DataPLAN: A Web-Based Data Management Plan Generator for the Plant Sciences

Xiao-Ran Zhou [1,*], Sebastian Beier [1], Dominik Brilhaus [2], Cristina Martins Rodrigues [3], Timo Mühlhaus [4], Dirk von Suchodoletz [3], Richard M. Twyman [5], Björn Usadel [1,6] and Angela Kranz [1,*]

1　IBG-4: Bioinformatics, Institute of Bio- and Geosciences, BioSC, CEPLAS, Forschungszentrum Jülich, 52428 Jülich, Germany; s.beier@fz-juelich.de (S.B.); b.usadel@fz-juelich.de (B.U.)

2　Data Science and Management, CEPLAS, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany; brilhaus@hhu.de

3　Computer Center, University of Freiburg, 79104 Freiburg, Germany; cristina.martins.rodrigues@rz.uni-freiburg.de (C.M.R.); dirk.von.suchodoletz@rz.uni-freiburg.de (D.v.S.)

4　Computational Systems Biology, University of Kaiserslautern-Landau, 67663 Kaiserslautern, Germany; timo.muehlhaus@rptu.de

5　TRM Ltd., P.O. Box 493, Scarborough YO11 9FJ, UK; richard@twymanrm.com

6　Institute for Biological Data Science, CEPLAS, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany

*　Correspondence: x.zhou@fz-juelich.de (X.-R.Z.); a.kranz@fz-juelich.de (A.K.); Tel.: +49-02461-61-85881 (X.-R.Z.); +49-2461-61-85833 (A.K.)

Abstract: Research data management (RDM) combines a set of practices for the organization, storage and preservation of data from research projects. The RDM strategy of a project is usually formalized as a data management plan (DMP)—a document that sets out procedures to ensure data findability, accessibility, interoperability and reusability (FAIR-ness). Many aspects of RDM are standardized across disciplines so that data and metadata are reusable, but the components of DMPs in the plant sciences are often disconnected. The inability to reuse plant-specific DMP content across projects and funding sources requires additional time and effort to write unique DMPs for different settings. To address this issue, we developed DataPLAN—an open-source tool incorporating prewritten DMP content for the plant sciences that can be used online or offline to prepare multiple DMPs. The current version of DataPLAN supports Horizon 2020 and Horizon Europe projects, as well as projects funded by the German Research Foundation (DFG). Furthermore, DataPLAN offers the option for users to customize their own templates. Additional templates to accommodate other funding schemes will be added in the future. DataPLAN reduces the workload needed to create or update DMPs in the plant sciences by presenting standardized RDM practices optimized for different funding contexts.

Keywords: DMP; planning; omics; phenotyping; botany; research data management

1. Introduction

A data management plan (DMP) sets out the procedures and policies for handling, sharing and analyzing research data throughout a project's life cycle and beyond [1–3]. DMPs present a clear and structured approach for the management of all types of anticipated research data, including strategies for data acquisition, organization, documentation, storage, access and sharing [4,5]. DMPs thus contribute to efficient and productive research data management (RDM) and enhance the findability, accessibility, interoperability and reusability (FAIR-ness) of data [5]. Many funding bodies require a DMP. A preliminary version of the DMP should be provided in the proposal, and then formalized during the project's early stages before the collection of data and metadata begins. This comprehensive approach ensures that a robust RDM strategy is integrated into the project, but the preparation of a DMP involves several challenges (Figure 1).
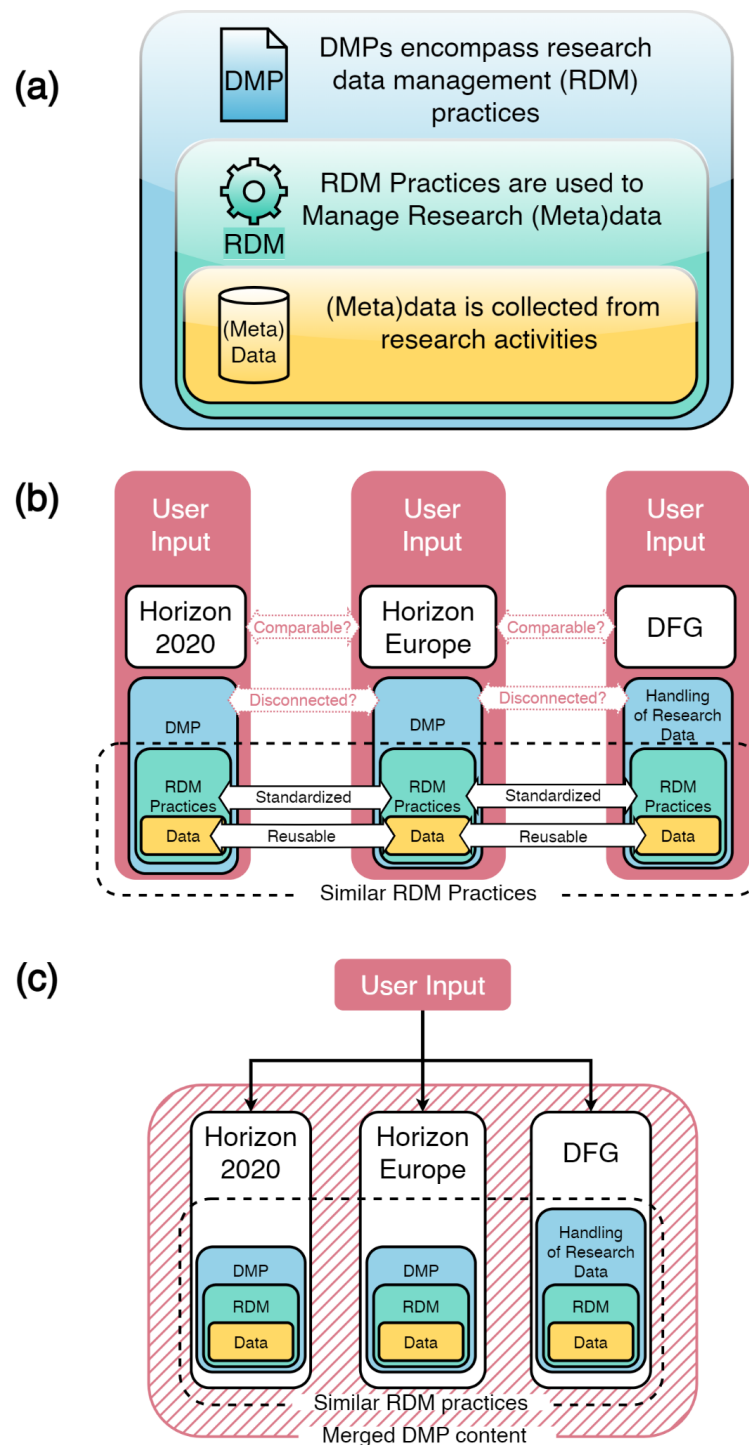
Figure 1. DMPs prepared for multiple projects can be merged if they all use standardized RDM and have reusable metadata and raw data. (**a**) DMPs encompass RDM practices for raw data and metadata. The content of a DMP is dependent on the RDM practices used. (**b**) Although DMPs encompass reusable standardized RDM practices in similar plant-related projects (blue, green and yellow boxes) [1], the contents of DMPs prepared for different projects or funding agencies are disconnected [6,7], and users must provide input (red boxes) multiple times even though the DMPs only have minor differences. (**c**) If similar or standardized RDM practices (blue, green and yellow boxes) are used in different projects, the content of different DMPs can be merged. The merged content can be provided for use in diverse projects with different funding agencies and programs. This reduces the user input (red box) compared to that shown in panel (**b**).

Major funding bodies and programs, such as Horizon Europe [8], the German Research Foundation (DFG) [9], the National Institutes of Health (NIH) [10], the National Science Foundation (NSF) [11] and the Swiss National Science Foundation (SNSF) [12], expect applicants to submit RDM-related documents as part of the proposal. The DMP requirements are set by funding bodies such as DFG or programs such as Horizon 2020 and Horizon Europe [13–15], and the questionnaires in each case are unique. The checklist of the DFG's "Handling of Research Data" document, which is treated as a DMP equivalent [16,17], consists of 19 questions, whereas the Horizon Europe DMP questionnaire comprises 45 questions. FAIR principles are mentioned in Horizon Europe, Horizon 2020 and DFG questionnaires; therefore, RDM practices should be comparable and partially connected. However, the connections between DMP contents in different projects are not always as evident as the standardized RDM practices, data and metadata (Figure 1b).

Compliance with the requirements of different funding bodies is time-consuming, and multiple tools have been developed to tackle the resulting challenges [2,18–29]. These tools help researchers to write comprehensive DMPs by providing a structured framework and guidelines. However, they vary in their support for the preparation of DMPs for specific funding bodies [2,18,19,30], the ability to focus on domain-specific RDM approaches [25,27,31], and the inclusion of relevant laws, directives and regulations [2,18,19,23,25,26,32,33]. To enable information exchange among these variants, the maDMP (Common Standard for Machine-actionable DMPs) [34] has been implemented in multiple tools [18–20,35–39]. This standard can be supplemented by decentralized RDM practices [40], because each research field may have its own set of standards, protocols, best practices for data management [41–47], and unique data management challenges and opportunities [7]. To our knowledge, no tools are yet available that can merge DMP components to match the requirements of different funding bodies or programs in the plant sciences (Figure 1c).

