

CONVENTION ON INTERNATIONAL TRADE IN ENDANGERED SPECIES
OF WILD FAUNA AND FLORA



Seventieth meeting of the Standing Committee
Rosa Khutor, Sochi (Russian Federation), 1-5 October 2018

CITES TRADE DATABASE: IMPROVING ACCESS TO ANNUAL REPORT DATA

This document has been submitted by the Secretariat and prepared by the United Nations Environment Programme – World Conservation Monitoring Centre (UNEP-WCMC) in relation to agenda item 26.2.*

* *The geographical designations employed in this document do not imply the expression of any opinion whatsoever on the part of the CITES Secretariat (or the United Nations Environment Programme) concerning the legal status of any country, territory, or area, or concerning the delimitation of its frontiers or boundaries. The responsibility for the contents of the document rests exclusively with its author.*

CITES Trade Database: Improving access to annual report data

CITES Trade Database: Improving access to annual report data

Published May 2018

Copyright UNEP-WCMC 2018

Citation UNEP-WCMC, 2018. CITES Trade Database: Improving access to annual report data.



The United Nations Environment World Conservation Monitoring Centre (UNEP-WCMC) is the specialist biodiversity assessment centre of the UN Environment, the world's foremost intergovernmental environmental organisation. The Centre has been in operation for over 30 years, combining scientific research with practical policy advice.

The contents of this report do not necessarily reflect the views or policies of UN Environment, contributory organisations or editors. The designations employed and the presentations of material in this report do not imply the expression of any opinion whatsoever on the part of UN Environment or contributory organisations, editors or publishers concerning the legal status of any country, territory, city area or its authorities, or concerning the delimitation of its frontiers or boundaries or the designation of its name, frontiers or boundaries. The mention of a commercial entity or product in this publication does not imply endorsement by UN Environment.

**UN Environment World Conservation Monitoring Centre
(UNEP-WCMC)**

219 Huntingdon Road,
Cambridge CB3 0DL, UK
Tel: +44 1223 277314
www.unep-wcmc.org

UN Environment promotes environmentally sound practices globally and in its own activities. Printing on paper from environmentally sustainable forests and recycled fibre is encouraged.

1. Context

Under Article VIII of the Convention, Parties are required to provide information regarding their trade in CITES-listed specimens through their annual reports, and the Secretariat makes this information available through the CITES Trade Database (trade.cites.org). This database currently contains over 18 million records of international trade in CITES-listed species. Parties recognise the importance of these reports as a tool for monitoring the implementation of the Convention, assessing the effectiveness of their wildlife management and trade policies, and to enhance the detection of potentially harmful or illicit trade.

In document SC 69 Doc.28.2 the Secretariat put forward a suggestion to provide the CITES Trade Database in a non-aggregated output including all data submitted by Parties in their annual reports, including permit numbers which are currently maintained confidentially in the database. Currently, three permit fields are stored in the CITES Trade Database: export permits, import permits and origin permits. Export permits are a required field in annual reports, whereas import permits are optional.

CITES permit numbers can help facilitate the tracking of shipments over time (e.g. where an export permit matches the origin permit in re-exports) and can enable matching of import and export shipments, which helps to identify discrepancies for further investigation. Permit numbers also provide the opportunity to determine whether a particular item in trade was shipped alongside other items on the same permit (which may help to indicate if certain species are often traded together or cases where parts of an animal, shipped together, are likely to represent the same individual).

At the 69th meeting of the Standing Committee, Parties welcomed the Secretariat's initiative to improve access to annual report data through the provision of an additional output; however, a number of Parties expressed concerns that making permit numbers public could lead to fraud and/or illicit trade. One solution proposed during discussions at SC 69 was to assign confidential permit numbers a unique identifier which could be made publically available, enabling CITES Authorities and researchers to match import and export shipments whilst avoiding confidentiality concerns.

