***Consultation Paper***

***Draft technical standards on access to data and aggregation and comparison of data across TR under Article 81 of EMIR***

**KDPW\_TR, 2016-02-01**

**General Comments**

Publishing the Consultation Paper on “Review of the EMIR Technical Standards on data access and operational standards for comparison and aggregation of data” will have a significant impact on previous proposals and agreements between TRs and ESMA concerning TRACE project. We realize that the feedback to this consultation might also result in changes to TRACE specification which might cause obstacles for the TRACE implementation. This consultation causes a lot of uncertainty regarding data access for regulators for TRs, in particular the relation between the review proposal and the TRACE project. While TRACE was planned to cover around 30 NCAs through 1 connection (with ESMA), this regulation shall be applied to ALL regulators and it indicates creating a number of SFTP connections between repositories and regulators. There is also many other discrepancies, which we believe must be explained, before any TR can start changing their system to meet the proposed regulatory requirements. To allow regulators harmonized data access, ESMA and all TRs shall work together to a complete and final specification for TRACE, and this should be in line with what is required in changes proposed in this consulted review.

All TRs are willing to cooperate and support regulators with what is required, but, to achieve this goal, those requirements have to be clear, accurate, consistent and realistic, so that every TR will interpret them in the same manner and has a chance to implement them in required time. Repositories already have experience in inter TR projects implementation; there is a process of reconciliation established between all six TRs serving millions of reports every day. We believe the experience gathered while implementing Inter TR reconciliation should now be used for data access harmonization. This can only be successful if repositories’ comments and suggestions are taken into considerations.

We consider the proposed time for implementation (3 months after publication) as much too challenging. The IT projects of that complexity require much more time for building proper specification, implementation and testing. That project will involve multiple participants: 6 TRs and several authorities and to achieve full compatibility and harmonization we would expect minimum 9-12 months for delivery, as it was previously stated in bilateral consultations with ESMA. Testing phase itself would require at least 3 moths time. However, the more exact time for delivery could be only discussed after full functional specification is published.

As proposed by TRs for TRACE project, we would suggest setting phased timeframes for delivery of the proposed functionalities: standard (recurrent) reports first, communication as phase 2 and ad-hoc reports (those would require more time for testing and implementation) as phase 3. That way, authorities could start using harmonized data earlier (from standard reports) and TRs would have sufficient time to deliver other functionalities. Setting unrealistic deadlines for requirements can cause unnecessary confusion and disorder that will have to be amended after start date, which will only make additional costs for all participants, costs that could be avoided. We also believe the implementation timelines should be adjusted to implementation of new RTS, which will also be significant for TRs in terms of available IT and human resources.

**Q1.**

**Do you foresee any technical issues with the establishment of secure FTP connections between trade repositories and authorities? What are the cost implications of the establishment of secure FTP connections? What other practical difficulties, if any, do you foresee? Please elaborate.**

We do not expect any technical issues with establishment of SFTP between TR and authorities; however we do not recommend using SFTP for communication purposes. Secure FTP is a good tool for delivering reports, but for communication we recommend using a messaging protocol, for example MQ, which is designed to support the request and response messaging. Queries and all the response messages, such as query validation confirmation and data delivery to SFTP confirmation should be sent by MQ, data outputs – results of query could be delivered via STFP connection to the specified, accessed only by an authority (encrypted) folder. MQs are designed to serve multiple participants, and the rule FIFO (first in first out) is used, which allows TR to reply for the queries in the exact order they come in. For SFTP connection those mechanisms would have to be delivered by TR as a part of the service.

**Q2.**

**Do you foresee any technical issues with the above mentioned data exchange supported by ISO 20022 methodology? Do you foresee any cost implication from the establishment of standardized data exchange? Do you foresee any additional benefit from establishing data exchange supported by ISO 20022 methodology? Please elaborate.**

XML is designed for messages, therefore we support the idea of using that protocol for communication – query/response, but we do not support the idea of using it for outputs delivery.

XML allows standardizing data and is a good tool for validation, however one should consider whether data submitted and kept by TRs really needs validation on a high level. Each TR already implemented level 2 validations, so the quality of data is now at possibly high level. One harmonized across all TRs file would serve the purpose. Such a solution does work for reconciliation process; serving communication between all 6 TRs.