Plant sciences have specific needs regarding RDM practices. Many data types are specific to the plant sciences or are shared with other disciplines but have aspects that are unique to plants. For example, growth and biomass accumulation occur in all living organisms, but organ-specific growth parameters such as leaf and root biomass [48–51] and the number of flowers and fruits are unique to plants. Photosynthetic performance is shared with some microbes, but it is measured in a unique way in plants [52]. High-throughput phenotyping [49,53] in plants generates large, highly-dimensional datasets [54,55] based on different types of imaging [56] among other data types [54,55], and these must be annotated with plant-specific ontologies [57]. Plant taxonomy is another example, particularly in the context of plant–microbe interactions [58]. These rich and diverse datasets must also be merged with fairly universal data types (such as genomics, transcriptomics, proteomics and metabolomics data), which generate heterogeneous workflows. RDM platforms for plants therefore need to accommodate these unique aspects.

Initiatives such as DataPLANT, FAIRagro and ELIXIR [59–63] have developed numerous tools to help plant scientists overcome RDM challenges. As part of the German National Research Data Initiative (NFDI), the DataPLANT consortium [59] provides a research data infrastructure for the plant sciences with a user-centric approach. The heavy workload in RDM can be reduced by using RDM platforms with workflow descriptions and data annotations. For example, numerous detailed descriptions can be included in the DMP by using the DataPLANT platform [47], which relies on the investigation–study–assay (ISA) standard [64,65] and common workflow language (CWL) [66]. Many research methods associated with the fundamental plant sciences, such as genome sequencing/assembly and phenotyping, follow a standardized workflow with little variation between plant species. This allows the use of prewritten content that can be intuitively chosen by users and integrated using a web-based tool.

Plant scientists must also comply with relevant legislation and recommendations that govern how research data should be managed, shared, and accessed, especially in the context of international collaborative projects that involve cross-border data transfer. For example, under EU legislation, the utilization of genetic resources requires permission from

the country of origin [67,68]. These requirements and standards can vary depending on the country, region and research topic. User input should be monitored automatically to detect related research objects and show warnings before exporting the output.

Here, we present DataPLAN, a web-based DMP generator for the plant sciences that can easily be used and adapted for other scientific disciplines with similar objectives, methods and data types. DataPLAN was developed as part of the NFDI DataPLANT framework [47] and allows users to select tools and services provided by DataPLANT. DataPLAN requires only one set of inputs to generate three DMPs modeled according to different funding agencies or programs: DFG [9], Horizon 2020 [69] and Horizon Europe [8]. Furthermore, DataPLAN includes a practical user guide, which is dynamically generated based on user input, providing valuable assistance throughout the project life cycle, offering insights, tips and reminders for effective RDM. The text modification function code for the generation of DMPs is implemented using client-side JavaScript. DataPLAN's source code is freely available on GitHub under a GNU General Public License (GPL-3.0) (https://github.com/nfdi4plants/dataplan, accessed on 25 September 2023), ensuring transparency and promoting more collaboration within the plant sciences research community.

2. Materials and Methods

2.1. DMP Template and Questionnaire Design

The prewritten DMP contents were prepared manually. The programming functions to convert user input into DMP output were manually coded in client-side (pure frontend) JavaScript. DMPs for DFG and Horizon Europe relied on DataPLAN templates, which were prepared in three consecutive steps. First, the questions in the DFG and Horizon Europe questionnaires were answered manually (Figure 2, Step 1). Second, the answers were adapted to align with typical data and metadata standards or data types that are relevant for the fundamental plant sciences (Figure 2, Step 2). Third, in order to merge DataPLAN's answers for different subdisciplines, we also studied underlying data generation and analysis methods, data formats, known minimum information standards, and repositories for data publishing (Table 1). The related information (e.g., which minimum information standard should be used to report metadata from specific high-throughput technologies) was incorporated into the templates.
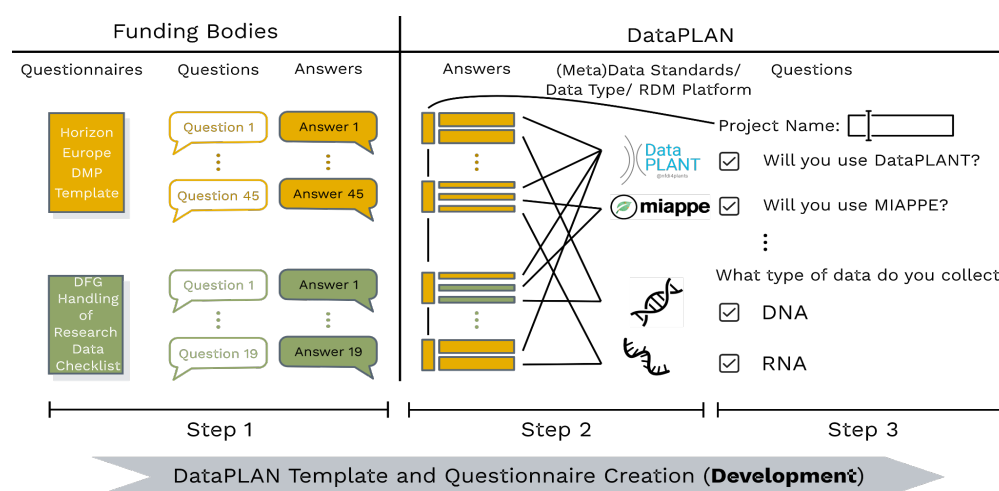

Figure 2. DataPLAN template and questionnaire design. Step 1: Manual checking and answering of questions in the DFG and Horizon Europe questionnaires. Step 2: Generation of reusable answer building blocks for each funding body. We ensure the answers comply with existing metadata standards, data types and RDM platforms so that they can be reused between different projects and funding bodies. Step 3: Design of the questions displayed by DataPLAN followed by matching them with the reusable answers generated in step 2.

To construct the questionnaire for DataPLAN, we first identified analogous questions in the DMP templates and guidelines provided by Horizon Europe and DFG. The similar questions are grouped and able to generate a repository of responses (Figure 2, Step 3). These responses can be seamlessly utilized to complete all prewritten templates, thus enabling the creation of DMPs that align with the criteria for Horizon Europe and DFG.

2.2. Software Development

The DataPLAN web-based software (https://plan.nfdi4plants.org/, accessed on 25 September 2023) was developed using client-side (frontend) JavaScript [70] and is designed as a self-contained single-page application (SPA). As shown in Figure 3a, we stored the interface, functions and reusable DMP content in a single HTML file. The interface is exposed, whereas the other content is hidden. The single HTML file can be executed on all modern web browsers, including Google Chrome, Microsoft Edge, Mozilla Firefox, Opera and Safari. A client-side SPA ensures data security by eliminating the need for software installation or information transfer to external servers. In terms of implementation, several open-source libraries were used to enhance the user interface (UI) and visualizations: Bootstraps 5 [71], bs5-intro-tour [72], d3 word cloud [73], FileSaver [74], and split.js [75]. DataPLAN is accessible via the DataPLANT website (https://plan.nfdi4plants.org/, accessed on 25 September 2023).
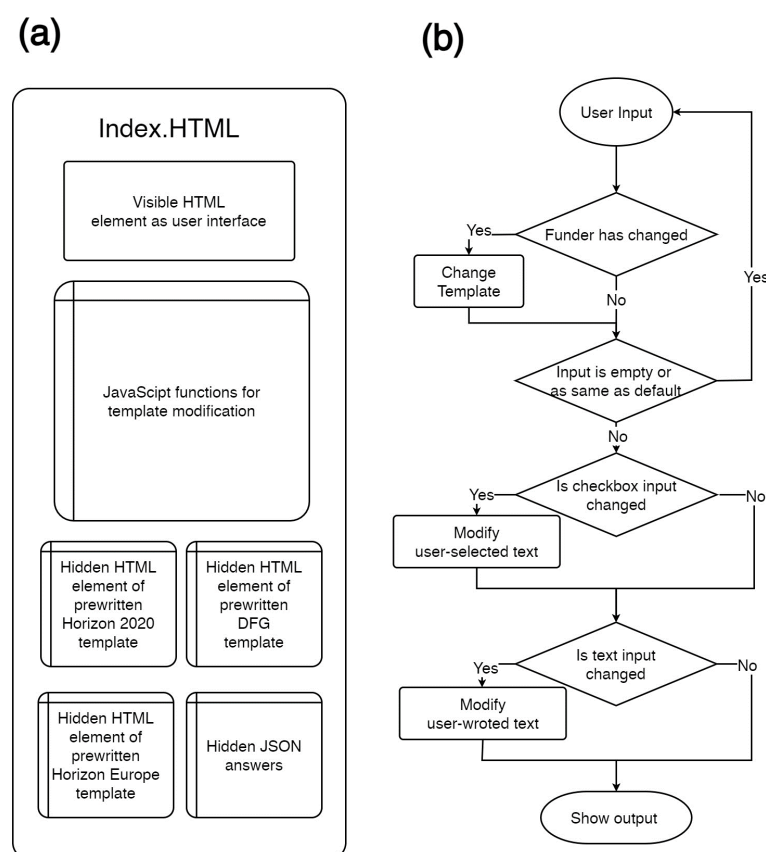

Figure 3. DataPLAN architecture and core function flowchart. (**a**) The architecture of DataPLAN consists of a single index.html document. The exposed HTML elements build the user interface. The integrated JavaScript functions are used to modify the user interface. The answers and prewritten DMP templates are stored as hidden elements in the same HTML file. (**b**) The core function of DataPLAN consists of four decisions to check which one of the three processes should be run. The four decision blocks are used to check for changes in the template, overall input, checkbox input, and text input. The three processes are (1) template change, (2) *user-selected text* modification, and (3) *user-written text* modification. After the correct processes, the final output is shown to users.