The Standing Committee requested the Secretariat to work with UNEP-WCMC to take into account the concerns expressed by Parties and to report back on this initiative at its 70th meeting. Following discussions with the Secretariat, UNEP-WCMC has investigated the feasibility of replacing confidential permit numbers with unique identifiers for the non-aggregated output, and presents these findings below. Methods for generating such identifiers, technical solutions for implementing the inclusion of the new identifiers within the non-aggregated outputs and key considerations are provided.

2. Proposed approach

Based on researching different potential solutions to this issue, UNEP-WCMC suggests that the most secure and appropriate way forward would be to:

- generate unique identifiers that are **ten alpha/numeric characters long**; and
- generate the identifiers using a **secure, non-reversible cryptographic hash algorithm**¹.

“Cryptographic hash algorithms” are a robust, secure and efficient method of generating randomised identification numbers which are in standard use in industry. Using an identifier that is ten alpha/numeric characters long provides up to 36¹⁰ unique character combinations which is more than ample in the scope of CITES permitting.

The format of CITES permit numbers varies between Parties and as such, the permit number fields in the CITES Trade Database are currently free text with no systematic data cleaning. Permit numbers may include characters such as symbols or punctuation marks which may be reported differently by different trading partners. For example the same permit number could be reported as '12/AB/1234', '12 AB 1234' or '12-AB-1234' etc. To minimise artificial differences between reported permit numbers, as part of the unique identifier generation process, import, export and re-export permit numbers would be converted to alphanumeric strings, and stripped of all punctuation and special characters (e.g. with "-" or "/" between characters).

A unique identifier would then be assigned to each unique permit number, irrespective of whether that permit number was reported as an import, export or re-export permit number. It is particularly important that permit numbers in the 'export' and 'origin' columns are considered together as permits in the 'origin' column are also export permits. The same unique identifier must be applied to a unique permit number regardless of whether it is considered an export permit or an origin permit for a particular shipment. This will ensure that any ability to track permits through the trade chain is not lost.

In order to convert the unique identifier back to the original permit number, a look-up table (containing the permit number and its associated unique identifier) would be generated for future look-ups. When the process is run for the first time, it would generate an initial look-up table. In subsequent runs, any new permit numbers and their unique identifier would be added to the bottom of this table. Therefore permit numbers would always retain the same unique identifier. Without this table, unique identifiers cannot be converted back into permit numbers; this look-up table would be stored securely.

Table 1 in Annex A contains an example of how the data output will be structured. In the next section we present two technical solutions for how the process to generate unique identification numbers could be implemented.

3. Options for technical implementation

¹ The specific cryptographic hash algorithm that would be used is **SHA-512**.

There are two options for the technical implementation of this proposed approach. The first option proposes an extension to the CITES Trade Database infrastructure to accommodate the new identifiers, which would enable for the management of the identifiers in one, integrated system; whereas, Option 2, while less costly, is a more low-tech option involving separate management processes.

Option 1 - Database-generated and stored identifiers

Option 1 would fully integrate the creation and storage of unique identifiers into the framework of the CITES Trade Database. Going forwards, unique identifiers would be generated for permit numbers at the point of upload of a shipment into the CITES Trade Database. The ID generation scripts would also be run on historical records, so that all permit numbers held in the CITES Trade Database would be assigned a unique identifier.

Benefits

Under this option, unique identifiers would be held in the database as additional fields alongside the original permit numbers, which would remain confidential.

- Integrating this step into the current data management and upload process means that the unique identifiers could automatically be included in any whole database download from the CITES Trade Database or other ad-hoc subsets of data which may be requested by Parties.
- Data basing the unique identifiers would mean that they would be searchable in the internal data management platform, which would expedite responses to queries from Parties on specific records based on unique identifiers.
- This option is more robust in that all the data would be held in one coherent system and would not require "ad hoc" scripts to be run each time an output is required.

Limitations

- This option would have budgetary implications, as it would require some development of the CITES Trade Database in order to integrate this process fully.

