Registration as a resource type in Datacite schema

About Datacite
Scholarly research is producing ever-increasing amounts of digital research data, where establishment of persistent identifiers and metadata describing each research finding is essential for long term preservation, findability, sharing, reuse, and discovery of scholarship. Datacite defines a persistent identifier as an association between a character string and a resource. Resources can be files, parts of files, persons, organisations, abstractions (concepts in a taxonomy, e.g.), and so on.

About COS and OSF
The Center for Open Science is a culture change organization that aims to align scholarly values with scholarly practices. It does so by developing and advancing a systems strategy for changing culture and behavior toward greater rigor, transparency, and sharing of research process, outputs, and outcomes.

To scale adoption of open behaviors by researchers, COS [1] provides open infrastructure that makes it possible to do the behaviors with the OSF, [2] conducts user-centered product development to make it easy to do the behaviors using OSF, [3] supports grassroots organizing to activate early adopters and make their behavior visible to shift community norms toward the behaviors, [4] offers solutions to journals, funders, and institutions to nudge their incentives to make it desirable to do the behaviors, and [5] provides and promotes a policy framework for stakeholders to make the behaviors required. Effective policy implementation requires effective infrastructure for doing the behaviors, and community buy-in to treat the behaviors as good practice rather than administrative burdens. These five levels of intervention are highly interdependent, each necessary, none sufficient.

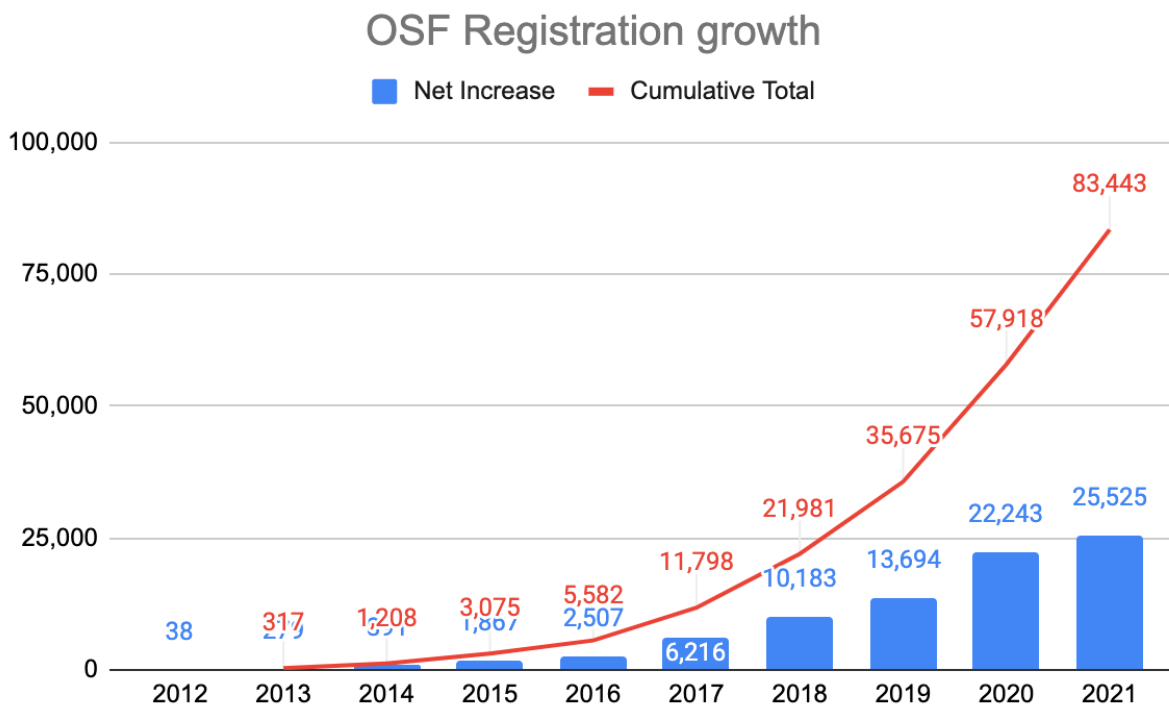**About Registrations and Preregistrations**
Preregistration is an explicit step in the research process where the research team specifies the research plan in advance of the study and submits it to a registry. The preregistration captures the study design, hypotheses, methods, and also often includes the pre-analysis plan, where the preregistration explicitly separates the hypothesis-generating (exploratory) from hypothesis-testing (confirmatory) research. This differentiation avoids conflicts as the same data cannot be used to generate and test the hypothesis, and by making this distinction in the planning, the quality and transparency of the research can be improved. The practice of preregistration increases the rigor of research, and decreases problematic research practices, including p-hacking, publication bias, data dredging, inappropriate post hoc analysis and HARKing.