To generate the DMP output, the entire template is searched multiple times. Figure 3b shows four diamond blocks. These decide which rectangular blocks need to run searches. Following each search, a portion of the text is modified based on the user input. Partial text modification is achieved using two placeholders: first, a regular placeholder, denoted by the $_ symbol, which simply inserts the *user-written text* into the sentence; and second, the use of more complex structures, where placeholders are embedded within conditional statements ("if" structures) that are controlled by either user input or the choice of template, which are stored in the hidden HTML elements (Figure 3a). These placeholders are identified using the window.find() function, with the search range extended to the left by three characters in order to locate the relevant conditional statement (#if or #if!). Additionally, once the #if has been found, the right boundary #endif must also be located. When these conditions have been met, a decision tree determines the appropriate changes to be made based on the user input.

2.3. Testing

To ensure DataPLAN functions robustly, we carried out comprehensive testing by combining automated technical assessments with user-centered evaluation. Automated testing involved the use of PerformanceMeasure [76] and the web development tool Light-House [77] to gauge the system's technical performance. Manual testing was conducted in two phases: first, developer testing, which scrutinized core functionalities, including placeholder replacement, logic condition of placeholders, and template switching; and second, testing by plant science experts, who served as potential users. This user testing assessed various critical aspects, including usability, time efficiency, user interface intuitiveness, template selection, customization options, adherence to RDM practices, collaborative features, version control capabilities, compliance checks, and import and export functionalities for DMPs.

3. Results

3.1. DMP Content Generation and Modification Using DataPLAN

Prewritten DMP content stored in DataPLAN is completed by user input in two stages. First, users are presented with a collection of manually curated prewritten templates, and these are customized based on user input. The content of the prewritten templates aligns with a unified questionnaire, minimizing the need for users to provide repetitive information. To facilitate this integration, we established a many-to-many relationship between each user response and the various components of DMP content, ultimately generating multiple documents tailored to the requirements of three funding agencies or programs. In the following sections, we describe the nature of these prewritten DMP templates, explain their formatting and categorization, and describe their adaptability for use across distinct funding agencies and programs.

3.1.1. Incorporation of RDM Practices and Platforms for the Plant Sciences

FAIR RDM practices are included in the prewritten DMP template as an option, and the advanced use of RDM platforms is recommended and set as the default. The prewritten templates adhere to FAIR principles by including RDM tools, data/metadata standards, RDM platforms, and endpoint repositories (Figure 2, Step 2). For example, to make data findable, unique identifiers, ontologies and annotations with metadata can already be used to answer questions such as "Will search keywords be provided that optimize possibilities for reuse?" in the Horizon 2020 questionnaire. To make the DMP more practical, the ontology services in DataPLANT are provided as a default to help users generate new ontology terms for data annotation. The implementation of DataPLANT as an RDM platform only requires the user to select one checkbox. By choosing DataPLANT [59,78], relevant concepts, tools and services such as the Annotated Research Context (ARC) [79], Swate [80], ARC Commander [81], and DataHUB [82] will be included in the output documents accordingly.

Consequently, if users follow the RDM practices offered by DataPLANT, this increases the FAIR-ness of their data.

DataPLAN aligns with widely recognized minimal information standards (Table 1), such as MIAME [45], MinSEQe [83], MIAPE [44], MSI [84] and MIAPPE [43]. These standards provide a common framework for the description and organization of data elements related to experimental planning, sample handling and data collection/analysis. By recommending these standards, DataPLAN enables researchers to produce structured metadata, enhancing data discoverability, reusability and reproducibility. Furthermore, the recommendations are also based on the data type selected in the DMP. Ontologies also play an important role in metadata annotations and data transformation. To enhance the annotation and transformation of metadata, DataPLAN will integrate DataPLANT Biology Ontology [85] and other relevant ontologies as a part of the DataPLANT platform.

Table 1. A collection of minimum information standards, endpoint repositories and data management platforms relevant to DataPLAN.

	Genomics	Transcriptomics	Proteomics	Metabolomics	Plant Phenotyping
Minimum information standard	MinSEQe [86]	MIAME [45] MinSEQe [86]	MIAPE [44]	MSI [84]	MIAPPE [43]
Endpoint repositories	ENA [87] EBI NCBI [88] DDBJ [89] SRA [90] GenBank [91]	GEO [88] SRA [90]	PRIDE ProteomeXchange [92]	Metabolights [93]	e!DAL-PGP [46] Gnpis [94] EURISCO [95]
RDM platform			DataPLANT [59]		

3.1.2. Categories of Prewritten DMP Content

The text content of the prewritten DMP templates is manually created and stored in several hidden HTML elements inside the index.html file. Based on their mapping to the user input, prewritten text can be assigned to three different functional groups, which are *static text* (black in Figure 4), *user-selected text* (green in Figure 4) and *user-written text* (text with yellow background in Figure 4).

The *static text* content within DataPLAN features prewritten text elements, including text derived from funding body templates and general introductory information. For example, the black text in the "Introduction" of Figure 4 represents static content that explains FAIR principles and the DMP's functions in this context. This *static text* content originates from a DMP document associated with a plant-focused Horizon 2020 project and is reusable across different DMP documents, thus remaining consistent and unaffected by user input variations. Although the content of *static text* is stable, its specific placement within the templates may vary. This reflects differences in the order of questions posed by different funding agencies or programs. Users also have the flexibility to modify the *static text* within user-defined templates according to their specific needs and preferences.

The second category of prewritten DMP content is *user-selected text* based on the checkbox option, which is shown in green in the right-hand panel of Figure 4. This text is designed to accommodate flexible scenarios contingent upon user choices. *User-selected text* includes guidance on metadata standards, handling specific data types, and other adaptable content that varies based on user selections. Each *user-selected text* block is linked to one or more selectable options in the user interface. Consequently, the content of *user-selected text* depends on user input. The *user-selected text* is stored alongside *static text* but within a structure that enables its inclusion or exclusion from the final DMP output. For example, the first sentence (in green) in the "Introduction" section (Figure 4) is *user-selected text* explaining that the project "...is a part of the Open Data Initiative (ODI) of the EU." This *user-selected text* is associated with the *EU project* option in questionnaire field 1.4 in

Figure 4. If the user selects the *EU project* checkbox, the green text will be incorporated into the DMP output, but otherwise it will be excluded.
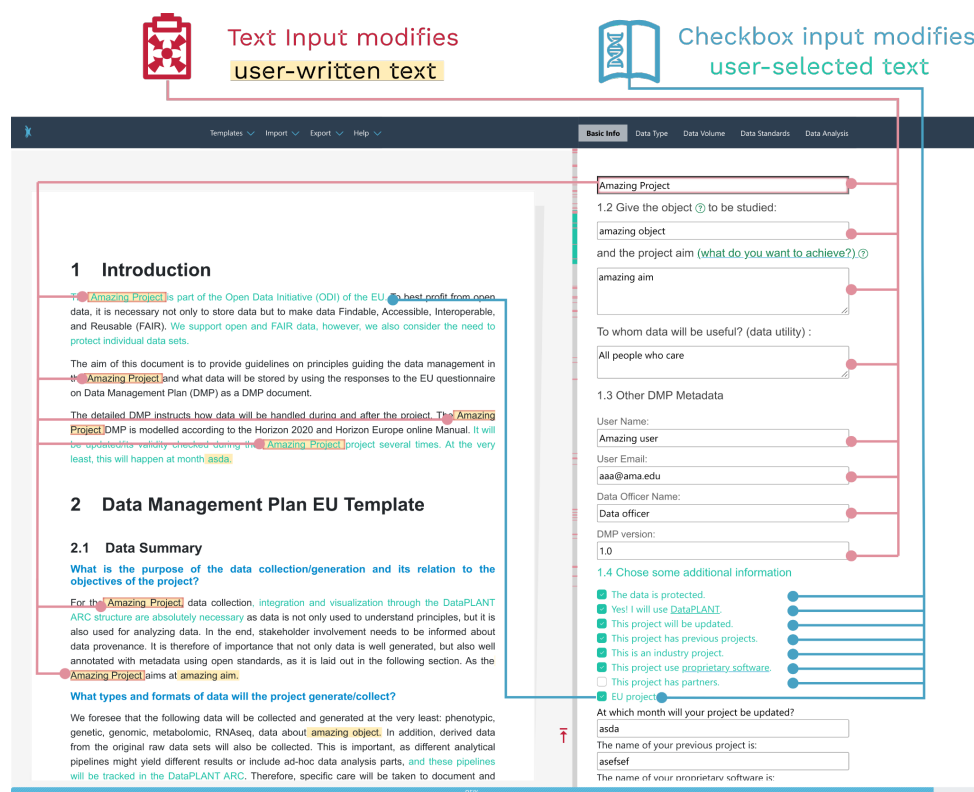

Figure 4. The web-based user interface of DataPLAN. The left panel displays a live preview of the DMP, while the right panel displays the DataPLAN questionnaire. In the left panel, the *static text* is shown in black, the *user-selected* text is shown in green, and the *user-written* text has yellow highlights. In the right panel, text inputs are indicated by red lines, and reusable answers in the checkbox format are indicated by blue lines. The red lines connect all the answers given by the text input, which is located at the top of the questionnaire in the right-hand panel, and a blue line connects the answer associated with selecting the *EU project* option in the checkbox. This connection between the left and right panels is also animated when the DataPLAN is used.

The third category of prewritten template content pertains to *user-written text*, which is essential project-specific information provided by the user. For example, the initial *user-written text* highlighted with a yellow background in Figure 4 is the project name "Amazing Project." This text is linked to question 1.1: "What is the project name or acronym?" User-written text can coexist with both *static text* and *user-selected text*, enabling users to tailor their DMPs for specific project needs while incorporating standardized and user-defined content as required.

