multi-act

Deliverable D2.2

# MULTI-ACT content and data collection manual

*"It will specify the data and content collection procedures and the annotation of content based on common expert-generated vocabularies. In this way, the relevant experts and stakeholders will continuously update the knowledge base of the project with the content that they collect."*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No 787570

| | |
|---|---|
| **PROJECT ACRONYM:** | A Collective Research Impact Framework and multi-variate models to foster the true engagement of actors and stakeholders in Health Research and Innovation |
| **CONTRACT NUMBER:** | 787570 |
| **DISSEMINATION LEVEL:** | Public |
| **NATURE OF DOCUMENT:** | Report |

| | |
|---|---|
| **TITLE OF DOCUMENT:** | MULTI-ACT Content and data collection manual. |
| **REFERENCE NUMBER:** | D2.2 |
| **WORKPACKAGE CONTRIBUTING TO THE DOCUMENT:** | WP2 |
| **VERSION:** | V0.4 |
| **EXPECTED DELIVERY DATE:** | 29/02/2020 |
| **DATE:** | 05/03/2020 |
| **AUTHORS (name and organization):** | Sofia Tsekeridou (Intrasoft International S.A.)<br>George Tsakirakis (Intrasoft International S.A.) |

Short description of the Deliverable (as in the DoA):

This document details the implementation of the acquired knowledge regarding the MULTI-ACT project, and more precisely the content collection, modelling and integration processes required to provide the MULTI-ACT toolbox with relevant, attractive and validated Collective Research Impact Framework related content.

| REVISION HISTORY | | | |
|---|---|---|---|
| **REVISION** | **DATE** | **COMMENTS** | **AUTHOR (NAME AND ORGANISATION)** |
| V0.1 | 10/02/2020 | Table of Contents and partners assignments for contribution | Sofia Tsekeridou (INTRASOFT International S.A)<br>George Tsakirakis (INTRASOFT International S.A) |
| V0.2 | 27/02/2020 | First Deliverable draft | Sofia Tsekeridou (INTRASOFT International S.A)<br>George Tsakirakis (INTRASOFT International S.A)<br>Paola Zaratin (FISM)<br>Valentina Tageo (FISM)<br>Nicolas Garcia Torea (UBU)<br>Natalia Matuszak (DiA) |
| V0.3 | 29/02/2020 | Internal review comments consideration | Marcin Kautsch (DiA)<br>Françoise Van Hemelryck (EBC) |
| V0.4 | 29/02/2020 | Final version for formal submission to EC | Sofia Tsekeridou (INTRASOFT International S.A)<br>George Tsakirakis (INTRASOFT International S.A) |
| V0.5 | 05/03/2020 | Editing and submission | Valentina Tageo (FISM) |

**FILENAME:** MULTI-ACT_D2.2_INTRASOFT_20200305_v0.5

**STATEMENT OF ORIGINALITY:**

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

## TABLE OF CONTENT

## TABLE OF FIGURES

# EXECUTIVE SUMMARY

Based on the outcomes of the Collective Research Impact Framework (CRIF) that was developed in WP3 and WP5, the initial version of the MULTI-ACT Toolbox v1 was created and its initial testing phase was commenced. This enabled further identification of all foreseen content that exceeds the strictly required data for the Toolbox use and core functionalities and aided to define the necessary mechanisms that will ensure the proper classification and annotation of all anticipated data and content.

These added functionalities will facilitate the ongoing enrichment of the Toolbox and the overall propagation of the Multiple Sclerosis (MS) Research Impact database, to later account for other brain diseases, guaranteeing the continuous enhancement of the Toolbox and its value in collecting valuable sources of information in its life course and making them available among its users.

The accumulation of all content will be done in parallel with the continuous updating of the expert vocabularies used for its annotation, ensuring that the users and stakeholders will benefit from the ongoing enrichment by the acquired knowledge.

# 1 INTRODUCTION

## 1.1 Purpose of this document

The purpose of this document is to present the overall data & content collection requirements and mechanisms related to MULTI-ACT Toolbox core functionalities as well as the organization of the content and data collection processes that will enable the accumulation of relevant Collective Research Impact Framework related structured content.

Furthermore, it provides the modelling and integration activities that were perceived, leading to the development of specific user roles and functionalities necessary for the collection and administration of all relevant data. This entails, apart from the procedures already identified in D2.1 Methodology & Design principles of the MULTI-ACT Toolbox deliverable, new sources and methods of data acquisition that will lead to creation and enrichment of a MS Research Impact database.

Finally, it analyses the manner in which all collected content is annotated based on the expert-generated vocabularies developed in the MULTI-ACT project which will ensure the proper classification of all the accumulated content so as to be validated and` usable in the MULTI-ACT Toolbox.

## 1.2 Structure of document

Section 2 describes the data requirements and the relevant content collection functions that were implemented to support the core functionalities of the MULTI-ACT Toolbox v1. That includes all user data, the initiatives' collected information as well as all data related to the impact assessment indicators.

Section 3 describes the additional data collection mechanisms and functions that were defined and are currently being developed in order to collect additional Collective Research Impact Framework related content.

Section 4 portrays the enriched Taxonomy that was defined to enable the annotation of all collected content as well as the required user roles and functionalities that were implemented in the Toolbox in order to administer the Taxonomy enrichment and subsequent use.

## 1.3 Acronyms and Abbreviations

| | |
|---|---|
| CMS | Content Management System |
| CRIF | Collective Research Impact Framework |
| EBC | European Brain Council |
| EY | Ernst & Young |
| FISM | Fondazione Italiana Sclerosi Multipla FISM Onlus |
| GUI | Graphical User Interface |
| MS | Multiple Sclerosis |
| RFPO | Research Funding and Performing Organizations |

| | |
|---|---|
| RRI | Responsible Research and Innovation |
| DiA | Dane i Analizy |
| UBU | University of Burgos |
| URL | Uniform Resource Locator |

# 2 MULTI-ACT Toolbox Data Requirements and Content Collection Functions

The data required for the implemented functionalities of the MULTI-ACT Toolbox v1 is hereby described.

## 2.1 User-Promoter

Upon registering, a new user, acting as the promoter of a multi-stakeholder research initiative provides only basic information (Figure 1 – User-Promoter Information) that consists of:

