



# Data Stewardship Wizard: Představení

Série školení pro akademické a výzkumné pracovníky UK v oblasti otevřeného přístupu k vědeckým informacím



CENTRUM PRO PODPORU  
OPEN SCIENCE  
Univerzita Karlova

**Marek Suchánek**

[marek.suchanek@ds-wizard.org](mailto:marek.suchanek@ds-wizard.org)

 [0000-0001-7525-9218](https://orcid.org/0000-0001-7525-9218)

# Motivace



Operations Checklist	
Parking Brake	Set
Fuel Flow	Cutoff
Battery Switch	On
Hydraulic Pump ON	On
Landing Gear	On
Flaps	Check
Spoiler	Up
Fuel Amount	Retracted
De-Ice	Check
Passenger Sign	Off
Check Weather	Off
	Flight Services
Transponder	Standby
Anti Collision Lighths	On
Engine Start Switches	Check
Throttle Switch	On
Master Switch	On

- Plánování správy dat (DMP) je důležité pro (vědecké) projekty
- Data Stewardship je komplikovaná problematika pokrývající různá témata
- DMP = co, jak a proč se bude dělat, aby v průběhu nevznikaly nejasnosti
- DMP ≠ dokument s dlouhým textem pro uspokojení grantové agentury

- Kterou šablonu (DMP Template) vyplnit?



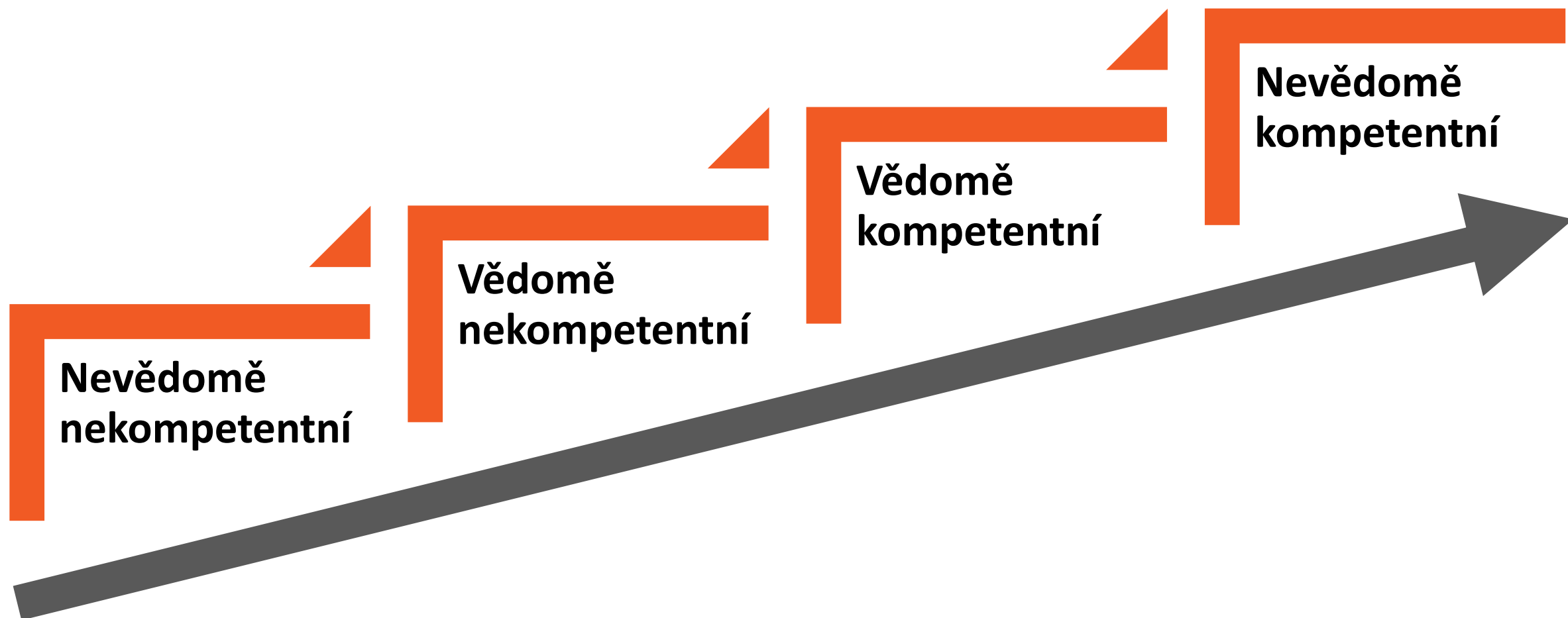
...a mnoho dalších

# Co je Data Stewardship Wizard?



- Open-Source nástroj vyvíjený v rámci ELIXIR
- Podporuje plánování správy dat s ohledem na současné best practice
- Není to nástroj pouze na „vyplnění DMP“, ale na sestavení plánu
- Funguje jako checklist před, během i po projektu