3.2. User Interface

3.2.1. Main Menu

The DataPLAN user interface features an intuitive layout, comprising a live preview on the left and a questionnaire on the right (Figure 4). Users can seamlessly navigate and interact with the tool within the same webpage throughout their session. The main menu above the live preview offers four dropdown menus: *Templates*, *Import*, *Export*, and *Help*, providing 22 clickable options. In the *Templates* section, users can switch between three current templates (H2020 DMP, Horizon Europe DMP, and DFG "Handling of Research Data"), access a practical guide, or load a custom user-defined template. The *Import* section allows users to import answers, clear user input, generate word clouds, update front page images, or load answers from cache. The *Export* section offers several choices, including

copying the text output, exporting answers to JSON, printing the document to pdf or docx, updating reminders, and saving answers to cache. In the *Help* dropdown menu, users can access Tutorial, Wiki, GitHub, Print Questions, and Changelogs. For new users, DataPLAN provides a guided tour to acquaint them with its layout and functionality.

3.2.2. Questionnaire (Right Panel)

Questions can be broadly assigned to two categories: those requiring text input and those requiring the selection of one or more options from a list of checkboxes (Figure 4, right panel). Text inputs are linked to *user-written text* such as the name of the project, the study topic, and the project's aim, because these details vary between projects and cannot be answered using prewritten options. Conversely, the questions listed under field 1.4 (Figure 4) can be answered by selecting checkboxes, which will insert prewritten answers into the DMP. In some instances, checkboxes may also require text input. For example, if the option "*This project will be updated*" is chosen in field 1.4 (Figure 4), a prompt will appear in the questionnaire asking the user to specify the month in which the update is planned.

The sequence of questions reflects both the order in which they appear in the DMP document and the logical progression of the research data life cycle. For example, the *study topic* and *project aim* questions precede the *data type* question because the first two are determined earlier in the process than the latter.

3.2.3. Live Preview (Left Panel)

The left panel of the interface presents a live preview of the DMP document (Figure 4, left). User-entered project-specific information is highlighted with a light-yellow background, while predefined responses based on selected checkboxes are indicated in green text. Both types of text are interactive, allowing users to hover over or click on them for further engagement. To enhance usability, automatic scrolling within the questionnaire is triggered only when users click on the corresponding text, reducing the likelihood of errors. Furthermore, clicking on an answer or question causes all instances of the same text to be marked in red on the scroll bar, helping users to locate pertinent information. This highlighting feature serves as a visual guide, revealing the relationship between the questionnaire and the resulting output.

3.3. DataPLAN Workflow

DataPLAN has five main options: input, DMP generation, template change, warning/reminder, and output. Two of these (DMP generation and warning/reminder) run in the background, while the others provide input by answering questions (Figure 5, green), select or customize a template (Figure 5, blue), or collect the output (Figure 5, black).
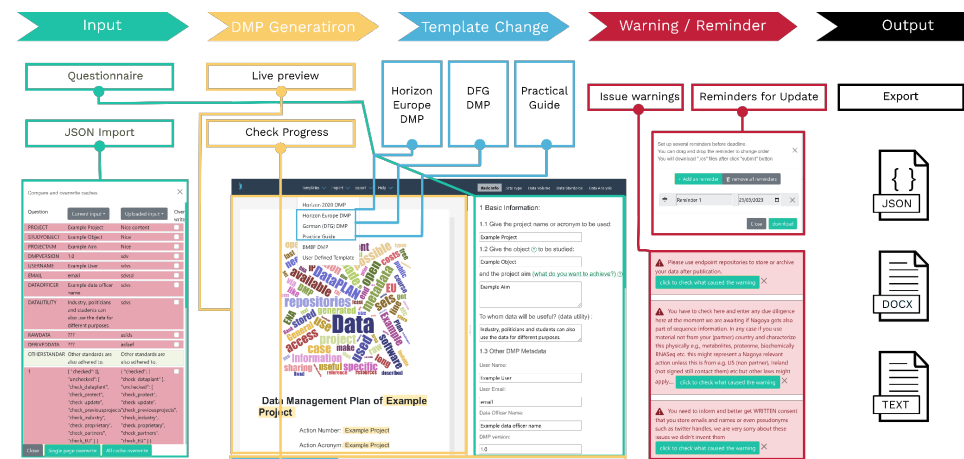

Figure 5. DataPLAN has five main steps. Input (green): Users can provide input either by completing the questionnaire manually or by importing saved input. DMP Generation (yellow): DMP generation

(core function) loads the user input into a prepared template. Template Change (blue): Users can change templates at any time. We have provided users with features to help create user-defined templates. Warning/Reminder (red): We use warnings to make users aware of potential hurdles. The text locations that cause the warning are shown in the live preview. Reminders are downloadable ics files that can be imported into a calendar. Output (black): The output of DataPLAN can be a text, doc, or JSON file.

3.3.1. Saving and Importing Data

Users can provide input either manually (by completing the questionnaire) or by importing previously saved responses. The *Export* dropdown menu features a *Download to JSON* function, allowing users to export responses in a machine- and human-readable JSON format compatible with other DMP JSON standards [96]. This enables users to import their data and resume the DMP generation process (Figure 5, yellow). When a JSON file is imported, the tool compares it with the current version of the DMP and presents several replacement options, such as replacing a single answer, replacing a group of cached answers, or fully replacing both the displayed and cached information.

3.3.2. Main Output (DMP-Related Documents)

DataPLAN's *Export* dropdown menu offers a range of options for DMP management: (1) copying text for further editing in a text editor, (2) exporting answers as a JSON file for reuse, (3) printing the document as a pdf or docx file, or directly to a printer, (4) setting an update reminder by generating an ics file that can be loaded into a calendar, and (5) saving currently displayed answers to the browser cache with five available slots. Option (1) enables users to copy one of three predefined DMP templates that can be used for Horizon 2020, Horizon Europe, and the DFG. The text can be pasted into any text editor or printed either as a pdf or HTML using option (3).

In addition to the main DMP documents, DataPLAN generates a practical guide based on user input. This guide offers detailed and optional measures to facilitate task assignments, tool selection, data format conversion, and timeline arrangements. As a supplement to the main DMP, the practical guide is dynamically updated based on the user's responses to questions. Examples of DMP documents generated by DataPLAN can be found in Supplementary Documents S1–S4.

3.3.3. Warnings

DataPLAN considers regulatory and legal considerations that arise during the creation of a DMP. For example, certain data types, such as international genomic resources and personal data, are subject to specific regulations and laws such as the Nagoya Protocol and the General Data Protection Regulation (GDPR) [97,98]. To assist users navigating these complexities, DataPLAN incorporates a warning system (Figure 5, red) that is activated when a selected data type or format falls within the purview of these regulations. Before printing or copying the DMP, these warnings serve as timely reminders, highlighting potential challenges in RDM and publication. Moreover, they include links to relevant sections in the text, allowing users to access the information conveniently.

3.4. Testing and Validating DataPLAN According to FAIR Principles of Software
3.4.1. Findability of the Software

To ensure DataPLAN complies with FAIR principles of software [99], we have implemented several measures to increase data findability. The tool is hosted on GitHub [100], so developers can easily track changes in the source code and collaborate on further modification. Furthermore, the use of HTML meta-tags ensures search engine optimization of the DataPLAN website, increasing the likelihood that users will discover the tool via search engine queries. Notably, DataPLAN was included in the online registry RDMkit [101] and bio.tools [102], further enhancing its findability in the research community.

3.4.2. Accessibility of the Software

The accessibility of DataPLAN has been optimized by ensuring that it can be accessed using communications protocols that are open, free and universal. Users can access Data-PLAN from desktop and mobile devices using a wide range of modern browsers, including Chrome, Firefox, Edge, Safari and Opera. For offline work, users can download the HTML, CSS or JavaScript code using the browser's *save* function and store them locally on their device. In addition, DataPLAN is hosted on GitHub, granting users the freedom to fork the repository and host their own personalized version. This flexibility enables users to customize and tailor the tool to suit their specific requirements. With no installation or login required, DataPLAN can be used as soon as the website has been downloaded (by their browser). The tool's lightweight source code (less than 1 MB) and fast loading also contribute to its accessibility. The assessment results from LightHouse (without "word cloud" generation) show that it has very good accessibility (Figure 6).

Figure 6. LightHouse results confirm that DataPLAN has an accessibility score of 95/100 (without word cloud generation). All other assessments (performance, best practices and search engine optimization) were also evaluated as good.

3.4.3. Interoperability of the Software

DataPLAN has been developed to put interoperability in mind by using only HTML and JSON to store data and metadata. Currently, the tool's metadata is compliant with the maDMP at minimal level and stores the DMP content in HTML files that are also machine-actionable. DataPLAN's purely client-side code enables integration into automated workflows. The ability to accept user-defined templates allows researchers to customize and tailor DMPs to their specific needs and requirements, increasing the flexibility and utility of the tool.

3.4.4. Reusability of the Software

DataPLAN's open-source licenses allow cost-free reuse by individuals and projects [103]. The code for DataPLAN is modular, allowing developers from other fields to reuse parts of the code for their own purposes. Even people without programming skills can effortlessly create their own templates and use them to prepare DMPs. In accordance with FAIR principles, DataPLAN's DMP content and metadata have a detailed and transparent provenance, they are accompanied by a license that allows reuse, and they meet domain-relevant community standards, further enhancing reusability.

4. Discussion

DataPLAN, a client-side web-based application with less than 1 MB of source code, facilitates the user-friendly [104] preparation of DMPs. DataPLAN comprises a two-panel webpage and three DMP templates that encompass RMD best practices in the plant sciences. The templates, which incorporate 12 data types, nine endpoint repositories, and one RDM platform [47], can be used for Horizon 2020, Horizon Europe and DFG projects.

4.1. Comparison with Existing DMP Tools

The DMP community acknowledges the importance of maintaining a diverse range of tools [1] to meet the specific needs of different research domains. Several tools have been developed to assist users in the preparation of high-quality DMPs, such as Data Stewardship Wizard [18], DMP online [105], DMP tools [2], DMP Canvas Generator, EzDMP [27],

DMPRoadmap [106], RDMO [28,33], Research Data Manager (UQRDM) [21], DataWiz [22], ARGOS [23], UWADMP [26], DMPTY [25], easyDMP [24] and DAMAP [39]. Table 2 summarizes their key characteristics, programming languages, current templates, customizability, openness, and convenience in comparison to DataPLAN.