- Name
- Surname
- Email address

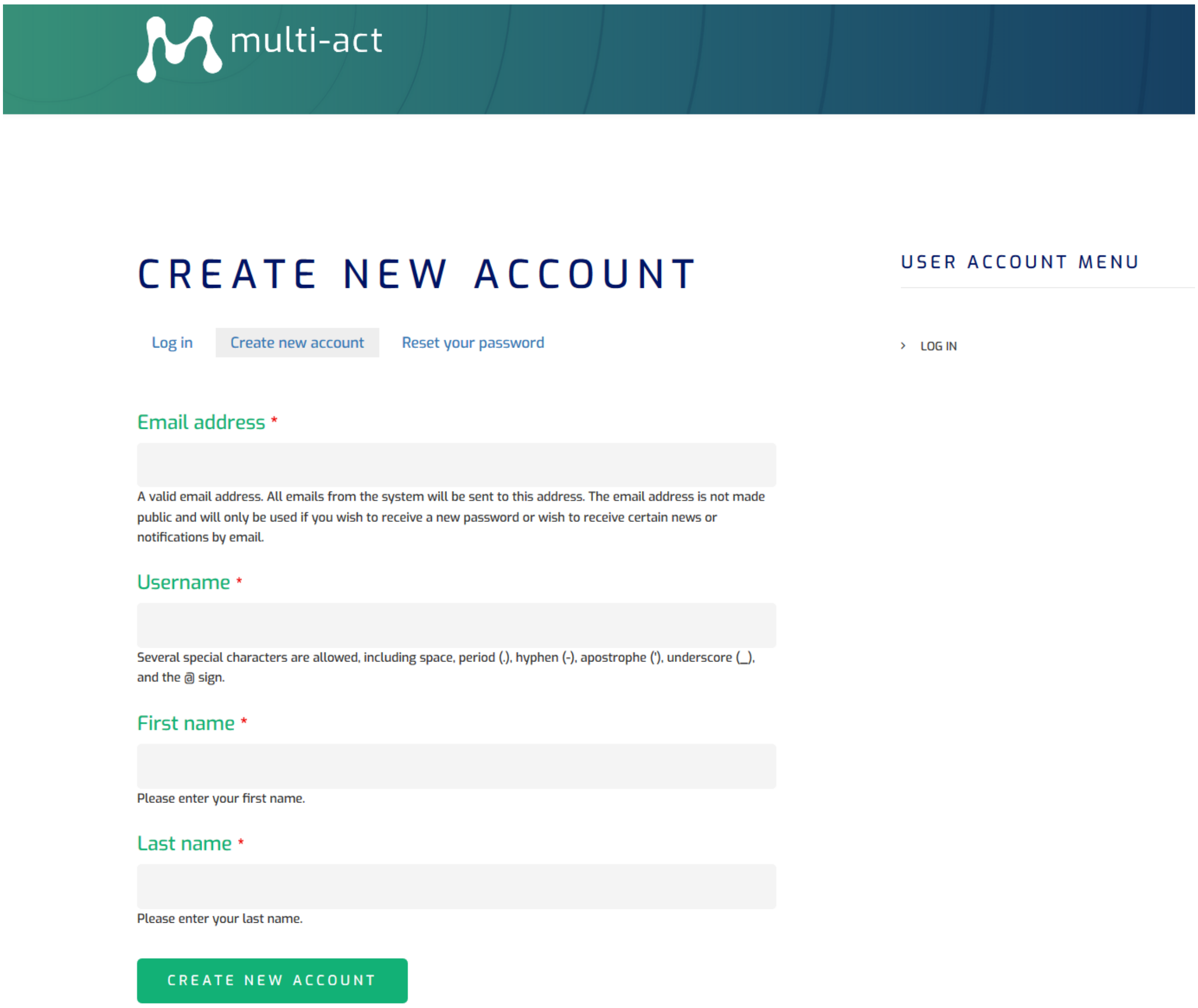



*Figure 1 – User-Promoter Information*

This minimum information is required to ensure the functionality of password recovery and can be updated at any point by the user using the profile editing option. (Figure 2 – User Information Edit)

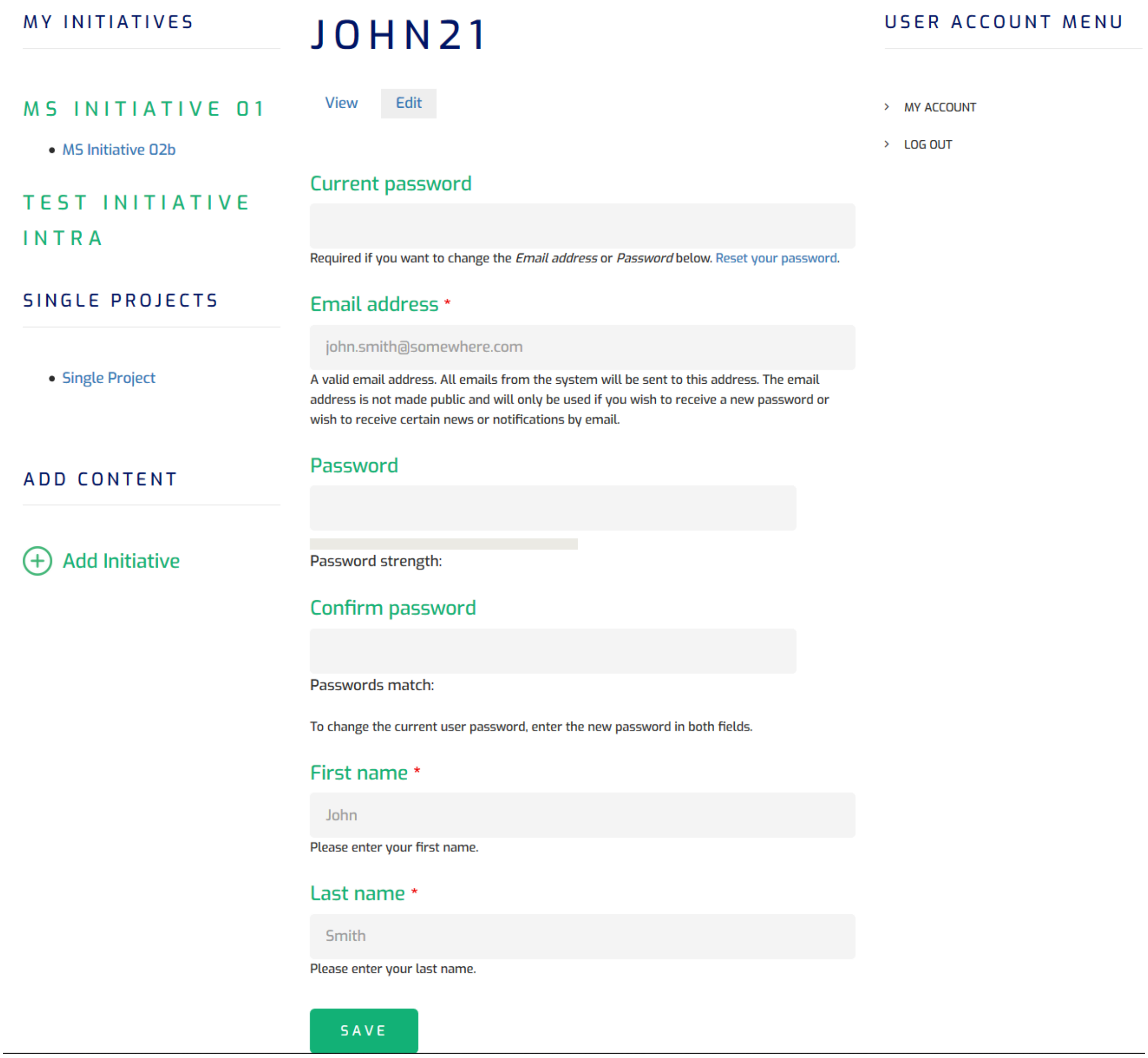



*Figure 2 – User Information Edit*

## 2.2 Initiatives

The Initiative's information collected upon its creation (Figure 3 – Initiative Information) consists of:

- Title
- Initiative Type
- Promoter's Stakeholder Type
    - o the different stakeholder types are the same as the ones illustrated under "Stakeholder Member Types" in Figure 3 – Initiative Information.
- Starting Year
- Number of Organizations involved
- Core Activities

- Stakeholder member types
- Geographical Coverage
- Parent Initiative
  - If a parent Programme initiative has completed the assessment process, for every child Project initiative the user is asked if the new Project initiative inherits/shares the same features as the parent Programme. If the user confirms the inheritance, the child Project does not have to undergo the Assessment process and subsequently shares the assessment status & score of its parent Programme.