- Nepsat = nejde o slohové cvičení, co nejméně psaní textu
- Guidance = vede uživatele v rámci chytrých dotazníků
- Flexibilita = lze upravovat obsah a snadno integrovat s jinými službami
- Otevřenost = kdokoli je může použít a tvořit také vlastní obsah
- Orientace na uživatele = vývoj se řídí výrazně na základě zpětné vazby




# Vytváření projektu (dotazníku)

## Create Project

### Name

### Knowledge Model

 **Common DSW Knowledge Model** 2.3.0  
DSW Knowledge Model originated from mindmap made by Rob Hooft ✕

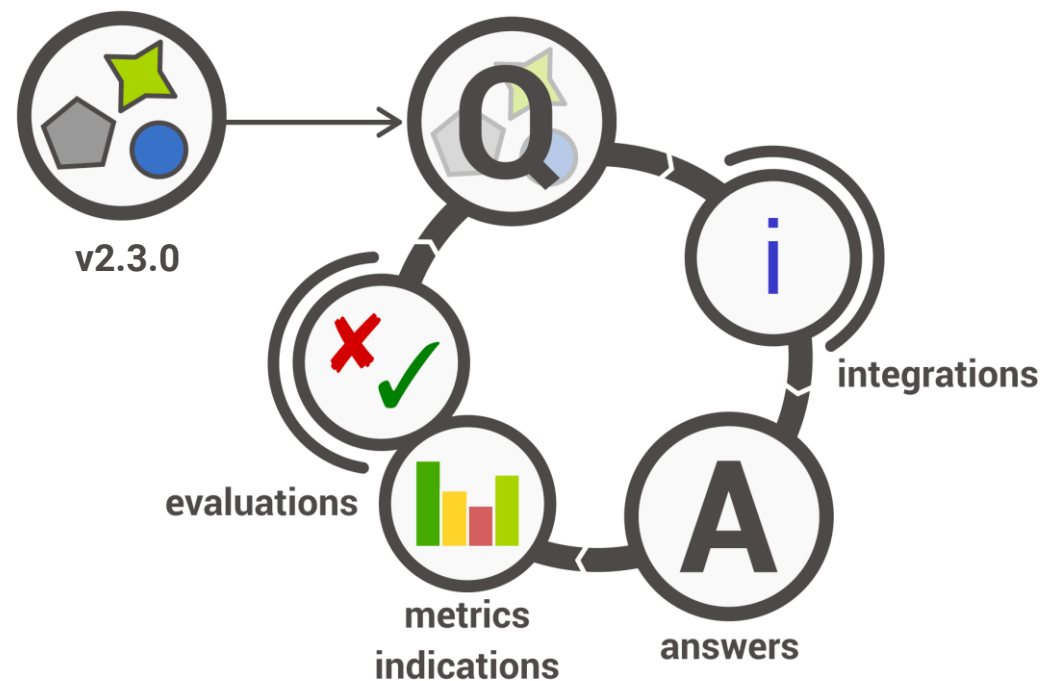
### Tags

- Horizon 2020 DMP
- Science Europe DMP
- maDMP

You can filter questions in the Questionnaire by tags. If no tags are selected, all questions will be used.

Cancel

Save





# DMP jako chytrý dotazník

## 1 Is there any pre-existing data? + !

Are there any data sets available in the world that are relevant to your planned research?

Desirable: *Before Submitting the Proposal*

Data Stewardship for Open Science: [atq](#)

External links: [Google dataset search](#), [Datacite Search](#)

a. No

b. Yes ☰

### 1.b.1 Will you be using any pre-existing data (including other people's data)? + !

Will you be referring to any earlier measured data, reference data, or data that should be mined from existing literature? Your own data as well as data from others?