Table 2. Existing DMP preparation tools compared to DataPLAN.

Name	Programming Language	Funding Body Templates	Customizable	Templates	Content Preview	Open Source
Data Stewardship Wizard (DSW) [18]	Haskell ELM	3	Yes, with programming	Yes	Yes [107]	No
DMP Canvas Generator [108]	JavaScript	0	No	No	No	No
DMPonline [19]	Ruby JavaScript	18	Yes, with programming	No	Yes [29]	No
DMP tools [2]	Ruby JavaScript	19	Yes, with programming	No	Yes [29]	No
DMProadmap	Ruby JavaScript	19	Yes, with programming	No	Yes [29]	No
RDMO [20]	Python (Django) and JavaScript (AngularJS)	6	Yes, with programming	No	Yes [30]	Yes
Research Data Manager (UQRDM) [21]	Not available	0	Not available	No	No	No
DataWiz [22]	JAVA	0	Not available	No	Yes [32]	No
ezDMP [27]	Not available	0	No	No	No	No
ARGOS [23]	JAVA and Typescript	0	Yes	No	Yes [29]	No
UWADMP [26]	Not available	0	NA	No	No	No
DMPTY [25]	JavaScript, HTML	0	No	Yes	No	No
easyDMP [24]	Python (Django)	1	No	No	No	No
DAMAP [39]	Typescript, Java	1	NA	NA	No	Yes
DataPLAN (the tool in this paper)	Frontend: JavaScript No backend	3	Yes, no need for programming	Yes	Yes	Yes

4.1.1. Technical Comparison

All the tools listed in Table 2 are web-based and use JavaScript for the frontend. However, the backend programming languages vary. DataPLAN and ARGOS both provide customizable templates without coding, whereas Data Stewardship Wizard, DMP online, DMP tools, DMPRoadmap and RDMO require code to be written for customized templates. Data Stewardship Wizard, DMPTY and DataPLAN provide a content preview, so users can see the output in real time before finishing the questionnaire. Tools such as RDMO, Data Stewardship Wizard, DMP online, DataWiz and DataPLAN can also be used offline. Among them, RDMO, Data Stewardship Wizard, DMP online and DataWiz require a local server, whereas DataPLAN can be used offline as soon as the webpage is loaded or downloaded. DataPLAN, along with UWADMP, provides DMP services without needing users to log in, whereas the other tools require user registration and login.

All the tools use open-end questions (text) similar to the questionnaires provided by the funding bodies, but some tools, such as ezDMP, Data Stewardship Wizard, Open DMP, UWA DMP, RDMO and DataWiz also ask closed-ended (checkbox) questions. In contrast, DataPLAN includes six closed-ended questions and 10–13 open-ended questions to collect project-specific information. DataPLAN offers a general questionnaire that is mapped to the templates provided by different funding bodies. Whereas other DMP tools have individual entry masks for each template, DataPLAN maintains a consistent user interface. Internally, the answers are organized and structured differently according to the

selected template. This provides a streamlined user experience, eliminating the need for users to navigate different entry masks for each template.

4.1.2. Content Comparison

A DMP is not only an RDM practice but also encompasses planning for all other RDM tasks, increasing the level of interconnectedness [15]. For example, DMP questionnaires from funding bodies prompt users to consider data sharing even before data collection begins. This holistic approach reminds users to avoid using unsuitable tools (proprietary or non-FAIR) during data collection. DataPLAN and DAMAP go beyond providing a blank canvas for users to fill in their DMP responses by offering prewritten answers for every question in the template (Figure 4, left panel). DataPLAN provides users with the flexibility to customize these prewritten answers to meet their project's needs (Figure 4), saving time and ensuring compliance with funding body guidelines. However, the incorporation of a new template into DataPLAN is more time-consuming. The questions must be aligned with the established question-and-answer corpus within DataPLAN. Then, possible answers must be generated and included to enable integration into the DataPLAN template system. Other DMP tools [2,18–20] do not offer prewritten templates because as tools (but not service providers), they are applicable to a wide range of research domains, which need to be self-hosted or customized to provide specific and practical RDM solutions. By focusing on the plant sciences, DataPLAN can provide more specific and practical answers to the questionnaires, helping researchers to create tailored DMPs. Furthermore, although the current DataPLAN template focuses on plant research, the tool can already be used by other research domains with similar methods and techniques.

Data Steward Wizard analyzes the FAIR-ness of its DMPs by assessing all user input, whereas DataPLAN provides RDM tools and platforms to improve FAIR-ness. RDMO, Data Steward Wizard and DataPLAN can also export DMP results in both human and machine-readable formats. All the European tools (Data Stewardship Wizard, DMP online, DMP tools, DMP Canvas Generator, DMPRoadmap, RDMO, DataWiz, ARGOS, DMPTY, easyDMP and DataPLAN) mention international agreements and regulations in their DMPs. DataPLAN includes warning notifications to make users aware of missing answers or problematic statements (e.g., those in conflict with the GDPR). It is important to be mindful of issues related to sensitive or personal information when creating a DMP because these can have significant legal and ethical implications. If users plan to collect or store personal information as part of their research, they must obtain written informed consent from individuals and must comply with relevant laws. Failing to do so could hinder the project. Another specific feature of DataPLAN is its handling of the Nagoya Protocol [109,110], an international agreement that defines specific requirements for the sharing of and access to research data, particularly regarding genetic resources. Researchers working on projects that involve materials from developing countries that have signed up to the Nagoya Protocol must understand and comply with its requirements to avoid potential violations. DataPLAN assists users by providing notifications to help address legal issues and ensure research is conducted ethically and in compliance with relevant laws and regulations.

DataPLAN is designed to integrate seamlessly with plant-focused RDM platforms to enhance data sharing, interoperability, and long-term preservation. The integration of DataPLAN with existing RDM infrastructures enables researchers to connect their DMPs with other RDM platforms. The current integration of DataPLANT [59,78] and its tools and concepts (such as ARC [79], Swate [80], ARC Commander [81], and DataHUB [82]) into DataPLAN templates can streamline and simplify data management for researchers. The tools and resources cover every stage of RDM, from data acquisition to publication. By providing links and guides for the use of such resources within DataPLAN, researchers can access the tools and resources more easily, allowing them to manage their data more effectively. DataPLAN also provides a step-by-step guide, helping researchers to use RDM platforms for data management throughout their research projects.

In summary, DataPLAN is unique because it focuses on plant science, has high user-friendliness [104], and can general multiple DMP documents at once. Compared to general tools which require customization, DataPLAN is both a tool and a service that can be used directly by the plant science community.

4.2. Outlook

The input and output of RDMO will be supported by DataPLAN in the future. Templates in RDMO and Data Stewardship Wizard will be also very helpful for future template development. Domain-specific templates such as biodiversity and emission RDMO templates [111,112] are helpful resources that will be referenced for additional template development. Input and output JSON will be compliant with maDMP.

To streamline data management, DataPLAN will enhance its integration with popular data analysis tools and platforms. This will enable researchers to seamlessly connect their DMPs with data analysis workflows, data visualization tools, and statistical analysis software. By linking DMPs with newly developed data analysis tools, researchers can ensure that their RDM practices align with the most advanced data analysis tools and processing requirements. DataPLAN will integrate DataPLANT Biology Ontology [85], DMP common standard ontology (DCSO) [113] and other relevant ontologies as a part of the DataPLANT data management platform.

External platforms such as GitHub or GitLab are currently used for collaborative communication in DataPLAN. This weakness is a disadvantage of pure frontend designs. The collaborative functions can be enhanced by adding an independent optional backend. DataPLAN will enhance its collaboration features, allowing researchers to collaborate on DMP development and maintenance without using GitHub or GitLab.

5. Conclusions

We have created DataPLAN, a user-friendly tool designed for the plant sciences that generates multiple DMP documents at once. Equipped with prewritten reusable answers and a single-page interface, DataPLAN enables researchers to create DMPs in minutes, regardless of their experience and expertise. We use a pure frontend design to prevent data transmission, enhancing data security and privacy. DataPLAN is an open-source and web-based tool, enabling customization by users and modification by developers. In the future, DataPLAN will be maintained and updated by DataPLANT and IBG-4 to integrate new technologies and to deepen the synchronization of evolving DataPLANT tools such as Swate, DataHub, and ARCCommander. Overall, DataPLAN is a valuable resource for researchers seeking to manage their data efficiently while maintaining high user-friendliness.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data are available in Supplementary Data S1, and the code is available on GitHub https://github.com/nfdi4plants/dataplan (accessed on 25 September 2023).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

ARC	annotated research context
DDBJ	DNA Data Bank of Japan
DFG	German Research Foundation (Deutsche Forschungsgemeinschaft)
DMP	data management plan
EBI	European Bioinformatics Institute
ENA	European Nucleotide Archive
EU	European Union
FAIR	findable, accessible, interoperable and reusable
GDPR	EU General Data Protection Regulation
GEO	Gene Expression Omnibus
MIAME	minimal information about a microarray experiment
MIAPE	minimum information about a proteomics experiment
MIAPPE	minimal information about plant phenotyping experiment
MinSEQe	minimum information about a high-throughput sequencing experiment
MSI	Metabolomics Standards Initiative
NCBI	National Center for Biotechnology Information
NFDI	National Research Data Infrastructure (of Germany)
PRIDE	Proteomics Identification Database
RDM	research data management
RNA-Seq	RNA sequencing
SRA	Sequence Read Archive