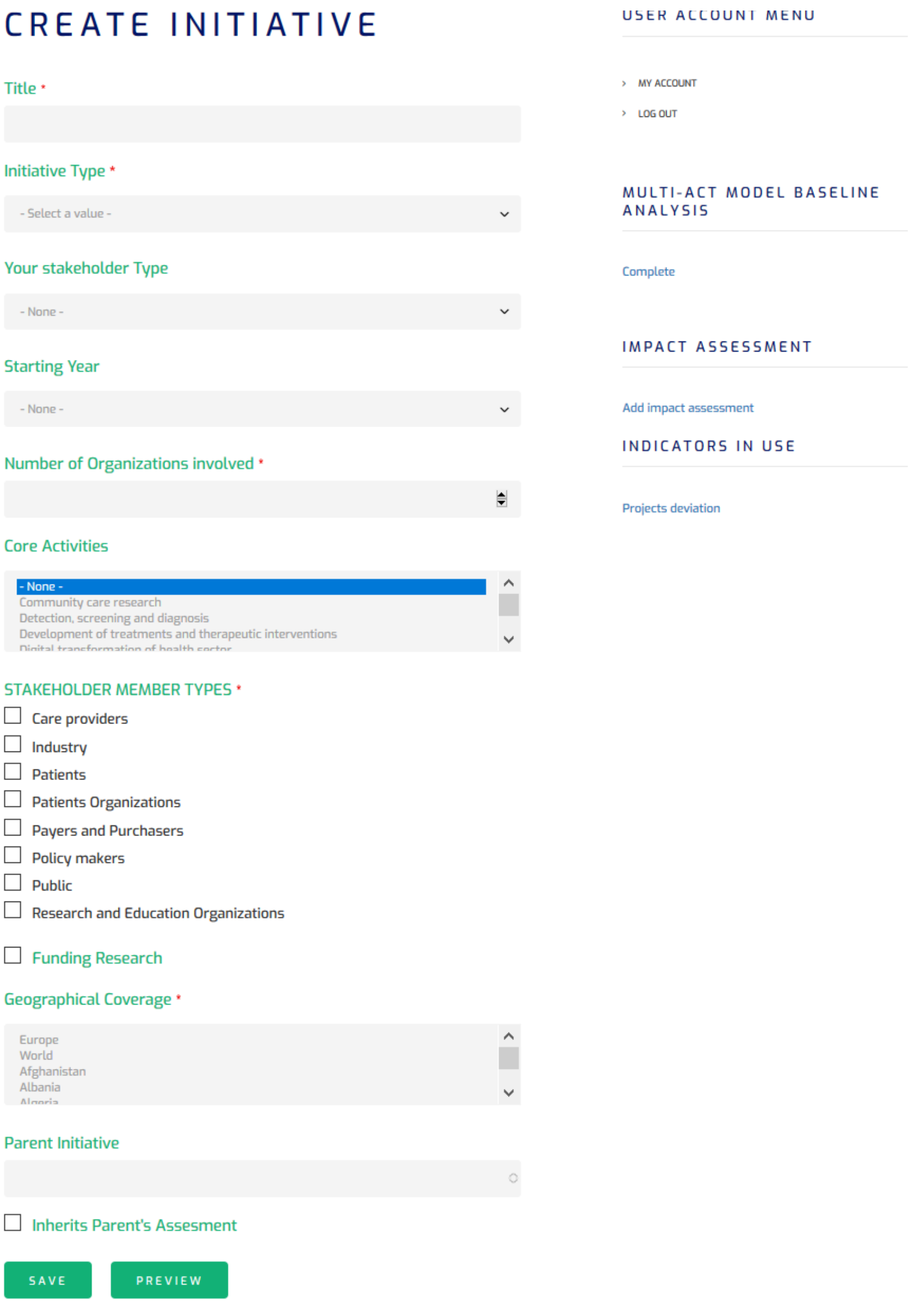



*Figure 3 – Initiative Information*

## 2.3 Indicators

The impact assessment of an initiative is facilitated with the use of the Master Scorecard developed in WP3. The Master Scorecard is a catalogue that provides a long list of indicators assessing the impact of health RRI initiatives considering key aspects of the CRIF dimensions: efficacy, economic, excellence social and patient-reported dimensions.

### 2.3.1 *Indicator Automatic Importer*

Due to the massive content of the Master Scorecard and its subsequent effort in the relevant data entry, an importing mechanism was developed that diminishes the human effort that would be otherwise required.

This importing mechanism parses the Master Scorecard input that was provided in an Excel form and automatically populates all relevant fields, rapidly carrying out complex actions that would require immense time and human effort, while eliminating the probable expected human errors that would be otherwise anticipated.

A sample video to show the output of this functionality is hosted at the Toolbox hosting server.

### 2.3.2 *Indicator Editing – Author Role*

The continuous enrichment of the Master Scorecard indicators as well as the general need of editing their content demanded for a new user role to be defined to be able to facilitate this need.

This role was then appointed to Academic members of the MULTI-ACT Consortium empowering them in continuous enhancement and validation of the indicators content.

The users bearing the role of Author can search for a given indicator via:

- Search function (Figure 4 – Indicator Search A). The keyword-based search is performed into all text fields of the Master Scorecard.
- Dropdown Menus based on the Indicator’s CRIF Taxonomy (Scorecard) reference. (Figure 5 – Indicator Search B)

