Purpose and Scope

The purpose of this document is to layout the expected outcoming for the data extraction MVP (minimum variable product). Referred to as the "Flora MVP Design Phase".
This document is intended to provide guidance to the vendors on the MVP but is not constrain them is case they want to include addition value added features.

Scope:

Extraction of data from selected Product Safety study reports to the prescribed level of granularity.
(In general, there will be 9 different Product Study report types for each Flora Form).

Not in scope:

All table contents from the AMES Product Safety Study Reports
Selected table content from other Product Safety Study Reports (as indicated in the expected outcome file).
Content from Product Chemistry study Reports required for completing a Flora Form.
The direct population of the Flora Form with the extracted data

Background:

In order to accelerate the approval process for formulated products in Brazil, ANVISA the Brazilian regulatory authority, has requested applicants to populate an Excel template with specific data applicable for each application. The Excel template is known as the "Flora Form".

The Flora Form consists of a series of tabs which need to be completed by the applicant.
Some cells within the Flora form are populated using picklists, some cells are auto populated by referencing other cells within the Flora form, and some cells contain calculations of numeric data from other populated cells.

The Flora form is password protected by the ANVISA. Therefore, we do not intend to populate the Flora Form directly at this stage. Our approach is to first confirm that we can extract the required data to the required level of granularity and then approach ANVISA to investigate if we can submit the completed Flora Forms in xml, JSON formats. Alternatively, if ANVISA insists on the Excel format, then we will discuss options with ANVISA for auto-populating the Flora Form in Excel format.

Requirements for the Flora MVP Design Phase

Participants should extract as much of the required data as possible to the required level of granularity as possible. Extracting too much data is considered better than not extracting enough data. This is because it is easier to delete unwanted data than search for missing data.

The data extract output can be supplies as a single file for the entire Flora Form or as a single file per Study Report. However, both options would be better.

Output format should be available in xml and JSON formats.
The initial Flora Forms will be populated using copy & past. Therefore, please support this by providing a scheme for the xml file or by providing a copy & past option (such as in Excel).

Syngenta will provide participants with the following:

- "Expected outcome file" which includes guidance on where to find the required content to be extracted.
- All relevant studies used to complete content which is in scope.
  That is: One clean set of studies and one marked up set of studies indicating the expected data extractions.
- A list of look up values and known anomalies between study reports layouts and text, known to date.
- Up to 9 Studies X 20 formulation =180 working documents.

Participants should aim to extract as much of the required data as possible.

In the cases where the data is intended to populate one of the tables in the Flora Form, please present the data so it is available for copy & pasting, as far as possible.

Expected Out Come

At the end of the MVP, the vendors will provide the data extracts for the 20 data sets in the required formats.
The vendor will provide access to the tool to selected Syngenta users, who will process a further 2 data sets of up to 9 studies each. The users will then be able evaluated the usability of the tool as described it he next section.

Usability of Solution

After initial setup of specific document types in the system, users should be largely autonomous when using the solution.

Ideally the proposed solution should allow users to:

- Upload document as required (for any document types which are set up in the solution).
- Easily validate the output against the study contents.
- Easily make corrections. To include missing data or to exclude unwanted data
- Work efficiently without system delays.
- Export the data extraction in the required output format.

APPENDIX 1 List of Product Safety Studies in Scope

- There are potentially 14 different Study types
- There are 8 Rule sets to apply
- In case a study does not have a specific rule set, apply the In-vitro Rule and apply the OECD number as indicated in the study report.