Desirable: *Before Submitting the Proposal*

Data Stewardship for Open Science: [ezi](#)

a. No

b. Yes ☰

# Kolaborace nad plánováním



## Share Project

### Users

Add users

 Vojtěch Knaisl	Owner ▾	✕
 Jan Slifka	Editor ▾	✕
 Marek Suchánek	Viewer ▾	✕

**Visible by all other logged-in users**  
Other logged-in users can **view** ▾ the Project.

**Public link**

Cancel

Save

# Kolaborace nad plánováním



HVSC: Hypothetical Vascular Study by a Chemist



Share

Questionnaire **TODOs 2** Metrics Preview Documents Settings

## Current Phase

Before Submitting the Proposal

## Chapters

- I. Administrative details 1
- II. Re-using data 3**
- III. Creating and collecting data 6
- IV. Processing data 3
- V. Interpreting data 1
- VI. Preserving data 6
- VII. Giving access to data 3

## II. Re-using data

Before you decide to embark on any new study, it is nowadays good practice to check all options to re-use existing available data, either collected or generated by yourself in an earlier project, or data from others (Barend Mons calls this "Other PEople's Data And Services" or OPEODAS). This can include reusable data that have been created for an earlier study, and also so-called "reference data" which is used by many projects.

It is not because we can generate massive amounts of data that we always need to do so. Creating data with public money is bringing with it the responsibility to treat those data well and (if potentially useful) make them available for re-use by others. And the circle is only complete if such data is actually re-used.

### 1 Is there any pre-existing data?

Are there any data sets available in the world that are relevant to your planned research?

Desirable: *Before Submitting the Proposal*

Data Stewardship for Open Science: [atq](#)

a. No

b. Yes

Clear answer

#### 1.b.1 Will you be using any pre-existing data (including other people's data)?

**TODO x**

Will you be referring to any earlier measured data, reference data, or data that should be mined from existing literature? Your own data as well as data from others?

Desirable: *Before Submitting the Proposal*

Data Stewardship for Open Science: [ezi](#)

a. No

b. Yes

# Základem je kvalitní „guidance“

Vysvětlení

## 1.a.4.b.1.a.1 What repository will this data be stored in?




Domain repositories often have the best functionality to make the data findable and reusable: even though it may look like a database that could be reused in a completely different field would be better findable in a generic repository, the limited availability of domain-specific metadata make that less valuable.

Many repositories are listed in <https://fairsharing.org/>

If a repository offers to give your data set a DOI or alternative persistent identifier it is a good idea to use that option.

Desirable: *Before Finishing the Project*

External links: [FAIRSharing](#), [Registry of Research data Repositories](#)

a. A domain-specific repository 

Findability

b. Our national repository


Findability

c. Our institutional repository

Findability

d. A special-purpose repository for the project 

Findability

 Clear answer

Disadvantage of a general purpose repository is the lack of data-specific features (e.g. 'play' instead of 'download' for an audio file) and limited findability

Výběr možností

FAIR metriky



Reference



Doporučení

# Nabídka odpovědí pomocí integrace

## 2.a.7 Funding + !

Add all the funding that are part of this project.

Desirable: *Before Submitting the Proposal*

### 2.a.7.a.1 Funder + ! 🗑️

Czech

Ministerstvo Obrany České Republiky

Grantová Agentura České Republiky

České Vysoké Učení Technické v Praze

Parazitologický ústav, Akademie Věd České Republiky

Ministerstvo Zdravotnictví České Republiky


🕒 a. Planned

# Provázání odpovědi s URI

## 2.a.7.a.1 Funder



Univerzita Karlova v Praze

 <http://dx.doi.org/10.13039/100007397>

Specify the name of the funder that you ask for funding for your project. If the funder is not present in the suggested list, please specify a complete URL to the funder web site.