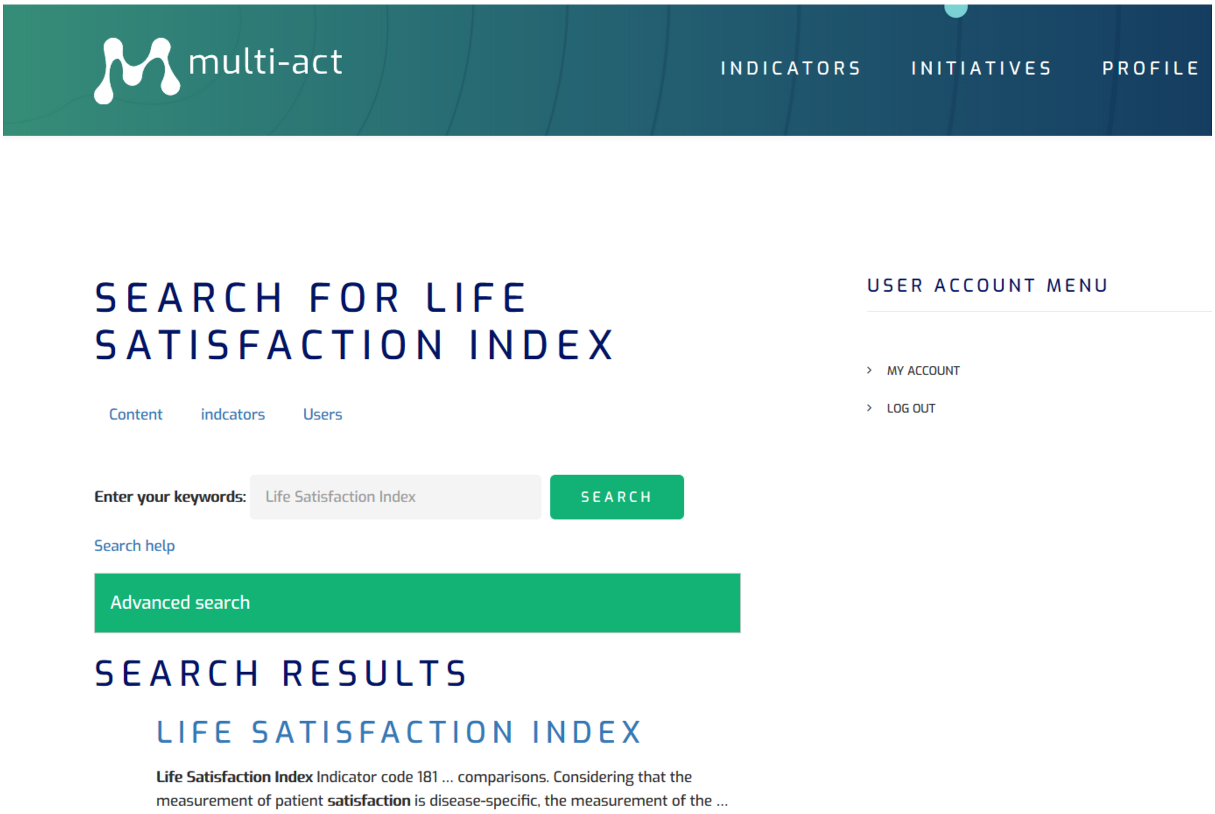



*Figure 4 – Indicator Search A*

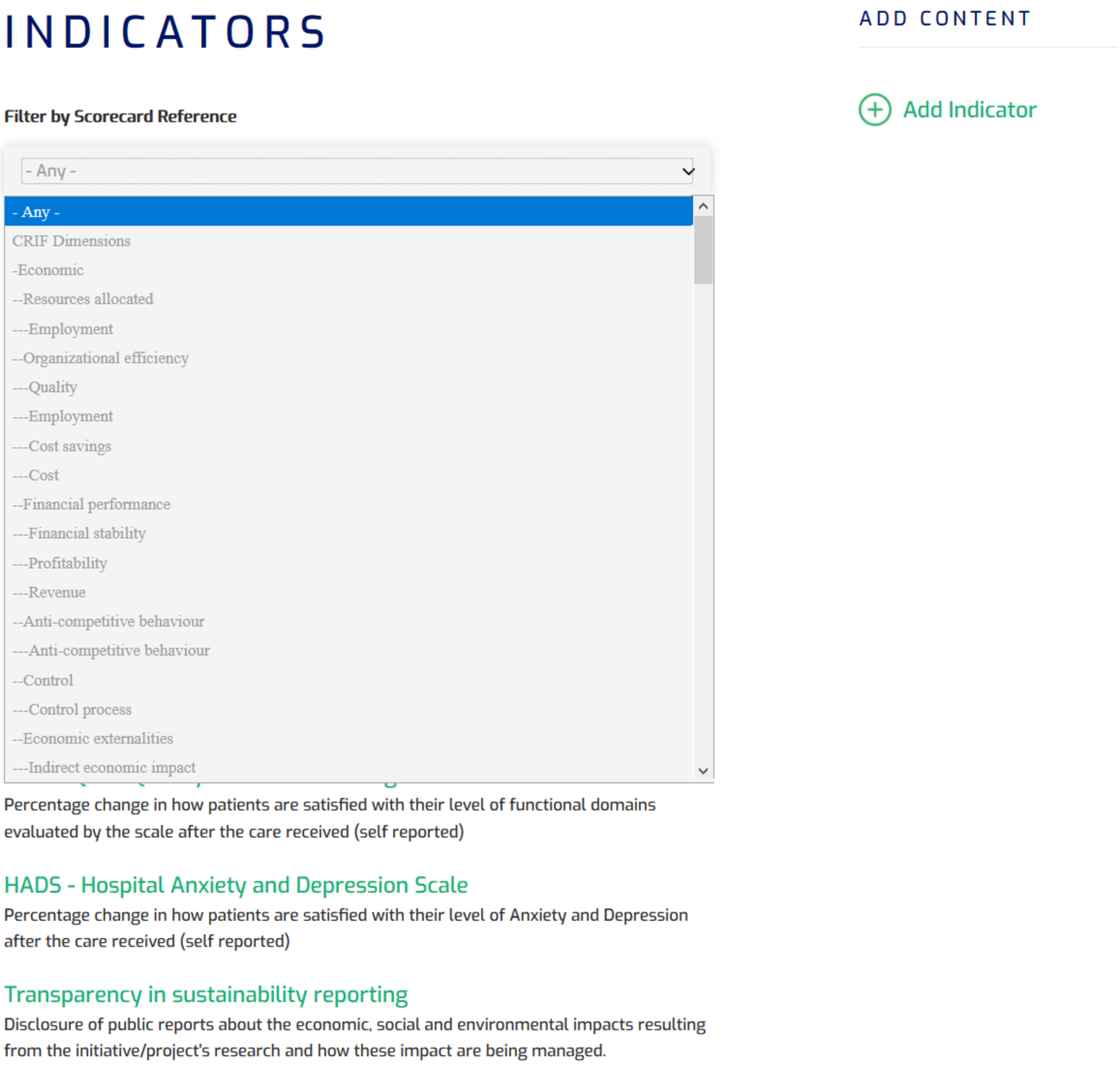



*Figure 5 – Indicator Search B*

Upon selecting an indicator, the users are able to update its content trough a GUI that offers rich HTML capabilities (Spelling Correction, Formatting, etc.) (Figure 6 – Indicator Editing).

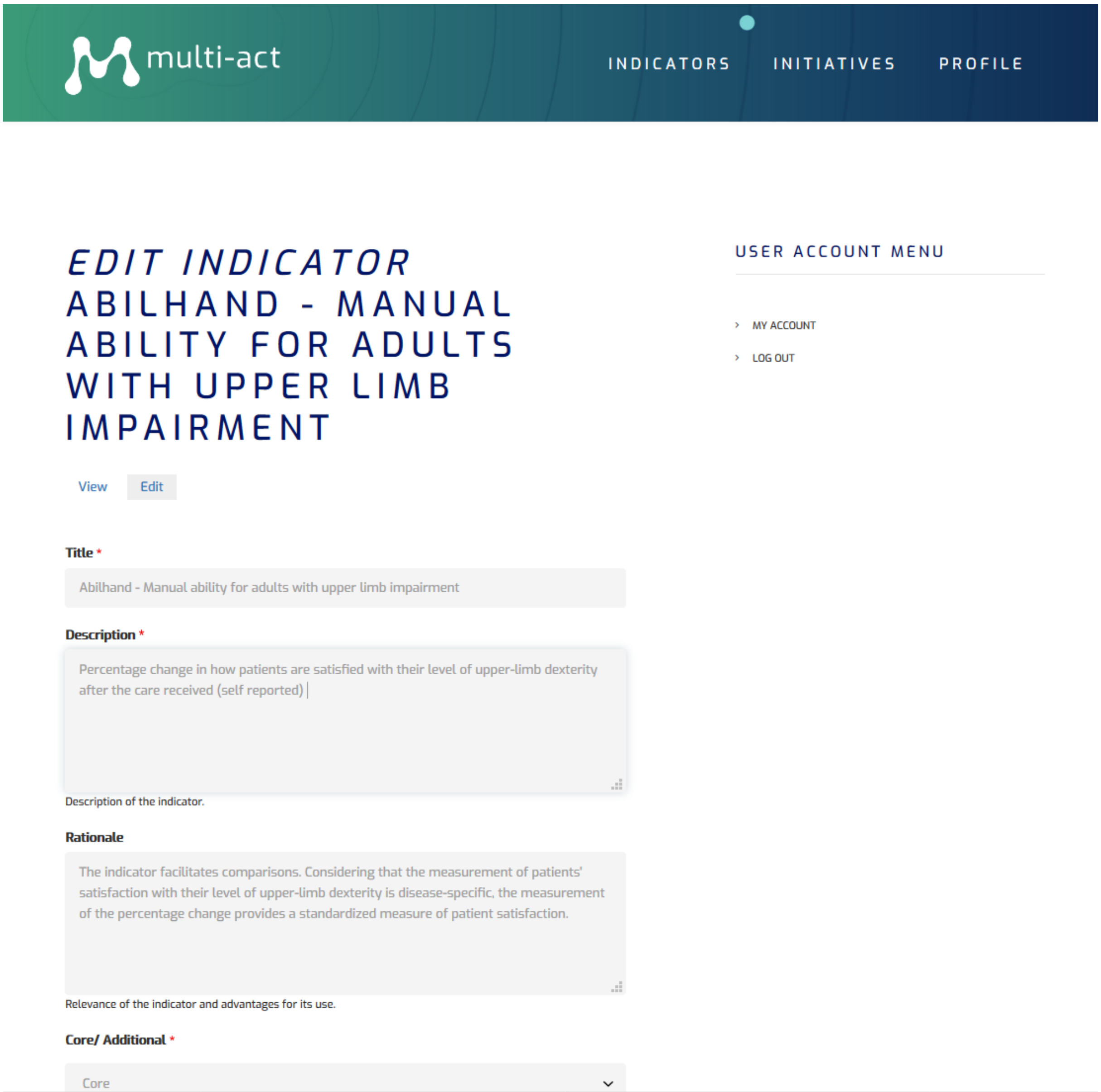



*Figure 6 – Indicator Editing*

# 3 MULTI-ACT Toolbox Additional Data and Content Collection Functions

## 3.1 Baseline Analysis Uploaded Content

Prior to a user being able to undergo the Impact Assessment of their initiative, they must complete the Baseline Analysis Questionnaire which assesses the level of adherence of the initiative with MULTI-ACT Framework's Governance Model.

During this process a multitude of information is collected, addressing the initiative's governance, vision, management etc.). The questionnaire also entails several complex functionalities aiming to make the process as user friendly as possible i.e.

- Dynamic content based on previous answers given.
- Pagination and completion status/percentage.
- Draft submissions so that a user can return to complete the questionnaire at a future point in time.
- Etc.