References

1. Jones, S.; Pergl, R.; Hooft, R.; Miksa, T.; Samors, R.; Ungvari, J.; Davis, R.I.; Lee, T. Data Management Planning: How Requirements and Solutions Are Beginning to Converge. *Data Intell.* **2020**, *2*, 208–219. [CrossRef]
2. Sallans, A.; Donnelly, M. DMP Online and DMPTool: Different Strategies Towards a Shared Goal. *Int. J. Digit. Curation* **2012**, *7*, 123–129. [CrossRef]
3. Miksa, T.; Simms, S.; Mietchen, D.; Jones, S. Ten Principles for Machine-Actionable Data Management Plans. *PLoS Comput. Biol.* **2019**, *15*, e1006750. [CrossRef] [PubMed]
4. Stewart, A.J.; Farran, E.K.; Grange, J.A.; Macleod, M.; Munafò, M.; Newton, P.; Shanks, D.R.; UKRN Institutional Leads. Improving Research Quality: The View from the UK Reproducibility Network Institutional Leads for Research Improvement. *BMC Res. Notes* **2021**, *14*, 458. [CrossRef]
5. Wilkinson, M.D.; Dumontier, M.; Aalbersberg, I.J.J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.-W.; da Silva Santos, L.B.; Bourne, P.E.; et al. The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Sci. Data* **2016**, *3*, 160018. [CrossRef]
6. Du, X.; Dastmalchi, F.; Ye, H.; Garrett, T.J.; Diller, M.A.; Liu, M.; Hogan, W.R.; Brochhausen, M.; Lemas, D.J. Evaluating LC-HRMS Metabolomics Data Processing Software Using FAIR Principles for Research Software. *Metabolomics* **2023**, *19*, 11. [CrossRef]
7. Gajbe, S.B.; Tiwari, A.; Gopalji; Singh, R.K. Evaluation and Analysis of Data Management Plan Tools: A Parametric Approach. *Inf. Process. Manag.* **2021**, *58*, 102480. [CrossRef]
8. Vieira, A. How to Comply with Horizon Europe Mandate for RDM. Available online: https://www.openaire.eu/how-to-comply-with-horizon-europe-mandate-for-rdm (accessed on 25 September 2023).

9. Handling of Research Data. Available online: https://www.dfg.de/en/research_funding/principles_dfg_funding/research_data/ (accessed on 25 September 2023).

10. NOT-OD-21-013: Final NIH Policy for Data Management and Sharing. Available online: https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html (accessed on 25 September 2023).

11. Preparing Your Data Management Plan. Available online: https://new.nsf.gov/funding/data-management-plan (accessed on 25 September 2023).

12. Data Management Plan (DMP)—Guidelines for researchers. Available online: https://www.snf.ch/en/FAiWVH4WvpKvohw9/topic/research-policies (accessed on 25 September 2023).

13. Monastersky, R. Publishing Frontiers: The Library Reboot. *Nature* **2013**, *495*, 430–432. [CrossRef]

14. Tenopir, C.; Birch, B.; Allard, S. *Academic Libraries and Research Data Services: Current Practices and Plans for the Future*; Association of College and Research Libraries: Chicago, IL, USA, 2012.

15. Sheikh, A.; Malik, A.; Adnan, R. Evolution of Research Data Management in Academic Libraries: A Review of the Literature. *Inf. Dev.* **2023**. [CrossRef]

16. Datenmanagementpläne. Available online: https://www.fu-berlin.de/sites/forschungsdatenmanagement/materialien/handreichungen/dmp/index.html (accessed on 25 September 2023).

17. Muster Datenmanagementplan für einen DFG-Antrag. Available online: https://www.cms.hu-berlin.de/de/dl/dataman/muster-dmp-dfg/view (accessed on 25 September 2023).

18. Pergl, R.; Hooft, R.; Suchánek, M.; Knaisl, V.; Slifka, J. "data Stewardship Wizard": A Tool Bringing Together Researchers, Data Stewards, and Data Experts around Data Management Planning. *Data Sci. J.* **2019**, *18*, 59. [CrossRef]

19. Getler, M.; Sisu, D.; Jones, S.; Miller, K. DMPonline Version 4.0: User-Led Innovation. *Int. J. Digit. Curation* **2014**, *9*, 193–219. [CrossRef]

20. Engelhardt, C.; Enke, H.; Klar, J.; Ludwig, J.; Neuroth, H. Research Data Management Organiser. In Proceedings of the 19th Conference EGU General Assembly, EGU2017, Vienna, Austria, 23–28 April 2017; p. 15760.

21. Research Data Manager (RDM). Available online: https://research.uq.edu.au/rmbt/uqrdm (accessed on 25 September 2023).

22. Blask, K.; Bölter, R. DataWiz. Available online: https://datawiz.leibniz-psychology.org/DataWiz/ (accessed on 25 September 2023).

23. Simpson, P.W. *Argos*; Month9books: Raleigh, NC, USA, 2016; ISBN 9780996890434.

24. Sigma, U. EasyDMP. Available online: https://easydmp.sigma2.no/ (accessed on 25 September 2023).

25. Trippel, T.; Zinn, C. DMPTY—A Wizard for Generating Data Management Plans. In Proceedings of the Selected Papers from the CLARIN Annual Conference 2015, Wroclaw, Poland, 14–16 October 2015; Linköping University Electronic Press: Linköping, Sweden, 2015; pp. 71–78.

26. UWA Library Guides: Research Data Management Toolkit: Welcome. Available online: https://guides.library.uwa.edu.au/RDMtoolkit (accessed on 25 September 2023).

27. Lehnert, K.; Ferrini, V.L.; Berman, H.; Gabanyi, M.; Stodden, V.; Morton, J.J. ezDMP: Data Management Planning Made Easy. In Proceedings of the AGU Fall Meeting, Washington DC, USA, 10–14 December 2018; Volume 2018, p. ED53C-01.

28. Neuroth, H.; Engelhardt, C.; Klar, J.; Ludwig, J.; Enke, H. Aktives Forschungsdatenmanagement. *ABI Tech.* **2018**, *38*, 55–64. [CrossRef]

29. Riley, B.; Rust, S.; Morrice, G.; Carrick, R. Roadmap: DCC/UC3 Collaboration for a Data Management Planning Tool. Available online: https://github.com/DMPRoadmap/roadmap (accessed on 25 September 2023).

30. Klar, J.; Michaelis, O.; Wallace, D.; Schröder, M.; Fütterer, H.; Lanza, G.; Martínez Muñoz, D.; PiIori, D.; Harry, E. Rdmo: A Tool to Support the Planning, Implementation, and Organization of Research Data Management. Available online: https://github.com/rdmorganiser/rdmo (accessed on 25 September 2023).

31. EDITORIAL. Everyone Needs a Data-Management Plan. *Nature* **2018**, *555*, 286. [CrossRef]

32. Leibniz-Institute for Psychology Information DataWiz: An Automated Assistant for the Management of Psychological Research Data. Available online: https://github.com/ZPID/DataWiz (accessed on 25 September 2023).

33. Klar, J.; Engelhardt, C.; Neuroth, H.; Enke, H.; Ludwig, J. RDMO—Research Data Management Organiser. In Proceedings of the EGU General Assembly, Vienna, Austria, 23–28 April 2017; p. 15760.

34. Miksa, T.; Oblasser, S.; Rauber, A. Automating Research Data Management Using Machine-Actionable Data Management Plans. *ACM Trans. Manag. Inf. Syst.* **2022**, *13*, 1–22. [CrossRef]

35. DMPTool. Available online: https://dmptool.org/ (accessed on 25 September 2023).

36. Fabry, C. Nouvelle version de DMP OPIDoR: Vers un DMP Machine Actionnable. Available online: https://zenodo.org/records/6760990 (accessed on 25 September 2023).

37. ARGOS Dmp. Available online: https://gitlab.eudat.eu/dmp (accessed on 25 September 2023).

38. EasyDMP—NIRD Data Planning. Available online: https://www.sigma2.no/data-planning (accessed on 25 September 2023).

39. Blumesberger, S.; Gänsdorfer, N.; Ganguly, R.; Gergely, E.; Gruber, A.; Hasani-Mavriqi, I.; Kalová, T.; Ladurner, C.; Macher, T.; Miksa, T.; et al. FAIR Data Austria—Abstimmung der Implementierung von FAIR Tools und Services. *Mitteilungen VÖB* **2021**, *74*, 102–120. [CrossRef]

40. Diepenbroek, M.; Glöckner, F.O.; Grobe, P.; Güntsch, A.; Huber, R.; König-Ries, B.; Kostadinov, I.; Nieschulze, J.; Seeger, B.; Tolksdorf, R.; et al. Towards an Integrated Biodiversity and Ecological Research Data Management and Archiving Platform: The German Federation for the Curation of Biological Data (GFBio). *Informatik* **2014**, 1711–1721.