There are many Registries that enable the creation of a preregistration, or registration, which is a time-stamped immutable declaration of the research plan prior to beginning the study. OSF supports the entire research lifecycle and continually works with communities to offer many templates for registering studies. The workflow on OSF is known as 'Registration', where a time-

stamped immutable declaration of the research is submitted to the OSF Registry or one of the Community Operated Registries. The registration workflow includes metadata about the study (title, abstract, contributors, license, discipline, institutional affiliation, tags) and responses to the registration template.

The registration template specifies the type of registration based on questions and the timing of the registration object in relation to data collection. A preregistration type of registration occurs prior to data collection, and can use different preregistration templates to support study methodology, such as the Qualitative Analysis Preregistration, Secondary Data Preregistration, and Preregistration in Social Psychology. Other templates offer more flexibility for archiving the research, reporting outcomes, or capturing changes to the research plan over time, such as the Open-Ended Registration template. Another registration workflow known as Registered Reports, which seeks to align scientific values and practices, leverages peer review on preregistered study plans before results are known. Registered Reports are a publishing format in which methods and proposed analyses are preregistered and peer reviewed prior to the research being conducted. In principle acceptance is given by the journal, along with peer feedback aimed at improving research rigor before the research commences.

Registration practices began in 2012 on OSF through a workflow where users can register their projects by completing a registration template and archiving the files. Since then many enhancement have been made to offer more standardardized templates and a more streamlined workflow. Since 2012, registrations have continued to grow, expecting to crest the 100,000 cumulative total in 2022.

OSF Registration growth

Current metadata
The current Datacite metadata schema doesn't offer an object type for registrations/preregistrations, making it difficult to pull them out as a unique step in the research process and connected to other outputs.

From the most recent version of the schema the current resource_type_general list the options are:
10.a resourceTypeGeneral
The general type of a resource Controlled List Values:
> Audiovisual
> Book
> BookChapter
> Collection
> Computational
> Notebook
> ConferencePaper
> ConferenceProceeding
> DataPaper
> Dataset
> Dissertation
> Event
> Image
> InteractiveResource
> Journal
> JournalArticle
> Model
> OutputManagementPlan
> PeerReview
> PhysicalObject
> Preprint
> **Report (used by Animal Registry)**
> Service
> Software
> Sound
> Standard
> Text
> Workflow
> **Other (used by OSF registrations)**

The addition of 'registration' to the type list would enable the minimum elements needed to capture registrations as a unique type of research output.

Within the 'registration' type data field, the ability to annotate that with the type of registration list:

>Preregistration
>Results Registration
>Registered Report
>Research Project Archive

Within the 'type of registration' list, an additional data field (text) to annotate a specific template with a schema for the template could provide additional specification:

>OSF Preregistration
>>https://api.osf.io/v2/schemas/registrations/5e795fc0d2833800195735d0/
>Secondary Data Analysis Preregistration
>AsPredicted Preregistration
>>https://api.osf.io/v2/schemas/registrations/5d2d2268d28338002c2432d2/
>Registered Report Protocol
>>https://api.osf.io/v2/schemas/registrations/5e795fc0d2833800195735cf/
>Pre-Data Collection Registration
>>https://api.osf.io/v2/schemas/registrations/564d31db8c5e4a7c9694b2c0/
>Qualitative Preregistration
>>https://api.osf.io/v2/schemas/registrations/5fa0ac510a7f38001c8ae854/
>Replication Recipe Preregistration
>>https://api.osf.io/v2/schemas/registrations/564d31db8c5e4a7c9694b2c2/
>Preplication Recipe Results Registration
>>https://api.osf.io/v2/schemas/registrations/564d31db8c5e4a7c9694b2c4/
>Social Psychology Preregistration
>>https://api.osf.io/v2/schemas/registrations/5730e99a9ad5a102c5745a8a/
>Project Archive (Open-ended Registration)
>>https://api.osf.io/v2/schemas/registrations/5df83f7dd28338001ac0ab0d/

The Animal Study Registry operated by the German Centre for the Protection of Laboratory Animals (Bf3R) at the German Federal Institute for Risk Assessment (BfR) is currently publishing their preregistrations as "Reports". The metadata desired on their study registrations:

- identifier
- creatorName
- givenName
- familyName
- title
- publisher
- publicationYear
- **resourceType**
- subject
- contributorName
- PublicationDate
- EmbargoDate: "Notified date of the end of the embargo"

- UpdateDate: "Date of the last update"
- language
- size
- format
- version
- rights
- description
- attachments: (yes/no/ number of attachments uploaded with the preregistration)

Specifically calling out the embargo date for registrations as an time to be included in the registration type metadata fields.