

## Plan Overview

---

*A Data Management Plan created using DMPTool*

**DMP ID:** <https://doi.org/10.48321/D18G89>

**Title:** Developments of the integrative hypothesis of specialization

**Creator:** Marco Mello - **ORCID:** [0000-0002-9098-9427](https://orcid.org/0000-0002-9098-9427)

**Affiliation:** Universidade de São Paulo ([www5.usp.br](http://www5.usp.br))

**Contributor:** Carsten Dormann, Sharlene Santana, Renata Muylaert, Cristina Kita, Sebastian Montoya-Bustamante, Natalya Zapata-Mesa, Nastaran Lotfi, Tanja Maria Straka, Luiza Santos Reis

**Funder:** São Paulo Research Foundation ([fapesp.br](http://fapesp.br))

**Template:** Digital Curation Centre ([português](http://português))

### **Project abstract:**

At the Ecological Synthesis Laboratory (SintECO), we are committed to studying interactions between organisms of different species. With an approach based on network science, in our previous project funded by FAPESP, we developed an innovative cognitive model, the integrative hypothesis of specialization (IHS), which sheds light on the assembly rules of interaction networks. With a successful proof of concept, our model proved to be able to explain how the four main topologies observed in interaction networks in nature arise. Now, in a new phase of the project, we intend to consolidate the IHS and develop it even further. First, we will present the IHS as a discursive model, that is, an efficient theory in the strict sense, framed in the semantic epistemological view. Second, we will explore some conceptual and operational developments of the IHS, aiming to create new tools for analysis of interaction networks. Third, we will investigate how IHS can help us to study emergence in interaction networks. Thus, we will consolidate a new theory that can be used as a tool not only for the study of interspecific interactions, but also of ecosystem services. With these next steps, we are committing to advance our line of research on the assembly rules of complex ecological systems, which is already starting to have a significant impact in several fields of study. Our project also has the potential

to help achieve some of the Sustainable Development Goals proposed by the United Nations.

**Start date:** 01-01-2024

**End date:** 12-31-2026

**Last modified:** 08-08-2023

**Copyright information:**

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customize it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

---

## Developments of the integrative hypothesis of specialization

### Coleta de Dados

---

#### Que dados serão coletados ou criados?

Serão compilados dados sobre interações ecológicas entre organismos de espécies diferentes (interações interespecíficas), tanto positivas quanto negativas, abrangendo diversos táxons. Os principais tipos de interações em que focaremos são a polinização, a dispersão de sementes, a herbivoria, o parasitismo e as doenças.

#### Como os dados serão coletados ou criados?

Os dados serão compilados a partir da literatura científica, de bases de dados de acesso aberto, do banco de dados do laboratório e dos bancos de dados dos nossos colaboradores. Também serão gerados dados *in silico* por simulação computacional.

### Documentação e Metadados

---

#### Que documentação e metadados acompanharão os dados?

Os dados serão contextualizados através de metadados sobre (i) a natureza, a frequência e o contexto dos eventos de interação registrados na natureza ou no laboratório; (ii) os organismos envolvidos nas interações, sua taxonomia, história natural e traços funcionais; (iii) os locais onde essas interações foram registradas e suas condições ambientais (naturais, urbanas ou rurais); e (iv) as fontes de onde as interações foram compiladas. Junto com dados e metadados, disponibilizaremos documentos explicativos para potenciais usuários, ao estilo dos metadados usados nos *data papers* da revista *Ecology*.

### Ética e Conformidade Legal

---

#### Como você administrará qualquer questão ética?

Como o projeto será baseado em dados secundários e dados gerados *in silico*, além de não envolver pesquisa com seres humanos, não há questões éticas envolvidas, como por exemplo necessidade de pedir licenças de pesquisa a órgãos ambientais ou médicos.

#### Como você vai gerenciar os direitos autorais e os direitos de propriedade intelectual (IP /

## IPR)?

Os dados compilados pelo projeto serão posteriormente disponibilizados para o público na forma de *data papers* ou de conjuntos de dados depositados em repositórios de acesso aberto, em sua maioria usando-se licenças [Creative Commons](#) ou licenças próprias de cada revista ou repositório. Os códigos de programação gerados pelo projeto serão depositados e disponibilizados no [GitHub](#) sob licenças Creative Commons.

## Armazenamento e Backup

---

### Como os dados serão armazenados e terão backup durante a pesquisa?

Os dados e metadados serão armazenados na forma de *data frames* (planilhas verticalizadas), seguindo a filosofia [tidy data](#), em arquivos eletrônicos em formato TXT ou CSV. O backup será feito em servidores na nuvem via Google Drive institucional e também de forma física em SSDs externos plugados aos computadores do laboratório, gerenciados via software de backup Time Machine.

### Como você vai gerenciar o acesso e a segurança?

O acesso aos dados, durante a execução do projeto, será exclusivo aos membros do projeto. Depois que o projeto for encerrado e as nossas principais descobertas forem publicadas, os dados serão disponibilizados ao público seguindo os padrões da filosofia [open science](#), como comentado anteriormente.

## Seleção e Preservação

---

### Quais dados são de valor a longo prazo e devem ser mantidos, compartilhados e / ou preservados?

Todos os dados sobre interações interespecíficas, assim como os metadados que os contextualizam, são de valor a longo prazo e, portanto, serão preservados tanto no laboratório quanto na forma de publicações de acesso aberto.

### Qual é o plano de preservação a longo prazo do conjunto de dados?

O plano de preservação a longo prazo está baseado nos backups feitos no laboratório (nuvem e físicos) e também na publicação dos dados na forma de *data papers*, assim como em seu depósito em repositórios estáveis e de acesso aberto.

## Compartilhamento de Dados

---

### Como você vai compartilhar os dados?

Na forma de *data papers* e conjuntos de dados depositados em repositórios estáveis e de acesso aberto.

### Existem restrições ao compartilhamento de dados requeridos?

Restrições serão aplicadas apenas até que sejam concluídas as principais publicações do projeto.

## Responsabilidades e Recursos

---

### Quem será responsável pelo gerenciamento de dados?

Os responsáveis pelo gerenciamento de dados serão o pesquisador responsável pelo projeto e também um bolsista de treinamento técnico.

### Quais recursos você precisará para entregar seu plano?

Bolsa de treinamento técnico para uma pessoa que será co-responsável pelo gerenciamento de dados, computador de mesa Apple de alta performance, SSDs externos com conexão USB-C para backups físicos locais, e acesso institucional com espaço ilimitado ao Google Drive.

---