIV		
PASSPORT				
NUMBER	COUNTRY			
04LE66241 (issued on 2005-11-10, valid until 2008-11-09)	CIV			
AE/088 DH 12 (Diplomatic passport, issued on 2002-12-20, valid until 2005-12-11)	CIV			
98LC39292 (issued on 2000-11-24, valid until 2003-11-23)	CIV			
C2310421 (Travel document issued on 2005-11-15, valid until 2005-12-31)	CHE			
CITIZEN				
COUNTRY				
CIV				

Figure 3: Exemplary sanction list entry

NAME						
LASTNAME	FIRST-NAME	MIDDLE-NAME	WHOLENAME	GENDER	TITLE	LANGUAGE
Bin Laden	Usama	Muhammed Awad		M	Shaykh	
			Usama Bin Muhammed Bin Awad, Osama Bin Laden			
			Abu Abdallah Abd Al-Hakim			
			Oussama Ben Laden			FR
Bin Laden	Usama				Hajj	
			Ben Laden Osama			
			Ben Laden Ossama			
			Ben Laden Usama			
			Bin Laden Osama Mohamed Awdh			
			Bin Laden Usamah Bin Muhammad			
			Shaykh Usama Bin Ladin			
			Usamah Bin Muhammad Bin Ladin			
			Al Qaqa			
			Usáma bin Ládin			CS
Bin Ládin	Usáma		Muhammed Awad			SK
BIRTH						
DATE	PLACE	COUNTRY				
1957-07-30	Jeddah	SAU				
1957-07-28	Jeddah	SAU				
1957-03-10	Jeddah	SAU				
1957-01-01	Jeddah	SAU				
1956	Jeddah	SAU				
1957	Jeddah	SAU				
1957-07-30		YEM				
1957-07-28		YEM				
1957-03-10		YEM				
1957-01-01		YEM				
1956		YEM				
1957		YEM				
CITIZEN						
COUNTRY						
SAU						
AFG						

Figure 4: Exemplary sanction list entry