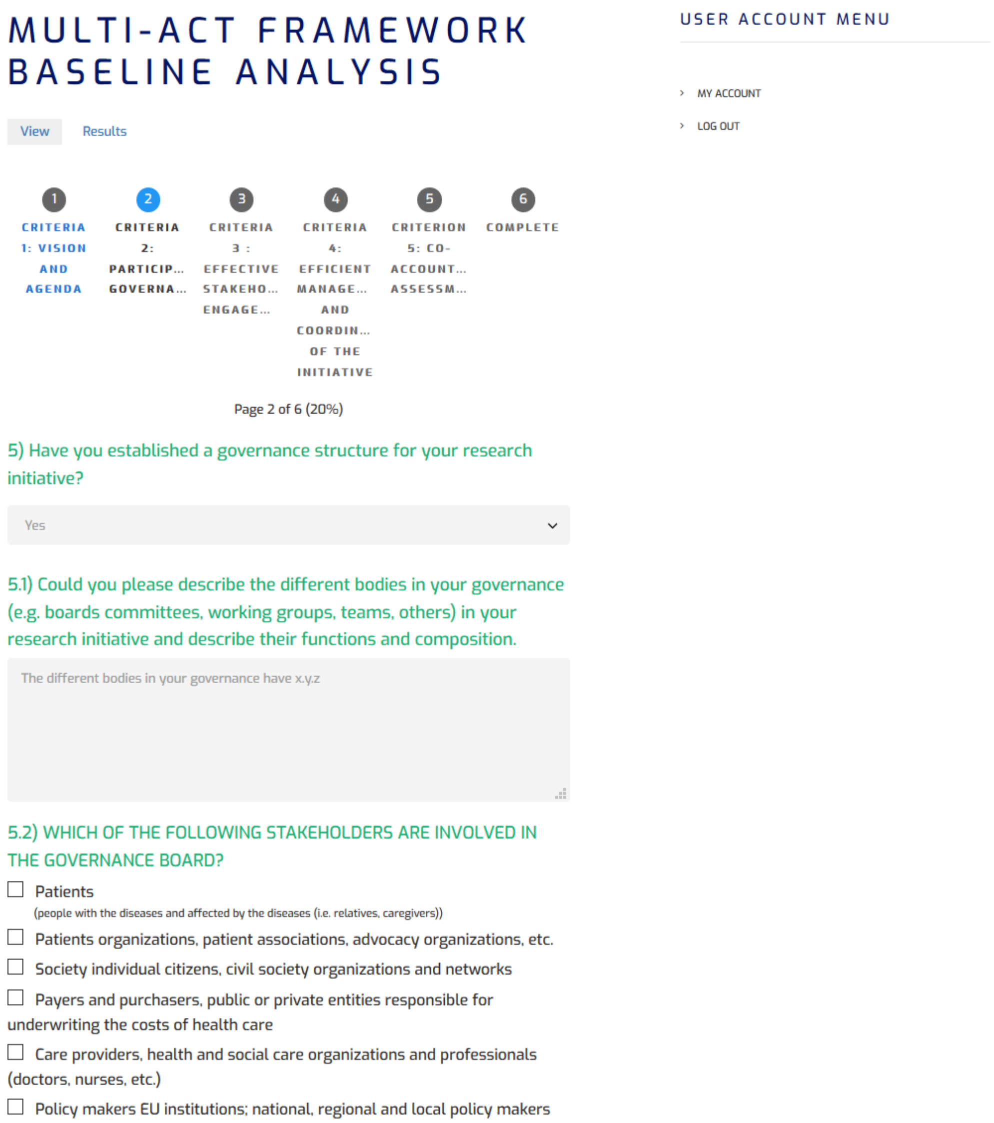



*Figure 7 – Baseline Analysis Questionnaire*

Furthermore, a document collection functionality was implemented in the Baseline Analysis Questionnaire to assist with document uploading for a variety of supporting documents (Mission Statements, Initiative's Agendas, Documentation of Initiative's procedures and mechanisms).

The user is able to upload the requested documentation, while defining multiple rules regarding acceptable file attributes (number of files, file size, file types etc.) (Figure 8 – Baseline Analysis Content Uploads)

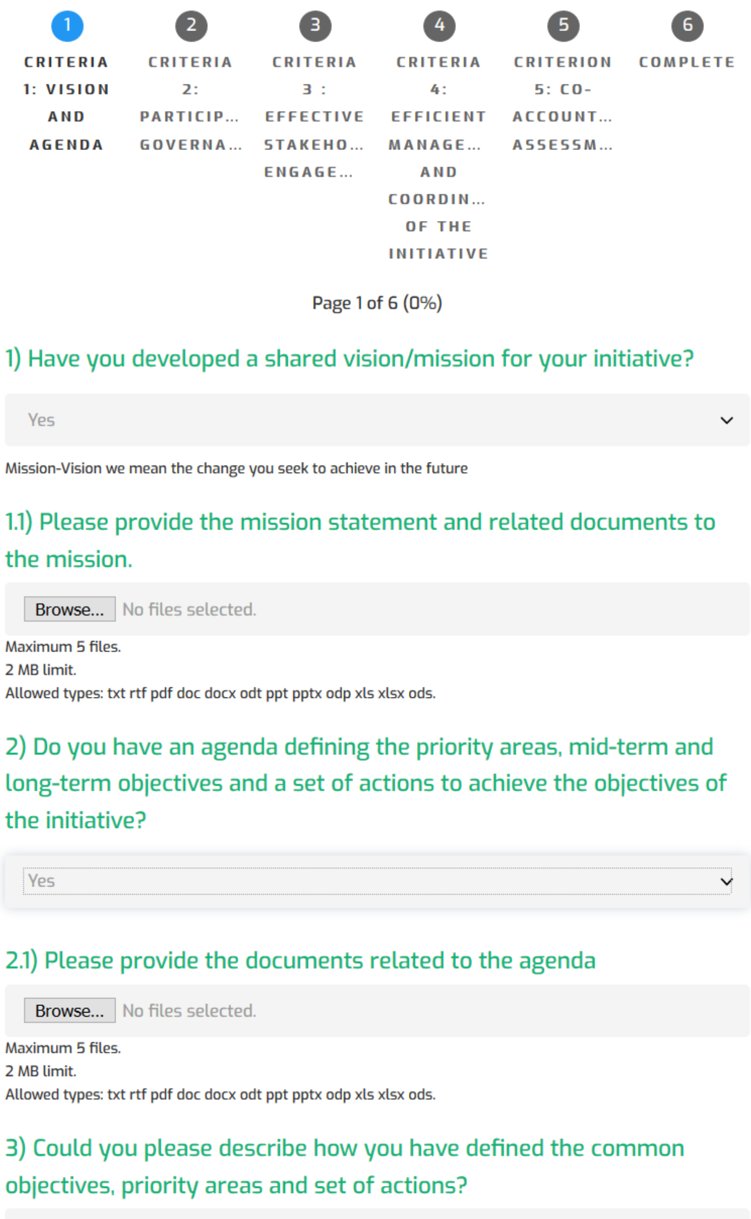



*Figure 8 – Baseline Analysis Content Uploads*

The accumulation of the Baseline Analysis results and documents provided by all the Toolbox users-initiatives, provides a major source of input of information for the population of the MULTI-ACT MS database.

## 3.2 Impact Assessment Support Documents

An additional source of user provided content that is currently being developed is document uploading as part of the Impact Assessment process. The users will be able to upload documents per selected indicator, properly annotating them with the use of the MULTI-ACT

Updated Taxonomy for Uploaded Content Annotation Use. The implemented content uploading and annotation functionalities will be documented in Deliverable D2.4 – MULTI-ACT Toolbox, v2.

## 3.3 Initiative Surveys

With the provided insight of the MS Care Unit that has commenced using the MULTI-ACT Toolbox, the specific need of an initiative to collect information from their stakeholders was revealed. The surveys present the first activity of the Research Initiative in order to collect vital inputs from their stakeholders, which are necessary for the governance of the research initiative as per the MULTI-ACT governance model. This presents the feasible setting for collecting and annotating content and data for the initiative's purposes, and subsequently the MULTI-ACT goals.