Study - OECD Guideline	**Required Format (from Flora pick list)**	**Rules to Apply**
OECD 402 (2017)	Nº 402: Acute Dermal Toxicity (09/10/2017)	Study specific rule
OECD 404 (2015)	Nº 404: Acute Dermal Irritation/Corrosion (28/07/2015)	Study specific rule
OECD 405 (2017)	Nº 405: Acute Eye Irritation/Corrosion (09/10/2017)	Study specific rule
OECD 425 (2008)	Nº 425: Acute oral Toxicity - Up-and-Down Procedure (03/10/2008)	Study specific rule
OECD 429 (2010)	Nº 429: Skin Sensitisation (23/07/2010)	Study specific rule
OECD 471 (1997)	Nº 471: Bacterial Reverse Mutation Test (21/07/1997)	Study specific rule

Study Guideline	**Required Format (from Flora pick list)**	**Rules to Apply**
OECD 403 (1981)	Nº 403: Acute Inhalation Toxicity (12/05/1981)	403 Rule
OECD 436 (2009)	Nº 436: Acute Inhalation Toxicity – Acute Toxic Class Method (08/09/2009)	403 Rule

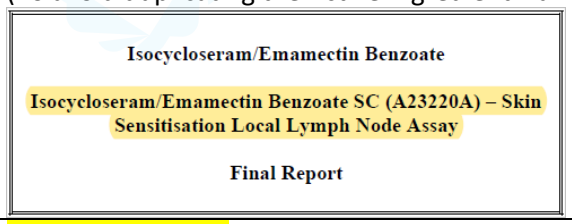
Study Guideline	**Required Format (from Flora pick list)**	**Rules to Apply**
OECD 406 (1992)	Nº 406: Skin Sensitisation (1992)	In-vitro Rule
OECD 428 (2004)	Nº 428: Split-Thickness Skin test (2004)	In-vitro Rule
OECD 438 (2018)	Nº 438: Eye Irritation (26/06/2018)	In-vitro Rule
OECD 439 (2019)	Nº 439: Skin Irritation (2019)	In-vitro Rule
OECD 474 (2016)	Nº 474: Micronucleus Bone Marrow Cells Rat (2016)	In-vitro Rule
OECD 487 (2016)	Nº 487: Micronucleus Human Lymphocytes (2016)	In-vitro Rule

Guideline variables

The OECD guideline numbers may be expressed in various ways (see examples below)

Study Guideline
OECD 402 (2017)
OECD No. 402 (2017)
Procedure 402
OECD, Method 403
OECD [Test Guideline, Number 402]

APPENDIX 2 Clarifications and Corrections to Extraction Requirements

Clarifications of Requirements	Response
Watermarked Source file	Many of the source files have watermarks. Non-watermarked source files an be provided. However, it would be better if the solution could manage watermarked files.
Output File Format?	Provide in Excel or xml or JSON xml and JSON will be required for the future process, so please demonstrate this in the MVP
Output layout as a single table format?	Provide all study report data a single file, if possible. First study followed by the second by the next etc. In numerical order. The user must be able to easily copy and paste the extracted data into an Excel file. If provided in xml output, please provide an xml viewer (to allow Copy and Paste to Excel)
Labels to be used in Output file	Use English labels only pre-pend the OECD report number as follows: 425-Study Title 425-Performing laboratory 425-Report Number
Date Format?	Number format as follows: dd/mm/yyyy
Text Groupings For In vitro Studies	For invitro studies, OECD 406, 428, 438, 439, 471 Take two blocks of test as follows: One for Study Design One for Results and Conclusion
Text Grouping for all other studies (general guidance)	Always bring “Conclusion(s)” as a single block of Text Bring Deviation from guideline as a numbered line per sentence. Refer to comments in Expected Extraction Results for other text.
Differences in highlighted text in the marked-up study versus the expected text in Excel file?	Study has too much mark-up which needs to be removed in the marked-up Study PDFs (this was due to reduction in scope). This is now corrected
Differences in wording in the marked-up study text versus the expected text in Excel file?	Always take the text/values in Study Report. (This due to so manual intervention in the original process and translate between English to Portuguese then Portuguese back to English).
Required Title Text? Can we include line one in the title box? (which includes the Active Ingredient names).	No, only take the title text. (To avoid duplicating the Active Ingredient Names)  Isocycloseram/Emamectin Benzoate Isocycloseram/Emamectin Benzoate SC (A23220A) – Skin Sensitisation Local Lymph Node Assay Final Report
How to consistently find Species in OECD 471.10	No longer in scope
How can we consistently find “Washing of treated areas” 404.14	No longer in scope
How can we consistently find “Was reliability test performed with positive control?” Component 429.13	No longer in scope
Where do we find the Table 1 contents for 403-436	No longer in scope

Appendix 3 Further Clarifications following initial analysis

Please find below some Flora data extraction updates on “Special cases” and corrections for consistency.