41. Zheng, Y. Methodologies for Cross-Domain Data Fusion: An Overview. *IEEE Trans. Big Data* **2015**, *1*, 16–34. [CrossRef]

42. Hannemann, J.; Poorter, H.; Usadel, B.; Bläsing, O.E.; Finck, A.; Tardieu, F.; Atkin, O.K.; Pons, T.; Stitt, M.; Gibon, Y. Xeml Lab: A Tool That Supports the Design of Experiments at a Graphical Interface and Generates Computer-Readable Metadata Files, Which Capture Information about Genotypes, Growth Conditions, Environmental Perturbations and Sampling Strategy. *Plant Cell Environ.* **2009**, *32*, 1185–1200. [CrossRef]

43. Papoutsoglou, E.A.; Faria, D.; Arend, D.; Arnaud, E.; Athanasiadis, I.N.; Chaves, I.; Coppens, F.; Cornut, G.; Costa, B.V.; Ćwiek-Kupczyńska, H.; et al. Enabling Reusability of Plant Phenomic Datasets with MIAPPE 1.1. *New Phytol.* **2020**, *227*, 260–273. [CrossRef]

44. Taylor, C.F. Minimum Reporting Requirements for Proteomics: A MIAPE Primer. *Proteomics* **2006**, *6*, 39–44. [CrossRef]

45. Brazma, A.; Hingamp, P.; Quackenbush, J.; Sherlock, G.; Spellman, P.; Stoeckert, C.; Aach, J.; Ansorge, W.; Ball, C.A.; Causton, H.C.; et al. Minimum Information about a Microarray Experiment (MIAME)—Toward Standards for Microarray Data. *Nat. Genet.* **2001**, *29*, 365–371. [CrossRef]

46. Arend, D.; Lange, M.; Chen, J.; Colmsee, C.; Flemming, S.; Hecht, D.; Scholz, U. e!DAL—A Framework to Store, Share and Publish Research Data. *BMC Bioinform.* **2014**, *15*, 214. [CrossRef] [PubMed]

47. Von Suchodoletz, D.; Mühlhaus, T.; Brillhaus, D.; Tschöpe, M.; Maus, O.; Grüning, B.; Garth, C.; Rodrigues, C.M. DataPLANT—Tools and Services to structure the Data Jungle for fundamental plant researchers. In *E-Science-Tage 2021: Share Your Research Data*; Vincent Heuveline, N.B., Ed.; heiBOOKS: Heidelberg, Germany, 2022; pp. 132–145. ISBN 9783948083540.

48. Arsova, B.; Foster, K.J.; Shelden, M.C.; Bramley, H.; Watt, M. Dynamics in Plant Roots and Shoots Minimize Stress, Save Energy and Maintain Water and Nutrient Uptake. *New Phytol.* **2020**, *225*, 1111–1119. [CrossRef] [PubMed]

49. Watt, M.; Fiorani, F.; Usadel, B.; Rascher, U.; Muller, O.; Schurr, U. Phenotyping: New Windows into the Plant for Breeders. *Annu. Rev. Plant Biol.* **2020**, *71*, 689–712. [CrossRef] [PubMed]

50. Bar-On, Y.M.; Phillips, R.; Milo, R. The Biomass Distribution on Earth. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 6506–6511. [CrossRef]

51. Lobet, G.; Pound, M.P.; Diener, J.; Pradal, C.; Draye, X.; Godin, C.; Javaux, M.; Leitner, D.; Meunier, F.; Nacry, P.; et al. Root system markup language: Toward a unified root architecture description language. *Plant Physiol.* **2015**, *167*, 617–627. [CrossRef]

52. Zhu, X.G.; Long, S.P.; Ort, D.R. What Is the Maximum Efficiency with Which Photosynthesis Can Convert Solar Energy into Biomass? *Curr. Opin. Biotechnol.* **2008**, *19*, 153–159. [CrossRef]

53. Bolger, M.; Schwacke, R.; Gundlach, H.; Schmutzer, T.; Chen, J.; Arend, D.; Oppermann, M.; Weise, S.; Lange, M.; Fiorani, F.; et al. From Plant Genomes to Phenotypes. *J. Biotechnol.* **2017**, *261*, 46–52. [CrossRef]

54. Cantelli, G.; Bateman, A.; Brooksbank, C.; Petrov, A.I.; Malik-Sheriff, R.S.; Ide-Smith, M.; Hermjakob, H.; Flicek, P.; Apweiler, R.; Birney, E.; et al. The European Bioinformatics Institute (EMBL-EBI) in 2021. *Nucleic Acids Res.* **2022**, *50*, D11–D19. [CrossRef]

55. Marks, R.A.; Amézquita, E.J.; Percival, S.; Rougon-Cardoso, A.; Chibici-Revneanu, C.; Tebele, S.M.; Farrant, J.M.; Chitwood, D.H.; VanBuren, R. A Critical Analysis of Plant Science Literature Reveals Ongoing Inequities. *Proc. Natl. Acad. Sci. USA* **2023**, *120*, e2217564120. [CrossRef]

56. Arend, D.; Junker, A.; Scholz, U.; Schüler, D.; Wylie, J.; Lange, M. PGP Repository: A Plant Phenomics and Genomics Data Publication Infrastructure. *Database* **2016**, *2016*, baw033. [CrossRef]

57. Arend, D.; Psaroudakis, D.; Memon, J.A.; Rey-Mazón, E.; Schüler, D.; Szymanski, J.J.; Scholz, U.; Junker, A.; Lange, M. From Data to Knowledge—Big Data Needs Stewardship, a Plant Phenomics Perspective. *Plant J.* **2022**, *111*, 335–347. [CrossRef] [PubMed]

58. Agrahari, R.K.; Singh, P.; Koyama, H.; Panda, S.K. Plant-Microbe Interactions for Sustainable Agriculture in the Post-Genomic Era. *Curr. Genom.* **2020**, *21*, 168–178. [CrossRef] [PubMed]

59. von Suchodoletz, D.; Mühlhaus, T.; Krüger, J.; Usadel, B.; Rodrigues, C.M. DataPLANT—Ein NFDI-Konsortium der Pflanzen-Grundlagenforschung. *BFDM* **2021**, *2*, 46–56.

60. Specka, X.; Martini, D.; Weiland, C.; Arend, D.; Asseng, S.; Boehm, F.; Feike, T.; Fluck, J.; Gackstetter, D.; Gonzales-Mellado, A.; et al. FAIRagro: Ein Konsortium in Der Nationalen Forschungsdateninfrastruktur (NFDI) Für Forschungsdaten in Der Agrosystemforschung: Herausforderungen und Lösungsansätze für den Aufbau einer FAIRen Forschungsdateninfrastruktur. *Informatik* **2023**, *46*, 24–35. [CrossRef]

61. Plant Sciences Community. Available online: https://elixir-europe.org/communities/plant-sciences (accessed on 25 September 2023).

62. Leonelli, S.; Davey, R.P.; Arnaud, E.; Parry, G.; Bastow, R. Data Management and Best Practice for Plant Science. *Nat. Plants* **2017**, *3*, 17086. [CrossRef]

63. Krantz, M.; Zimmer, D.; Adler, S.O.; Kitashova, A.; Klipp, E.; Mühlhaus, T.; Nägele, T. Data Management and Modeling in Plant Biology. *Front. Plant Sci.* **2021**, *12*, 717958. [CrossRef]

64. Sansone, S.-A.; Rocca-Serra, P.; Brandizi, M.; Brazma, A.; Field, D.; Fostel, J.; Garrow, A.G.; Gilbert, J.; Goodsaid, F.; Hardy, N.; et al. The First RSBI (ISA-TAB) Workshop: "Can a Simple Format Work for Complex Studies? *OMICS* **2008**, *12*, 143–149. [CrossRef]

65. Rocca-Serra, P.; Brandizi, M.; Maguire, E.; Sklyar, N.; Taylor, C.; Begley, K.; Field, D.; Harris, S.; Hide, W.; Hofmann, O.; et al. ISA Software Suite: Supporting Standards—Compliant Experimental Annotation and Enabling Curation at the Community Level. *Bioinformatics* **2010**, *26*, 2354–2356. [CrossRef]

66. Amstutz, P.; Crusoe, M.R.; Tijanić, N.; Chapman, B.; Chilton, J.; Heuer, M.; Kartashov, A.; Leehr, D.; Ménager, H.; Nedeljkovich, M.; et al. Common Workflow Language, v1.0. Available online: https://research.manchester.ac.uk/files/57032695/cwl_1.0_tool.pdf. (accessed on 25 September 2023).