To facilitate such need, three surveys of the MS Care Unit were implemented as part of the Toolbox functionalities, associated with the respective initiative "digital space" in the Toolbox, and are awaiting their finalization in order to be distributed among the initiative's stakeholders. Through this process, crucial knowledge is to be gained as to the future needs of relevant research initiatives and their promoters in terms of required data for the governance of the research initiative as per the MULTI-ACT defined governance model.

The three surveys are to be addressed to Academic and non-academic Healthcare Units as well as to Practicing Neurologists and will enable the collection of various clinical data. (Figure 9 – MS Care Unit Survey)


*Figure 9 – MS Care Unit Survey*

The current designed implementation is that the initiative will initially populate their list of stakeholders in the Toolbox, so that it can then select the recipients that will undertake a survey.

Afterwards a unique survey web link (URL) will automatically be created for every stakeholder thus ensuring the valid identity of the participants enforcing the one-vote per stakeholder rule.

## 3.4 MULTI-ACT Guidelines

The outcome of the Baseline Analysis Questionnaire will result to a scoring assessment that will portray the initiative's adherence with the MULTI-ACT vision. Additionally, for every criterion or in some cases specific sets of questions per criterion that a passing score has not been achieved, specific guidelines will be given on how to achieve better adherence.

The Governance guidelines were delivered as the outcome of WP5 and are documented in Deliverable D5.4 *MULTI-ACT model for collaborative initiatives*, whereas the Patient Engagement Guidelines are documented in Deliverable D1.6 *Final version of the MULTI-ACT Patient Engagement in Health R&I guidelines.* The Toolbox will eventually present these guidelines in a user-friendly, contextualized and accessible manner, respecting at the same time any information that may be confidential or sensitive. The implementation of the guidelines as part of the Toolbox functionalities will be presented in D2.4 – MULTI-ACT Toolbox, v2.

# 4 Content Annotation Processes

## 4.1 Taxonomy Enrichment – Admin Role

In order to facilitate the need of the enrichment and overall editing of the MULTI-ACT Taxonomy after its initial definition in D2.1, the core Taxonomy module was installed on Drupal[1]. This enables a user holding the role of Admin to administer all MULTI-ACT taxonomies by organising entities in categories and subcategories in a tree form. This has already been implemented in the *Master Scorecard* Taxonomy (Figure 10 – Master Scorecard Taxonomy)