 **FAIRsharing.org**  
standards, databases, policies

 **Crossref**

**ROR** 

  
**WIKIDATA**







## Summary Report

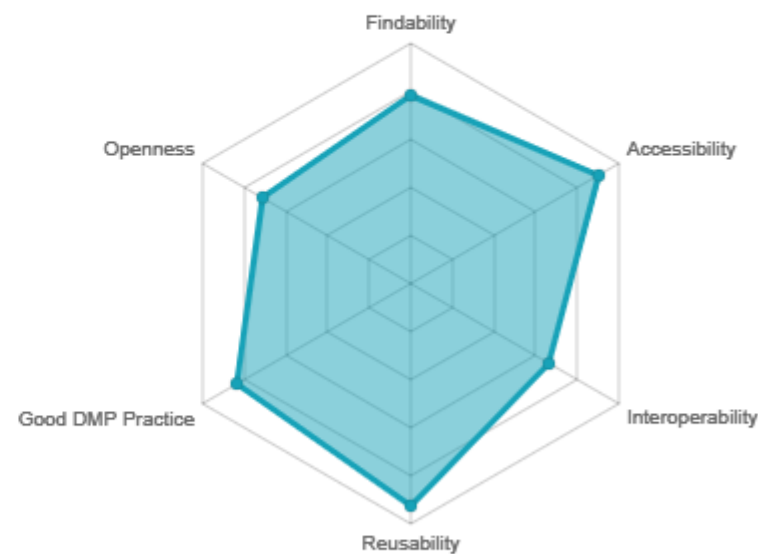
Answered (current phase): 100/100



Answered: 293/305



Metric	Measure	
Findability	0.78	
Accessibility	0.90	
Interoperability	0.67	
Reusability	0.92	
Good DMP Practice	0.84	
Openness	0.72	