67. Mason, P.G.; Barratt, B.I.P.; Mc Kay, F.; Klapwijk, J.N.; Silvestri, L.C.; Hill, M.; Hinz, H.L.; Sheppard, A.; Brodeur, J.; Vitorino, M.D.; et al. Impact of Access and Benefit Sharing Implementation on Biological Control Genetic Resources. *Biocontrol* **2023**, *68*, 235–251. [CrossRef]
68. GFBio e.V FAR-DSI: Feasibility Assessment of Regulation for Digital Sequence Information. Available online: https://www.gfbio.org/gfbio_ev/far-dsi-project/ (accessed on 25 September 2023).
69. European Commission Data Management—H2020 Online Manual. Available online: https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm (accessed on 25 September 2023).
70. Garfolo, B.T. JavaScript. In *Encyclopedia of Information Systems*; Elsevier: Amsterdam, The Netherlands, 2003; pp. 715–735. ISBN 9780122272400.
71. Otto, M.; Thornton, J. Bootstrap. Available online: https://getbootstrap.com/ (accessed on 25 September 2023).
72. Yaras bs5-Intro-Tour: Extension for Bootstrap 5 Which Allows to Build Intro Tours. Available online: https://github.com/yaras6/bs5-intro-tour (accessed on 25 September 2023).
73. Davies, J. d3-Cloud: Create Word Clouds in JavaScript. Available online: https://github.com/jasondavies/d3-cloud (accessed on 25 September 2023).
74. Grey, E. FileSaver.js: An HTML5 saveAs() FileSaver Implementation. Available online: https://github.com/eligrey/FileSaver.js (accessed on 25 September 2023).
75. Split.js. Available online: https://split.js.org/ (accessed on 25 September 2023).
76. Performance: Measure() Method. Available online: https://developer.mozilla.org/en-US/docs/Web/API/Performance/measure (accessed on 25 September 2023).
77. Lighthouse Overview. Available online: https://developer.chrome.com/docs/lighthouse/overview/ (accessed on 25 September 2023).
78. von Suchodoletz, D.; Krüger, J.; Mühlhaus, T.; Usadel, B.; Gauza, H.; Rodrigues, C.M. *Data Stewards as Ambassadors between the NFDI and the Community*; Universitätsbibliothek: Heidelberg, Germany, 2021.
79. Mühlhaus, T.; Garth, C.; Brilhaus, D.; Von Suchodoletz, D. ARC-Specification. Available online: https://github.com/nfdi4plants/ARC-specification (accessed on 25 September 2023).
80. Frey, K. Swate: Excel Add-in for Annotation of Experimental Data and Computational Workflows. Available online: https://github.com/nfdi4plants/Swate (accessed on 25 September 2023).
81. Weil, L.; Maus, O. ARCCommander: Tool to Manage Your ARCs. Available online: https://github.com/nfdi4plants/arcCommander (accessed on 25 September 2023).
82. Weil, H.L.; Schneider, K.; Tschöpe, M.; Bauer, J.; Maus, O.; Frey, K.; Brilhaus, D.; Martins Rodrigues, C.; Doniparthi, G.; Wetzels, F.; et al. PLANTdataHUB: A Collaborative Platform for Continuous FAIR Data Sharing in Plant Research. *Plant J.* **2023**. [CrossRef] [PubMed]
83. Rustici, G.; Williams, E.; Barzine, M.; Brazma, A.; Bumgarner, R.; Chierici, M.; Furlanello, C.; Greger, L.; Jurman, G.; Miller, M.; et al. Transcriptomics Data Availability and Reusability in the Transition from Microarray to next-Generation Sequencing. Available online: https://www.biorxiv.org/content/biorxiv/early/2021/01/03/2020.12.31.425022 (accessed on 25 September 2023).
84. Fiehn, O.; Sumner, L.W.; Rhee, S.Y.; Ward, J.; Dickerson, J.; Lange, B.M.; Lane, G.; Roessner, U.; Last, R.; Nikolau, B. Minimum Reporting Standards for Plant Biology Context Information in Metabolomic Studies. *Metabolomics* **2007**, *3*, 195–201. [CrossRef]
85. Dumschott, K.; Brilhaus, D.; Tschöpe, M. nfdi4plants_ontology: A Intermediate Ontology for Plants Used by DataPLANT to Fill the Ontology Gap. Available online: https://github.com/nfdi4plants/nfdi4plants_ontology (accessed on 25 September 2023).
86. Brazma, A.; Ball, C.; Bumgarner, R.; Furlanello, C.; Miller, M.; Quackenbush, J.; Reich, M.; Rustici, G.; Stoeckert, C.; Trutane, S.C.; et al. MINSEQE: Minimum Information about a High-throughput Nucleotide SeQuencing Experiment—A Proposal for Standards in Functional Genomic Data Reporting. Available online: https://zenodo.org/record/5706412 (accessed on 25 September 2023).
87. Li, W.; Cowley, A.; Uludag, M.; Gur, T.; McWilliam, H.; Squizzato, S.; Park, Y.M.; Buso, N.; Lopez, R. The EMBL-EBI Bioinformatics Web and Programmatic Tools Framework. *Nucleic Acids Res.* **2015**, *43*, W580–W584. [CrossRef] [PubMed]
88. Barrett, T.; Wilhite, S.E.; Ledoux, P.; Evangelista, C.; Kim, I.F.; Tomashevsky, M.; Marshall, K.A.; Phillippy, K.H.; Sherman, P.M.; Holko, M.; et al. NCBI GEO: Archive for Functional Genomics Data Sets—Update. *Nucleic Acids Res.* **2012**, *41*, D991–D995. [CrossRef]
89. Miyazaki, S.; Sugawara, H.; Ikeo, K.; Gojobori, T.; Tateno, Y. DDBJ in the Stream of Various Biological Data. *Nucleic Acids Res.* **2004**, *32*, D31–D34. [CrossRef]
90. Kodama, Y.; Shumway, M.; Leinonen, R.; International Nucleotide Sequence Database Collaboration. The Sequence Read Archive: Explosive Growth of Sequencing Data. *Nucleic Acids Res.* **2012**, *40*, D54–D56. [CrossRef]
91. Benson, D.A.; Cavanaugh, M.; Clark, K.; Karsch-Mizrachi, I.; Lipman, D.J.; Ostell, J.; Sayers, E.W. GenBank. *Nucleic Acids Res.* **2013**, *41*, D36–D42. [CrossRef]
92. Hermjakob, H.; Apweiler, R. The Proteomics Identifications Database (PRIDE) and the ProteomExchange Consortium: Making Proteomics Data Accessible. *Expert Rev. Proteom.* **2006**, *3*, 1–3. [CrossRef]
93. Steinbeck, C.; Conesa, P.; Haug, K.; Mahendraker, T.; Williams, M.; Maguire, E.; Rocca-Serra, P.; Sansone, S.-A.; Salek, R.M.; Griffin, J.L. MetaboLights: Towards a New COSMOS of Metabolomics Data Management. *Metabolomics* **2012**, *8*, 757–760. [CrossRef]
94. GnpIS. Available online: https://urgi.versailles.inra.fr/gnpis (accessed on 25 September 2023).

95. Weise, S.; Oppermann, M.; Maggioni, L.; van Hintum, T.; Knüpffer, H. EURISCO: The European Search Catalogue for Plant Genetic Resources. *Nucleic Acids Res.* **2017**, *45*, D1003–D1008. [CrossRef]

96. RDA DMP Common Standard for Machine-Actionable Data Management Plans. Available online: https://zenodo.org/records/4036060 (accessed on 25 September 2023).

97. Sherman, B.; Henry, R.J. The Nagoya Protocol and Historical Collections of Plants. *Nat. Plants* **2020**, *6*, 430–432. [CrossRef] [PubMed]

98. Voigt, P.; von dem Bussche, A. *The EU General Data Protection Regulation (GDPR)*; Springer International Publishing: Cham, Switzerland, 2016.

99. Barker, M.; Chue Hong, N.P.; Katz, D.S.; Lamprecht, A.-L.; Martinez-Ortiz, C.; Psomopoulos, F.; Harrow, J.; Castro, L.J.; Gruenpeter, M.; Martinez, P.A.; et al. Introducing the FAIR Principles for Research Software. *Sci. Data* **2022**, *9*, 622. [CrossRef] [PubMed]

100. Zhou, X. Dataplan: DataPLAN Is the Data Management Plan (DMP) Generator Developed in DataPLANT. Available online: https://github.com/nfdi4plants/dataplan (accessed on 25 September 2023).

101. D'Anna, F.; Faria, D. Your Tasks: Data Management Plan. Available online: https://rdmkit.elixir-europe.org/data_management_plan (accessed on 25 September 2023).

102. Ison, J.; Ienasescu, H.; Chmura, P.; Rydza, E.; Ménager, H.; Kalaš, M.; Schwämmle, V.; Grüning, B.; Beard, N.; Lopez, R.; et al. The Bio.tools Registry of Software Tools and Data Resources for the Life Sciences. *Genom. Biol.* **2019**, *20*, 164. [CrossRef] [PubMed]

103. Hasselbring, W.; Carr, L.; Hettrick, S.; Packer, H.; Tiropanis, T. From FAIR Research Data toward FAIR and Open Research Software. *It-Inf. Technol.* **2020**, *62*, 39–47. [CrossRef]

104. Becker, C.; Hundt, C.; Engelhardt, C.; Sperling, J.; Kurzweil, M.; Müller-Pfefferkorn, R. Data Management Plan Tools: Overview and Evaluation. *Proc. Conf. Res. Data Infrastruct.* **2023**, *1*, CoRDI2023-96. [CrossRef]

105. Donnelly, M.; Jones, S.; Pattenden-Fail, J.W. DMP Online: The Digital Curation Centre's Web-Based Tool for Creating, Maintaining and Exporting Data Management Plans. *Int. J. Digit. Curation* **2010**, *5*, 187–193. [CrossRef]

106. Rice, R.; Fergusson, D. *LEARN Toolkit of Best Practice for Research Data Management*; Research Data Management at the University of Edinburgh: How is it done, what does it costs? CS17; UCL: London, UK, 2017; pp. 89–93.

107. Suchánek, M.; Knaisl, V.; Pergl, R. Ds-Wizard: DSW Common Repository. Available online: https://github.com/ds-wizard/ds-wizard (accessed on 25 September 2023).

108. SIB Swiss Institute of Bioinformatics/Vital-IT DMP Canvas Generator. Available online: https://dmp.vital-it.ch/#/login (accessed on 25 September 2023).

109. Morgera, E.; Tsioumani, E.; Buck, M. *Unraveling the Nagoya Protocol: A Commentary on the Nagoya Protocol on Access and Benefit-Sharing to the Convention on Biological Diversity*; Martinus Nijhoff Publishers: Leiden, The Netherlands, 2014; ISBN 9789004217188.

110. Rourke, M.; Eccleston-Turner, M. The Pandemic Influenza Preparedness Framework as a "specialized International Access and Benefit-Sharing Instrument" under the Nagoya Protocol. *N. Ir. Legal Q.* **2021**, *72*, 411–447. [CrossRef]

111. Rothe, R.; Lindstädt, B. RDMO4Life im Projekt EmiMin—Die Anpassung von Datenmanagementplänen an lebenswissenschaftliche Fachspezifika. Available online: https://opus4.kobv.de/opus4-bib-info/frontdoor/index/index/docId/16229 (accessed on 25 September 2023).

112. GFBio e.V GFBio Data Management Plan Tool. Available online: https://www.gfbio.org/plan/ (accessed on 25 September 2023).

113. Cardoso, J.; Castro, L.J.; Ekaputra, F.J.; Jacquemot, M.C.; Suchánek, M.; Miksa, T.; Borbinha, J. DCSO: Towards an Ontology for Machine-Actionable Data Management Plans. *J. Biomed. Semant.* **2022**, *13*, 21. [CrossRef]