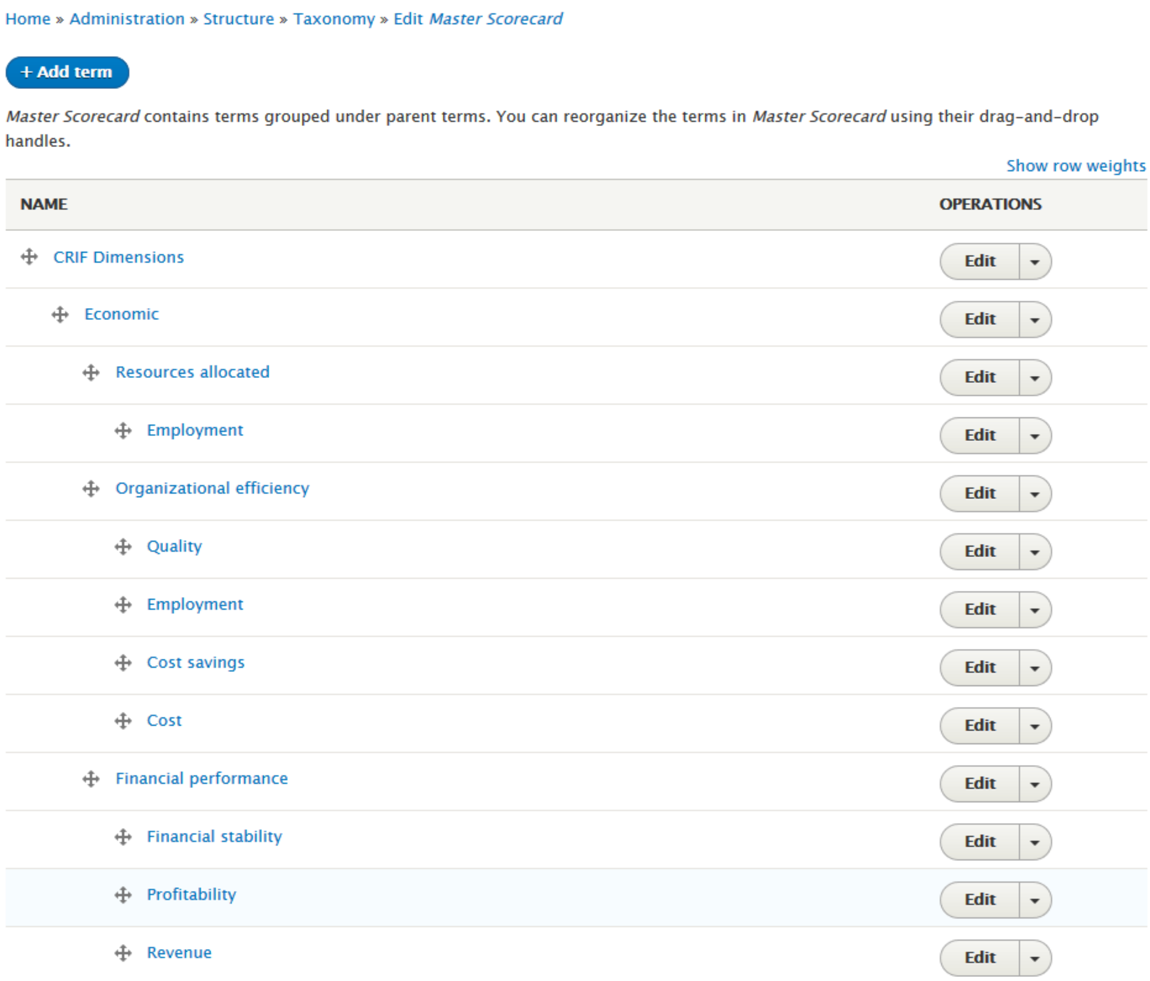


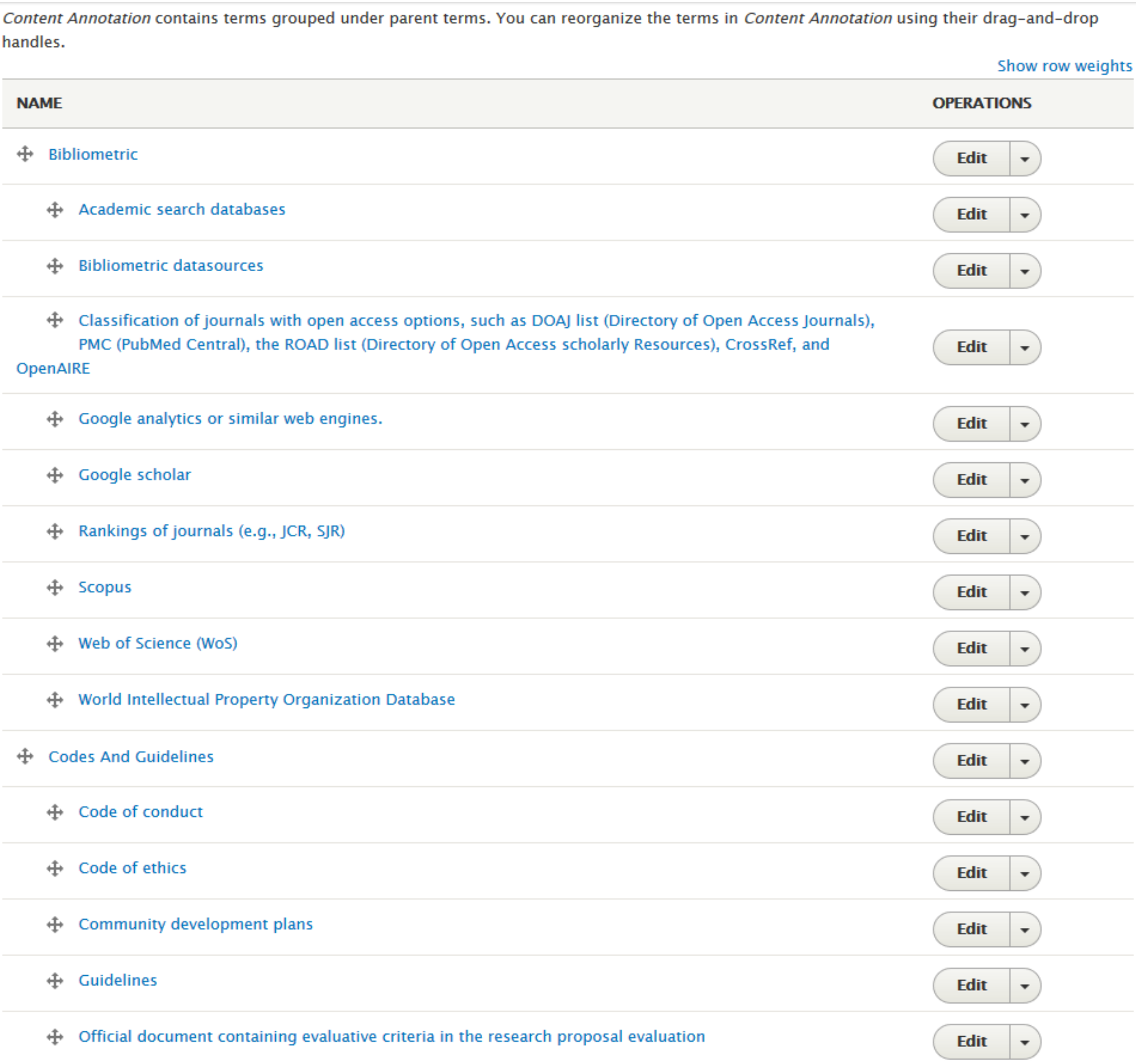

*Figure 10 – Master Scorecard Taxonomy*

The use of taxonomy enables regrouping and reorganizing of categories and subcategories by a simple drag and drop mouse movement, leading to the desired restructuring and automatically moving all the subsequent nodes.

[1] https://www.drupal.org/

## 4.2 Updated Taxonomy for Uploaded Content Annotation Use

For the document annotation needs, an annotation taxonomy has been collaboratively produced, which can be used in the classification of <u>all collected documentation</u>. A sample of the taxonomy outcome can be seen in Figure 14 - Content Annotation Taxonomy whereas the complete taxonomy is presented in APPENDIX 1 MULTI-ACT Annotation Taxonomy.



*Content Annotation* contains terms grouped under parent terms. You can reorganize the terms in *Content Annotation* using their drag-and-drop handles.

Show row weights

| NAME | OPERATIONS |
|---|---|
| Bibliometric | Edit |
| Academic search databases | Edit |
| Bibliometric datasources | Edit |
| Classification of journals with open access options, such as DOAJ list (Directory of Open Access Journals), PMC (PubMed Central), the ROAD list (Directory of Open Access scholarly Resources), CrossRef, and OpenAIRE | Edit |
| Google analytics or similar web engines. | Edit |
| Google scholar | Edit |
| Rankings of journals (e.g., JCR, SJR) | Edit |
| Scopus | Edit |
| Web of Science (WoS) | Edit |
| World Intellectual Property Organization Database | Edit |
| Codes And Guidelines | Edit |
| Code of conduct | Edit |
| Code of ethics | Edit |
| Community development plans | Edit |
| Guidelines | Edit |
| Official document containing evaluative criteria in the research proposal evaluation | Edit |

*Figure 11 – Content Annotation Taxonomy*

Subsequently, a user, upon uploading a document, will be able to easily annotate it, either via an auto-complete input field (Figure 12 – Auto-Complete Annotation) or by selecting a value in a tree like depiction of the annotation taxonomy. (Figure 13 – Annotation Tree-like Method)

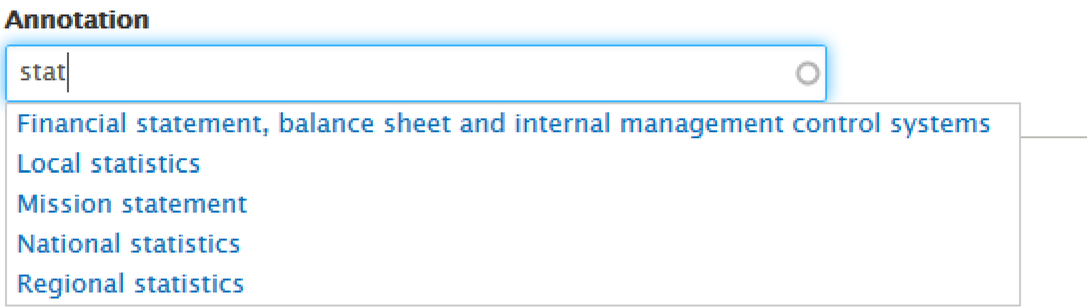



*Figure 12 – Auto-Complete Annotation*

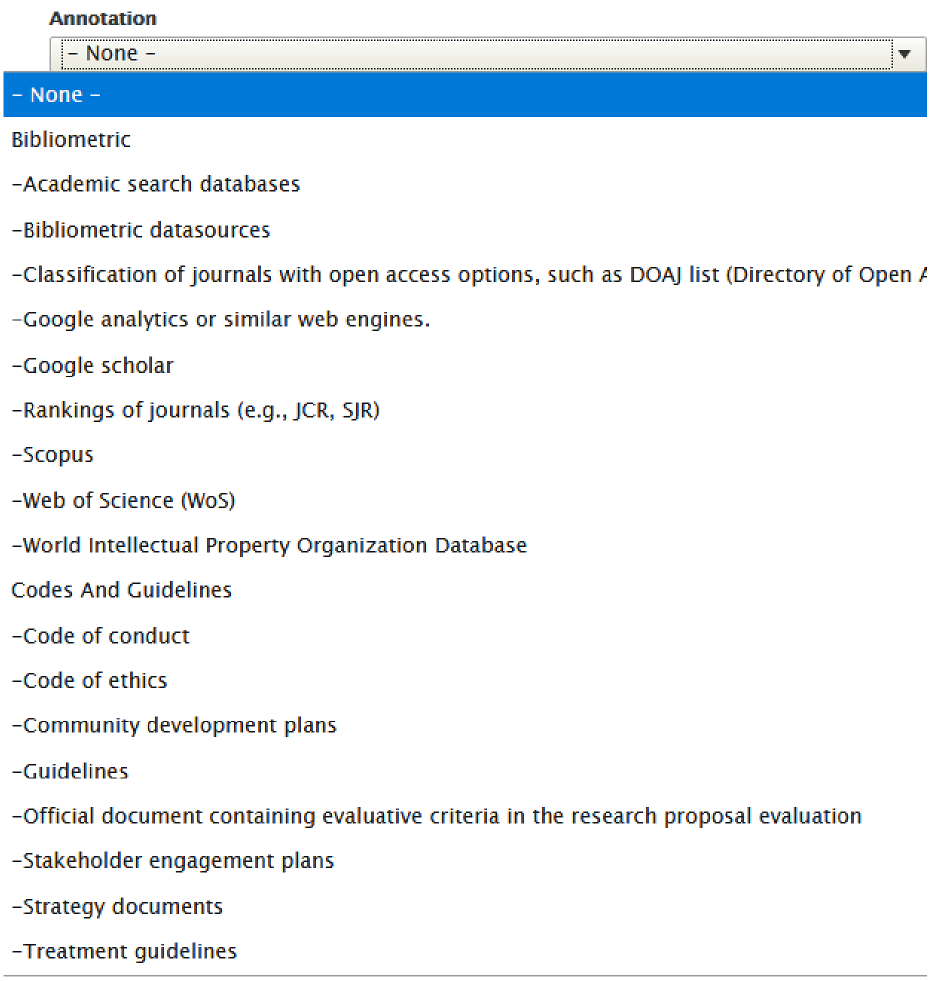



*Figure 13 – Annotation Tree-like Method*

# 5 CONCLUSIONS

Deliverable D2.2 has presented the data and content collection requirements and mechanisms related to MULTI-ACT Toolbox core functionalities as well as the organization of the content and data collection processes that will enable the accumulation of relevant Collective Research Impact Framework related structured content. It also described designed and in implementation Toolbox features and functions, supporting differentiated user roles and rights, introducing necessary author roles with content uploading, annotation and enrichment capabilities, as the feasible approach to collect and annotate content and data is through domain experts, users from relevant initiatives, etc.