# Exportování DMP pomocí šablony



## New document

### Name

Answered (current phase): 100/100



Answered: 293/305



### Template



**Horizon 2020 DMP 0.1.0**

Data Management Plan according to the H2020 template

### Format

HTML

PDF Document

LaTeX Document

MS Word Document

OpenDocument Text

Markdown Document

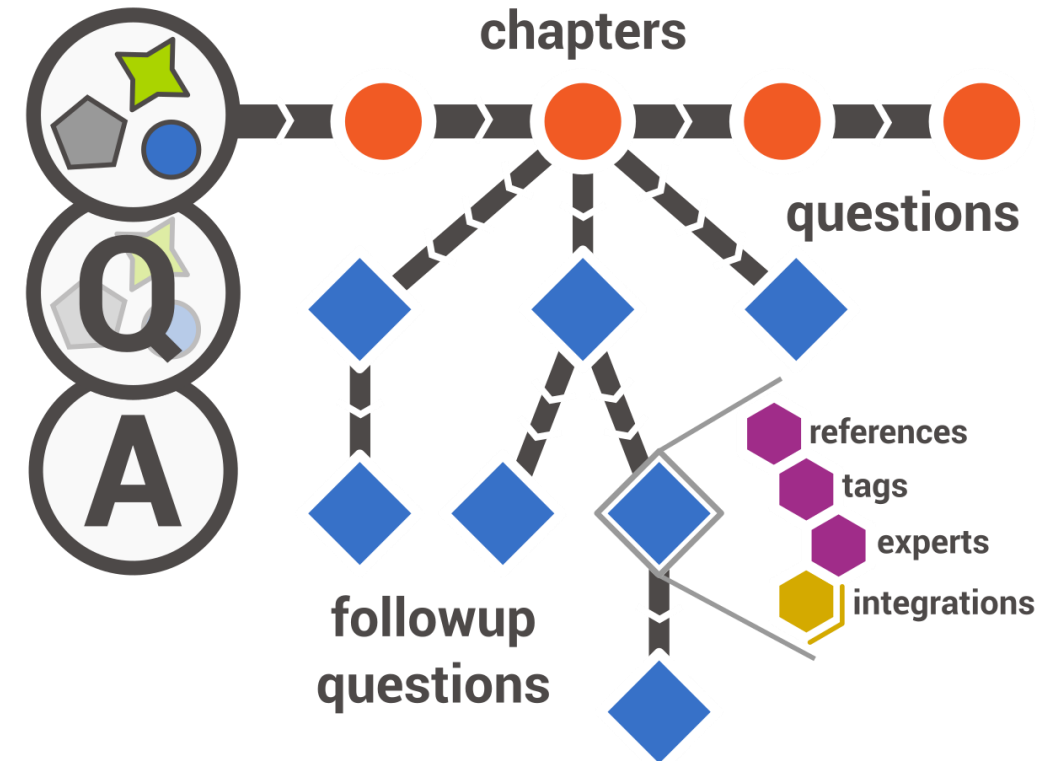
Cancel

Create



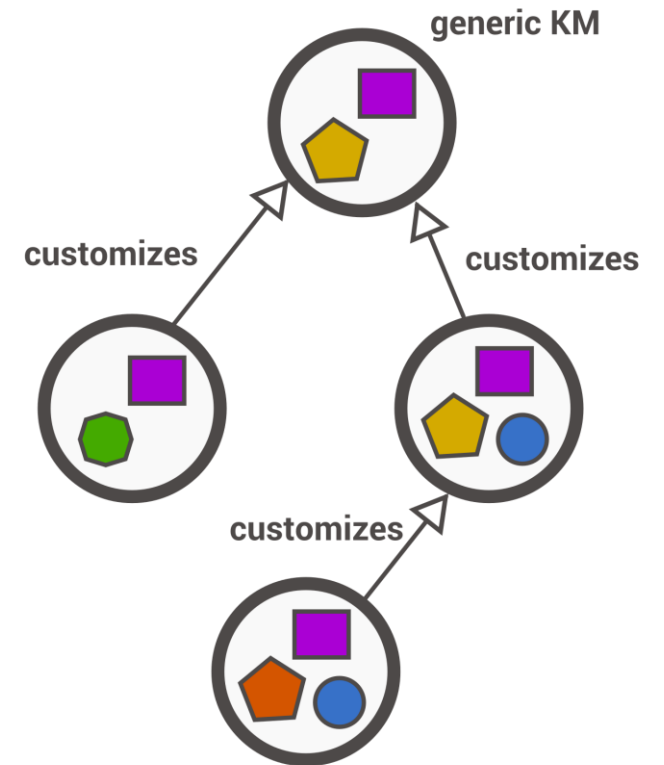
# Co je to „Knowledge Model“

- Struktura dotazníku v DSW není pevně daná
- Knowledge Model obsahuje:
  - Kapitoly
  - Otázky (různé typy)
  - Odpovědi (možnost výběru)
  - Reference a kontakty na experty
  - Tagy (štítky pro označení otázek)
  - Integrace pro napovídání odpovědí

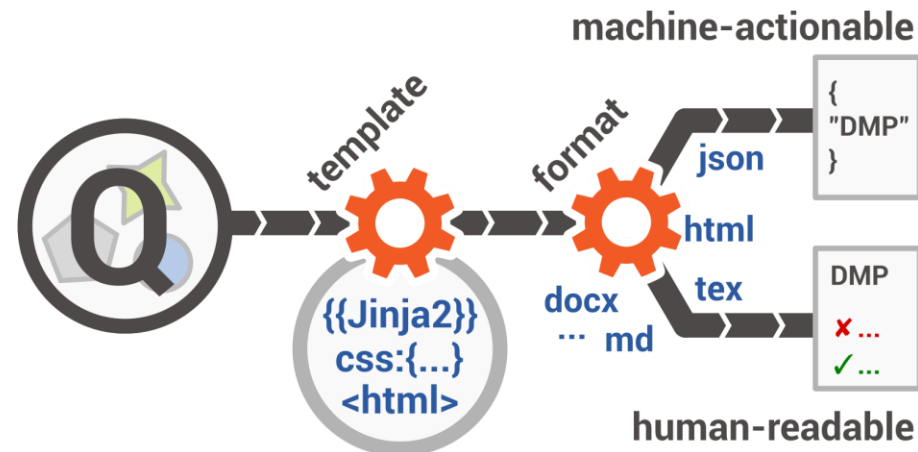


# Úprava na míru instituci

- V DSW je přímo zabudován editor Knowledge Modelů
- Lze si vytvořit zcela nový „from scratch“
- Lze ale také lokalizovat již existující:
  - Odstranit otázky, které nejsou relevantní pro danou instituci
  - Přidat nové otázky, které naopak jsou důležité pro instituci
  - Změnit existující otázky
  - Případně samozřejmě také měnit strukturu



- Šablony rovněž nejsou přímo v kódu a lze je tak měnit či vytvářet vlastní
- Jsou vytvářené v běžném šablonovacím jazyce Jinja2
- K dispozici je „DSW Template Development Kit“ pro snazší vývoj