Estimated budget

The estimated budget for Option 1 is USD 15,000.

Option 2 - Off-line script run each time an output is generated

Option 2 would generate unique identifiers in an offline process, using a script written in the programme R. This script would be run on a download from the CITES Trade Database once a year in conjunction with the production of the non-aggregated output of data. We would suggest that this download takes place annually in April, as this is when, on average, approximately 75% of Parties reports have been received for the most recent year of reporting.

This script would replace the confidential import, export and origin permit numbers with 10-digit alphanumeric unique identifiers; the non-aggregated data output would therefore only contain the unique identifiers and would exclude the reported confidential permit numbers. All other data would remain unchanged. This method could be run on any shipment-level output, and could be readily applied to historic data.

The first time the script is run, it would be run on the complete data download (approx. 18 million records) while subsequently, the script would just be run on additional trade data records uploaded since the last time the script was run. The look-up table, containing the permit number and unique identifier, would be generated for future look-up (along with a date-stamped back-up) and would be stored securely. This would mean that subsequent outputs would always apply the same unique identifiers to the same original permit each time the scripts were run, thereby maintaining continuity between outputs.

Benefits

- As this option does not involve additional development to the CITES Trade Database, it is less resource-intensive and can therefore be implemented within the framework of the existing support contract between the CITES Secretariat and UNEP-WCMC for the maintenance of the CITES Trade Database.

Limitations

- Using offline scripts is more manual than building unique identifiers into the CITES Trade Database framework. As a result scripts would have to be run every time the download was produced and there would be a yearly associated cost (which would not be the case in Option 1).
- Unique identifiers would not be data-based and therefore not easily searchable. Queries from Parties based on unique identifiers would be more time-consuming to investigate through the look-up table.
- Producing trade data outputs in response to ad-hoc requests from Parties would be require the script to be run every time.

Estimated budget

Initial cost to implement solution: USD 2000

Annual on-going cost to process data for the annual output: USD 700

Implementing Option 2 could be accommodated under *Activity II – Support to maintenance and increased accessibility of the CITES Trade Database* of the existing support contract between the CITES Secretariat and UNEP-WCMC assuming the level of funding for this activity is maintained at a level comparable to the 2017-2019 levels.

Annex A

Table 1. Example of current format of shipment by shipment data output from CITES Trade Database, including confidential permit numbers (*hypothetical examples*).

Year	App.	Taxon	Term	Qty	Unit	Imp	Exp	Origin	Purpose	Source	Reporter Type	Import Permit	Export Permit	Origin Permit
2014	II	<i>Alligator mississippiensis</i>	leather products (small)	20		GB	US		T	W	E		US123	
2014	II	<i>Alligator mississippiensis</i>	leather products (small)	20		GB	US		T	W	I		US123	
2015	II	<i>Alligator mississippiensis</i>	leather products (small)	10		AU	GB	US	T	W	E		GB123	US123
2015	II	<i>Alligator mississippiensis</i>	leather products (small)	10		AU	GB	US	T	W	I		GB123	US123

Table 2. Example of proposed format of shipment by shipment data output from CITES Trade Database with confidential permit numbers replaced by unique identifiers (*hypothetical examples*).

Year	App.	Taxon	Term	Qty	Unit	Imp	Exp	Origin	Purpose	Source	Reporter Type	Import Permit unique ID	Export permit Unique ID	Origin Permit Unique ID
2014	II	<i>Alligator mississippiensis</i>	leather products (small)	20		GB	US		T	W	E		50c458f4b4	
2014	II	<i>Alligator mississippiensis</i>	leather products (small)	20		GB	US		T	W	I		50c458f4b4	
2015	II	<i>Alligator mississippiensis</i>	leather products (small)	10		AU	GB	US	T	W	E		a41588f649	50c458f4b4
2015	II	<i>Alligator mississippiensis</i>	leather products (small)	10		AU	GB	US	T	W	I		a41588f649	50c458f4b4