In addition to capitalizing on internal consortium such domain expertise, in order to further and enrich the content and taxonomy related inputs and functionalities of the Toolbox, a first such activity has also been initiated with the first case study of MULTI-ACT, MS Care Unit. It has been agreed that respective surveys will be implemented in the Toolbox for use by the MS Care Unit and their users to fill in and collect vital information. This presents a first step towards knowledge acquisition of methods for data acquisition and in the long-term towards the creation and enrichment of an MS Research Impact database. Finally, the Deliverable further presented the enriched MULTI-ACT taxonomy for content annotation and the methods introduced in the Toolbox empowering users to easily annotate content and data, as well as easily enrich the taxonomy if necessary to ensure the proper classification of all the accumulated content so as to be validated and` usable in the MULTI-ACT Toolbox.

# APPENDIX 1 MULTI-ACT Annotation Taxonomy

Figure 14 visually presents the graph of the Content Annotation Taxonomy for annotating uploaded Documents/Data to the MULTI-ACT Toolbox, whereas the subsequent table presents in more detail the Taxonomy dimensions and values/elements.

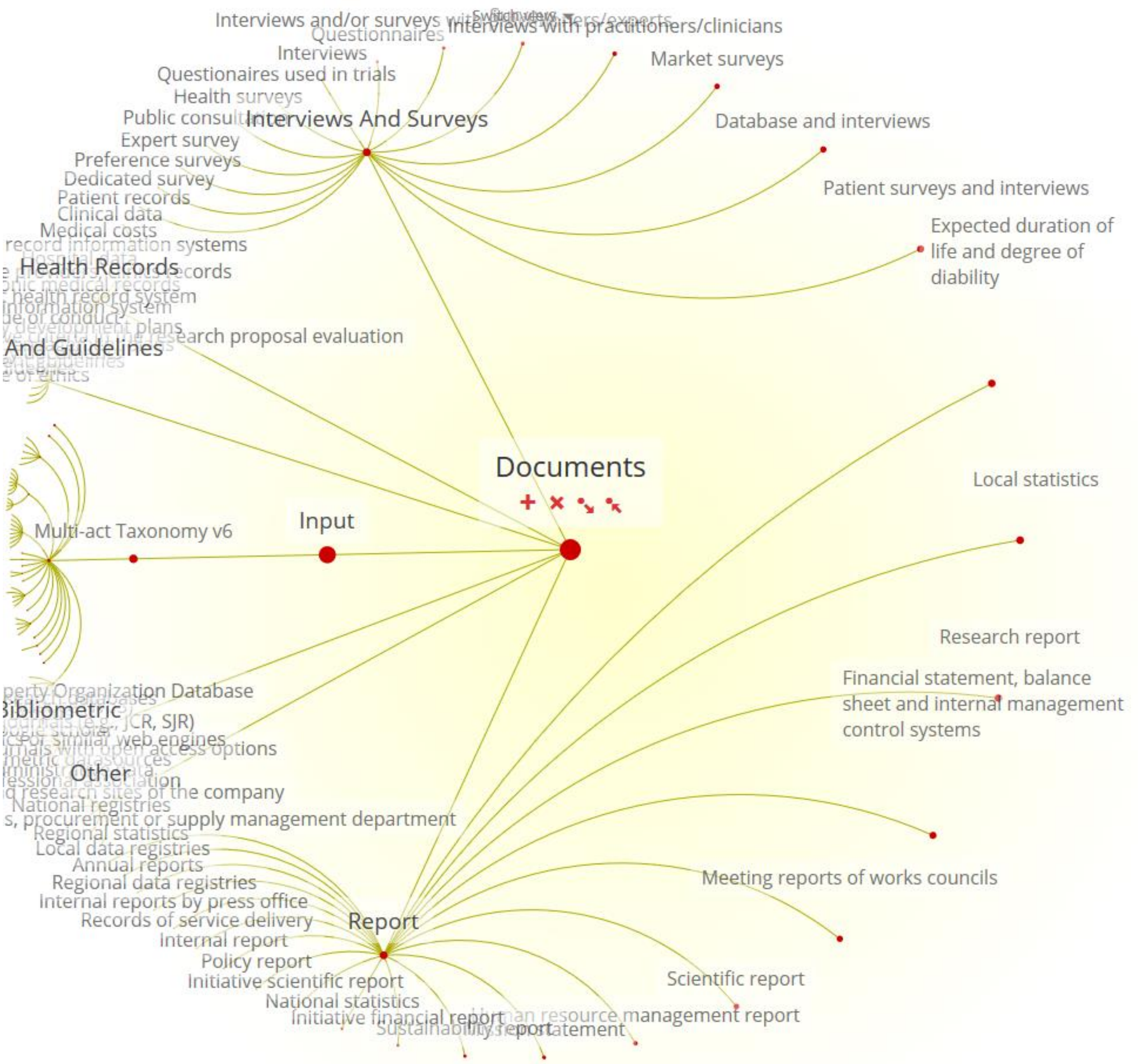



*Figure 14 - Content Annotation Taxonomy*

| | | |
|---|---|---|
| **Bibliometric** | | |
| | Academic search databases | |
| | Bibliometric data sources | |

| | |
|---|---|
| | Classification of journals with open access options, such as DOAJ list (Directory of Open Access Journals), PMC (PubMed Central), the ROAD list (Directory of Open Access scholarly Resources), CrossRef, and OpenAIRE |
| | Google analytics or similar web engines. |
| | Google scholar |
| | Rankings of journals (e.g., JCR, SJR) |
| | Scopus |
| | Web of Science (WoS) |
| | World Intellectual Property Organization Database |
| **Codes and Guidelines** | |
| | Code of conduct |
| | Code of ethics |
| | Community development plans |
| | Guidelines |
| | Official document containing evaluative criteria in the research proposal evaluation |
| | Stakeholder engagement plans |
| | Strategy documents |
| | Treatment guidelines |
| **Health Records** | |
| | Clinical data |
| | Clinical information system |
| | Electronic health record system |
| | Electronic medical records |
| | Health care providers, clinics records |
| | Hospital data |
| | Hospital record information systems |
| | Medical costs |
| | Patient records |
| **Interviews and Surveys** | |
| | Database and interviews |
| | Dedicated survey |
| | Expert survey |
| | Health surveys |
| | Interviews |
| | Interviews and/or surveys with practitioners/experts |
| | Interviews with practitioners/clinicians |
| | Market surveys |
| | Patient surveys and interviews |
| | Preference surveys |
| | Public consultation |
| | Questionnaires used in trials |

| | |
|---|---|
| **Reports** | Questionnaires |
| | Surveys |
| | |
| | Annual reports |
| | Expected duration of life and degree of disability |
| | Financial statement, balance sheet and internal management control systems |
| | Human resource management report |
| | Initiative financial report |
| | Initiative scientific report |
| | Internal report |
| | Internal reports by press office |
| | Local data registries |
| | Local statistics |
| | Meeting reports of works councils |
| | Mission statement |
| | National statistics |
| | Policy report |
| | Records of service delivery |
| | Regional data registries |
| | Regional statistics |
| | Research report |
| | Scientific report |
| | Sustainability report |
| **Other** | |
| | Administrative data |
| | Billing and accounting systems, procurement or supply management department |
| | National registries |
| | Production and research sites of the company |
| | Professional association |