More information with regard to the XML output validation that will be included in the schema is required. Each TR already holds data accepted under differing validation rules, i.e. pre-Level 1, post-Level 1 but pre-Level 2, post-Level 2 under current RTS, new RTS. Building one XML schema serving all those sets of validation would result in the schema that serves pre-L1 validation rules, which means – almost all fields optional, but then, is there any sense in using XML?

Another feature of XML files to be considered is data file sizes. Our experience with data delivered to authorities is that XML files are a 3-4 times bigger than files containing exactly the same information, but saved as CSV file. Obviously data outputs for authorities are large sized files from nature, as regulators have access to millions of records. We believe this should also be considered when choosing the appropriate file type as it will have impact on costs of keeping the output files, both for TRs and regulators. For XML files (if that would be the choice) we strongly recommend some type of file compression.

**Q3.**

**Do you foresee any technical issues with the establishment of recurrent and predefined queries?**

**If so, how would authorities be able to compare and aggregate data across TRs in absence of standardized queries and how would they be able to make use of TR data for the exercise of their duties if they are not able to properly and immediately access TR data? What are the cost implications stemming from the establishment of the proposed predefined and ad-hoc queries?**

We do not envisage any technical issues with establishment of recurrent and pre-defined queries, what is more KDPW\_TR already services already cover both types of queries. However, the cost of implementation can only be calculated when more detailed specification is available.

Extending the ad-hoc queries to what is set in the proposal (in particular the requirement for any combination of 34 fields used as filters) will demand vast investments in the TRs’ IT infrastructure. Currently accessible IT products allow to build query facilities that would fulfill the requirements, however the amount of data kept by TRs are significant and it is growing dynamically over time, so there will be the need to future investments also in the upcoming years to keep the efficiency and availability and avoid delays in data delivery.

Querying the databases consisting of millions of records depends on many factors, such as database engine parameters and configuration, but also on the clarity of project requirements and close cooperation of all interested participants. It is very important for TRs to understand authorities’ needs, to avoid misinterpretations, avoid unnecessary costs and ensure that the delivered solution fulfills the requirements and serves the purpose.

**Do you agree with the proposed minimum set of queries?**

We understand the authorities’ needs to be able to filter the data with all reported fields (or sets of fields) that are reasonable to be filtered. For TRs however, more complexity means more costs. It would be practical to consider which data will be most used for filtering, which is really ‘must have’ and set the realistic (compromised for both sides) list of such fields. We also suggest to implement ‘must have fields’ first and then possibly extend the list in later stages when we see the solution is stable, working properly and is used by regulators in every day work.

Much more detailed specification re. “combination of fields” would be much appreciated. For example: combination might mean “OR”, “AND” between filters, lists of filters etc. This should all be specified on a much deeper scale. We believe this is very important to be agreed between TRs and regulators in detail, so that both sender and receiver understand what data is actually required in the same manner.

We believe that also some time windows and maximum data files sizes should be set in order to avoid generation unexpectedly high size files, which are challenging to be moved and stored. We expect regulators to define query with understanding of their own needs and goals, and those ad-hoc ones in particular. We understand ad-hoc query is designed to get the strictly defined data fragments that should be well described, not vast amounts of data (e.g. no more than number of submitted reports daily).

**What would be the maximum number of recurrent queries which a single authority could submit in a given day? What would be the maximum number of ad-hoc queries which a single authority could submit in a given day? Please elaborate.**

The aforementioned TRACE project sets the maximum number of queries: 5 ad-hoc and 3 recurrent per day per NCA. This was agreed at ESMA – TRs level. We agree with those numbers but we suggest TRs should reject all NCA’s queries after maximum number was reached. The idea of serving those queries in a later time will cause unnecessary disorder in serving other NCA’s queries as well. In particular recurrent queries could never be served, as those would wait until any of allowed 3 queries was removed.

**Q4.**

**Do you agree with the proposed frequency to provide data to the relevant authorities? Please elaborate.**

We appreciate the necessity of immediate data access for regulators, but having in mind that around 60 regulators might require 8 queries to be run every day (maximum), which basically means almost 500 queries run every day on million records keeping databases. We understand that all queries could deliver the data as of COB previous day. It would be much more comfortable for TRs to extend 7.00 UTC limit to allow for post COB processes. Thereby regulators would be served with most up-to date data and TRs would have sufficient time for data processing and output deliveries. The detailed time frames for delivery can only be discussed when more detailed specification of the functionalities is delivered. Performance testing can be done to set the exact limits for delivery.