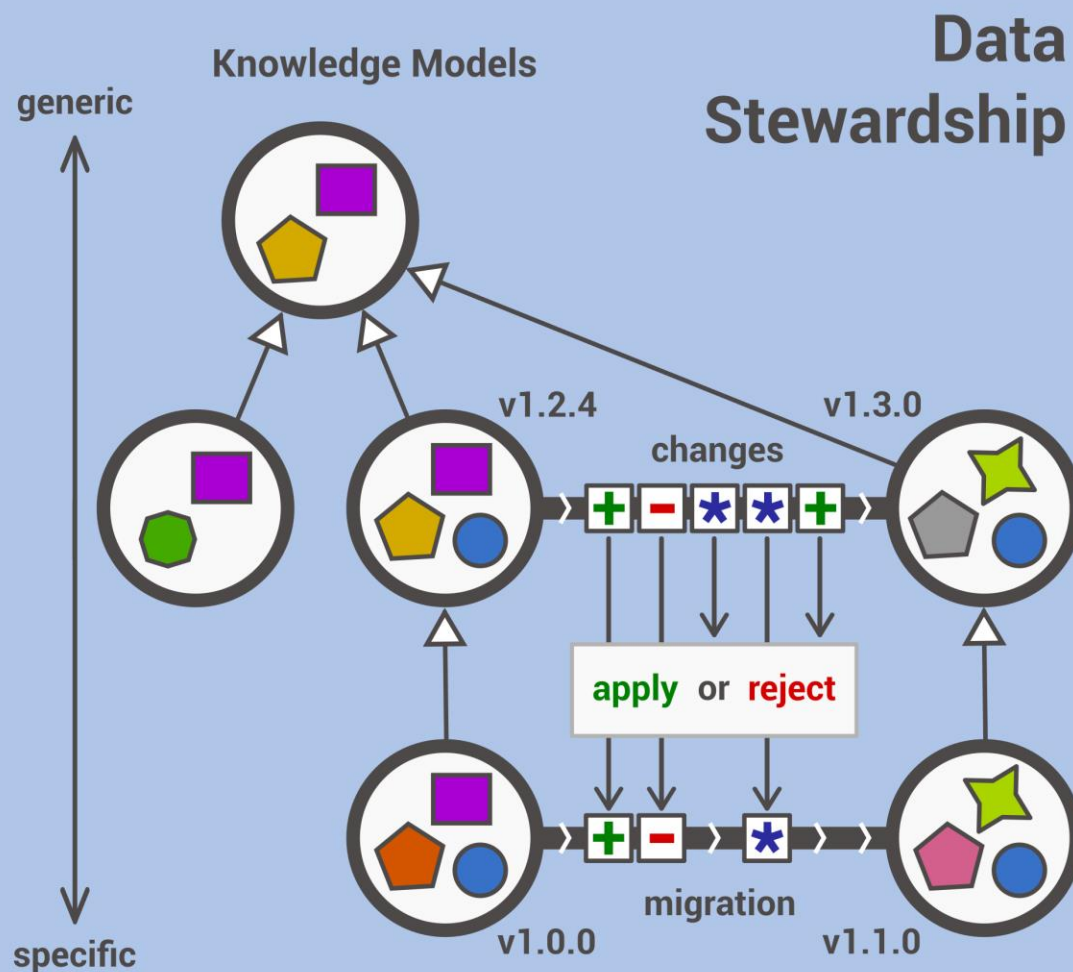


# A co když se něco aktualizuje?



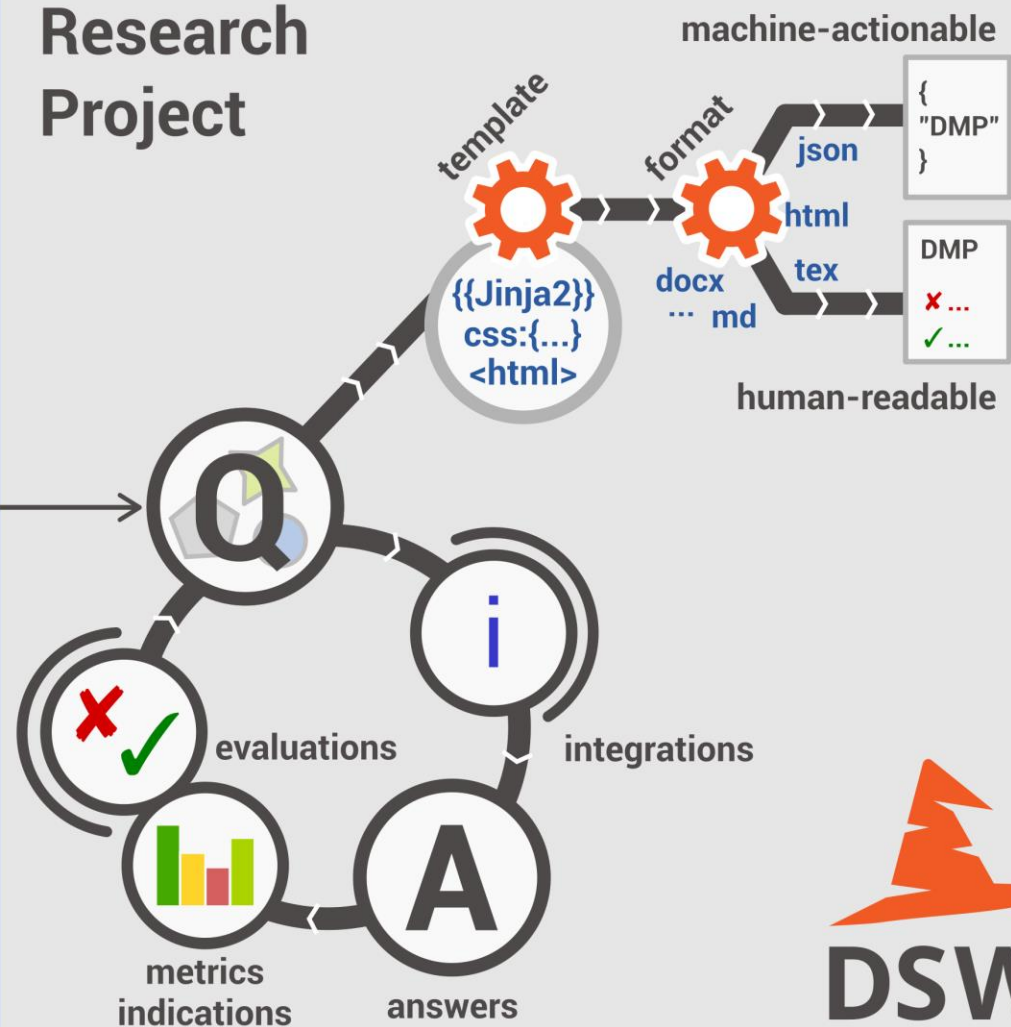
- DSW umí komplexní migrace na více úrovních
- Pokud je aktualizován rodičovský KM, lze zmigrovat změny do lokalizace
- Pokud je aktualizován KM, lze zmigrovat dotazníky na novou verzi
- Při změně vnitřních struktur v DSW jsou implementovány příslušné migrace pro databázi i import (např. pro KM z exportovaného souboru)

# Jak to všechno funguje



## Data Stewardship

## Research Project



- Zjistili jste mnoho o plánování správy dat, FAIR datech, Open Science, ...
- Tím jste se dostali minimálně do fáze „vědomě nekompetentní“
- Další krok je získat kompetenci v práci s daty
- Data Stewardship Wizard se bude snažit Vám tyto kroky usnadnit a vést Vás
- Stačí si zkusit vyplnit pro sebe dotazník pro svůj (imaginární) projekt
  - Navštivte [ds-wizard.org](https://ds-wizard.org), začněte plánovat, použijte „**Common DSW Knowledge Model**“
  - Dejte nám vědět, jak to šlo a co můžeme zlepšit

# Poděkování



Vývoj a provoz DSW byl podpořen z projektu výzkumné infrastruktury ELIXIR CZ (MŠMT grant č.: LM2018131).

---

Tato prezentace byla podpořena projektem  
[OP VVV Rozvoj kapacit pro výzkum a vývoj UK II,](#)  
CZ.02.2.69/0.0/0.0/18\_054/0015222



UNIVERZITA  
KARLOVA



EVROPSKÁ UNIE  
Evropské strukturální a investiční fondy  
Operační program Výzkum, vývoj a vzdělávání





# Otázky a diskuze

*[info@ds-wizard.org](mailto:info@ds-wizard.org)*