1. **Guideline numbers:**
   In some cases the guideline number listed on the cover page does not contain a date.
   Example: OECD 403.
   Therefore the correct guideline number can not be looked up as there are multiple years for OECD 403.
   OECD 403 (2009) and OECD 403 (1981)
   Therefore, in such cases bring please Exact text e.g. “OECD 403” if possible.

2. **404 and 405 Details of Reported changes (inconsistently of text presentation)**
   Please bring as a single block of text for both cases:

   404 “Details of reported changes” is a single block of and is correct:

404-Detailing of reported changes(?)	The primary irritation index was 0.00. No local dermal signs were observed in the treated animals throughout the study. No clinical signs of systemic toxicity were observed in the animals during the study, and no mortality occurred. As no clinical signs were observed at 72 hours after patch removal, the study was terminated after 72 hours of observation of the second and third rabbits. The body weights of all rabbits were considered to be within the normal range of variability.

   405 “Details of reported changes” is a line per sentence which is <u>not</u> correct: (Please bring as single block of text as above)

405-Detailing of reported changes 1	The second and third animals were symptom-free 2 days after treatment, so the study was terminated after the observation period of 3 days (72 hours).
405-Detailing of reported changes 2	Fluorescein staining was negative in eyes treated with a test item during the study.
405-Detailing of reported changes 3	The control eye of each animal was symptom-free during the study, fluorescein staining was negative during the study.
405-Detailing of reported changes 4	No mortality occurred during the study and no clinical signs of systemic toxicity were observed in any animal in this study.
405-Detailing of reported changes 5	The bodyweights were considered to be within the normal range of variability in the animals.

   This is corrected A23220A - Expected Extraction Results (Version 09.11.2022).xlsx

3. **429 “Preliminary Test Results” Line 1 was accidently deleted**

429-Study Design (Main Study)	In the study the test substance Isocycloseram/Emamectin Benzoate SC (A23220A) formulated in 1% aqueous Pluronic® was assessed for its possible skin sensitising potential. For this purpose a local lymph node assay was performed using test substance concentrations of 5, 10, and 25% (w/w). The highest concentration tested was the highest
429-Results (Main Study)	The animals did not show any signs of systemic toxicity during the course of the study and no cases of mortality were observed. The animals showed a very slight erythema of the ear skin on test day 3 only (Score 1). In this study Stimulation Indices (S.I.) of 0.8, 1.3, and 1.6 were determined with the test substance at concentrations of 5, 10, and 25% in 1% aqueous Pluronic®, respectively.
429-Preliminary test results 2	At the tested concentrations the animals showed an erythema of the ear skin (score 1 to 2). Signs of systemic toxicity included increased activity, piloerection, fur loss, hunched posture, partially closed eyes, decreased activity, elevated tail, tremor, and fasciculations.
429-Preliminary test results 3	The animal treated with 100% test item concentration was euthanised on day 4 due to deterioration of clinical symptoms (tremor, tippy toe walk, moribund appearance, substantial body weight loss).
429-Preliminary test results 4	Therefore, a second pre-experiment was performed using concentrations of 10 and 25%.
429-Preliminary test results 5	At the tested concentrations the animals did not show any signs of systemic toxicity. At the tested concentration of 25% the animal showed a very slight erythema of the ear skin (score 1).
429-Preliminary test results 6	The animal treated with 10% test item concentration did not show any signs of local irritation.

   This is corrected A23220A - Expected Extraction Results (Version 09.11.2022).xlsx

4. **Non- English Document**
   Please ignore any non-English documents as they are not in scope of this MVP.

5. **Testing sets - Missing and Duplicate files.**
   Please delete the 3 training set folders highlighted below.
   They will be replaced with new folders and files in the coming days.