**Q5.**

**Do you agree with this proposal? Please elaborate.**

We agree, encryption of data is required to fulfill privacy policy.

**Q6.**

**Do you agree with this proposal? Please elaborate.**

15 minutes time is enough for response when MQ solution is used. KDPW\_TR already serves many participants using MQ and response time is usually much less than 15 minutes. However, it might be more difficult to achieve when SFTP solution was expected for communication. All TRs would have to search through uploaded queries folder every 13-14 minutes. TRACE project expects TRs to download queries every hour and this seems to be realistic. Searching the incoming folders every 15 minutes, while TR still have got at least several hour time (up to 7.00 next day) to deliver the data is impractical. We strongly recommend MQ tool for communication, as it allows “immediate” response and start of the data delivery process.

**Q7.**

**Do you foresee any technical issues with the implementation of xml template in accordance with the ISO 20022 methodology? Do you foresee any technical issues in translating data received in non xml format to an xml template in accordance with ISO 20022 methodology? Do you foresee any benefit from establishing standardized xml template in accordance with ISO 20022 methodology for the aggregation and comparison of data? Would any other data standard fulfil to the same extent the requirements set out in paragraph 48 with respect to the aggregation and comparison of data by authorities? Please elaborate.**

As much as we appreciate implementation of xml files as a tool to supporting data standardization and harmonization on reporting level, we do not agree with using xml protocol for output files for the reasons mentioned earlier in this document (see Q2). In our view, data submitted currently to TRs is of comparably good quality, due implemented and harmonized validations. As we explored while building inter-reconciliation processes, large data exchange is much more convenient with csv packed files. Size of exchanged file really does matter, not only in terms of the connection (smaller files are uploaded and downloaded faster) but also in terms of data keeping – smaller files are less costly to be kept both by TRs and regulators. The policy of storing and archiving the output files should be developed and harmonized across TRs as well.

Re. data aggregation and comparability, we believe that harmonization of csv file structure across TRs on top of high level harmonized validation rules for records submitted to TRs should be sufficient to enable data aggregation from all TRs and its comparability.

As proposed in the draft of regulations, TRs should not withdraw the currently delivered csv and txt files. This is something we cannot agree with. We believe that once a suitable solution is delivered there will be no reasons to maintain currently available solutions, except weekly reports that were designed for regulators which will not be covered by the regulation, e.g. rejection reports. TRs might be willing to maintain current facilities temporarily, but once the new solution is stable and regulators ready to use it, current data deliveries will have to be stopped in order to avoid the unnecessary costs and allow for better efficiency of new functionalities.

**Other comments**

Although the proposed Amendments to Commission Delegated Regulation (EU) No 151/2013 do not include any changes related to Art. 2 of this Regulation (i.e. data access by relevant authorities), in the light of the upcoming new RTS/ITS, which still fails to explain some important issues, there is a need for further clarification in this regard.

The following issues need to be clarified:

**Beneficiary ID (field 1.12 in the new RTS/ITS)**

Under the new RTS/ITS it is possible to populate field 1.12 (Beneficiary ID) with the Client Code, if beneficiary is a private individual. It should be noted that in this case granting the access rights to relevant authorities, in accordance with Art. 2 of the Regulation No 151/2013, would be impossible due to the difficulties with identification of the private individual’s country of origin. In our opinion this information is a key for correct definition of the scope of access rights granted to competent authorities.

**Underlying identification (field 2.8)**

Under the new RTS/ITS it is allowed to populate the field 2.8 (underlying identification) with the ‘NA’, if the actual value is not accepted or the full name of the index, if ISIN is not available. In such cases, granting the access rights to relevant authorities, in accordance with Art. 2 of the Regulation No 151/2013, would be difficult, as TR will not be able to identify the country code of the underlying issuer. Therefore, it seems to be crucial to define how the access rights for those trades will be granted, i.e. whether it should be granted to all regulators, as it is now for all derivatives trades where the underlying is identified with an “I” or a “B” (according to Q&A TR Answer 37 d) .