

A reproducible open platform for a collective process

Translating CIDOC CRM 7.1.2 in French

GUILLEM Anais, KRUMMEICH Raphaëlle, DAVID Bertrand, VAN RUYMBEKE Muriel



Useful links

Gitlab Link of the project:

<https://gitlab.huma-num.fr/bdavid/doc-fr-cidoc-crm>

Gitlab README of the translation project:

<https://gitlab.huma-num.fr/bdavid/doc-fr-cidoc-crm/-/blob/translate/README.md>

Gitlab wiki of the translation project:

<https://gitlab.huma-num.fr/bdavid/doc-fr-cidoc-crm/-/wikis/home>

Translation: you can find here the translation as it is continuously integrated

<https://cidoc-crm-fr.mom.fr/>

Zotero bibliography (shared Zotero bibliography on CIDOC CRM, ontologies, LOD etc)::

https://www.zotero.org/groups/4641239/cidoc_crm_bibliographie

People and Institutions

- https://gitlab.huma-num.fr/bdavid/doc-fr-cidoc-crm/-/project_members?page=1

+/- 15 researchers from different cultural and research institutions (CNRS, MOM, universities etc)

- Institutional support: MASA Consortium
- Technical support:
 - Huma-Num (Gitlab hosting);
 - Maison de l'Orient et de la Méditerranée (MOM) (translation site web)



Objectives: reproducible open platform and a collective process

- Full translation of the CIDOC CRM 7.1.2/3 (future ISO version), including introductory texts, figures, tables etc
- Project of translation as:
 - **F**indable
 - **A**ccessible
 - **I**nteroperable and **O**pen
 - fully **R**eproducible and **R**eusable
- Building a CIDOC CRM user community among Digital Humanity research institutions
 - Translation as learning experience
 - Translation as collaboration project
 - Translation as a Digital Humanities research project

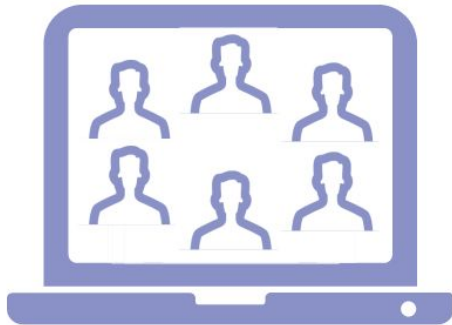
automated update of
a static webpage with
the actual translation

Method and workflow

collective synchronous work

- CIDOC translation issues
- Group's methods & other issues

RENDEZ-VOUS
par RENATER



Friday 3-5PM
collective validation sessions



Huma-Num

git push

CI process

E28_Objet_conceptuel
E29_Projet_ou_Procédure
E30_Droit
E31_Document
E32_Document_d'autorité
E33_Objet_linguistique
E34_Inscription
E35_Titre
E36_Élément_visuel
E37_Marque
E39_Acteur-riche-x
E41_Appellation
E42
E52_Laps_de_temps
E53_Lieu
E54
E55_Type
E56_Langue
E57_Matériau
E58_Unité_de_mesure
E59
E60
E61
E62
E63_Début_d'existence
E64_Fin_d'existence
E65_Création

individual asynchronous work

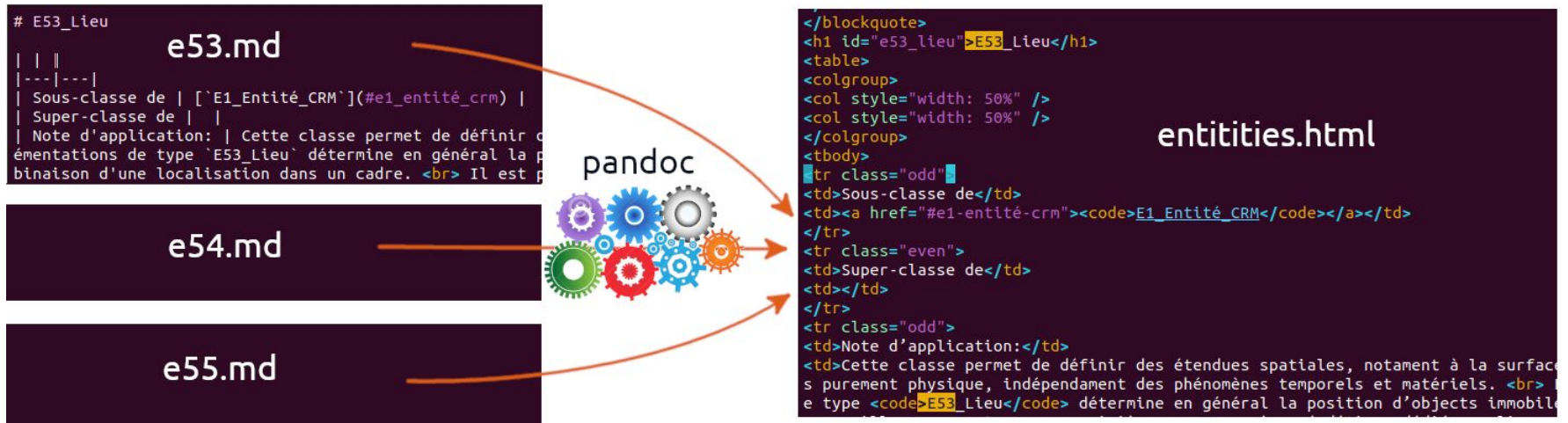
- translations (non git or git users)
- technical issues (git users)

Continuous Integration Process

.gitlab-ci.yml :

2 Stages:

- build -> run a **pandoc** docker container to generate html from markdown files (Entities, Properties), generate and concat index and menus
- deploy -> adjust CSS settings, copy files on web-server



Gitlab- Huma-Num/Doc fr CIDOC CRM

files/texts to translate

project management

project info

technical elements

The screenshot displays the GitLab interface for the 'Doc fr CIDOC CRM' project. The sidebar on the left contains navigation options: Project information, Repository, Issues (201), Merge requests (0), CI/CD, Security & Compliance, Deployments, Packages and registries, Infrastructure, Monitor, Analytics, Wiki, Snippets, and Settings. The main content area shows the project name 'Doc fr CIDOC CRM' with a globe icon, Project ID: 380, and a 'Leave project' link. It also displays statistics: 761 Commits, 2 Branches, 0 Tags, and 164.1 MB Project Storage. The project description is 'Traduire en français la documentation du CIDOC CRM'. Below this, there is a search bar with 'translate' selected and a dropdown for 'doc-fr-cidoc-crm'. There are buttons for 'Find file', 'Web IDE', and 'Clone'. A recent commit by 'màj graph groupes logiques' (Raphaelle Krummeich) is shown. Below the commit, there are buttons for 'README', 'CI/CD configuration', and several 'Add' buttons for LICENSE, CHANGELOG, CONTRIBUTING, and Kubernetes cluster. At the bottom, a table lists files and their commit history.

Name	Last commit	Last update
assets	amélioration mise en forme (CSS)	4 months ago
entities	Merge branch 'translate' of https://gitlab.hu...	1 month ago
figures	màj typo figures	4 months ago
html	suite modif pour l'intégration continue (CI)	2 months ago
logos	création footer.html	4 months ago
properties	passé avec toc-title	3 weeks ago
stats	màj graph groupes logiques	1 day ago

Gitlab issues and boards

translation workflow

validation workflow

Open general issues

To be translated

Translation ongoing

Translated, waiting for validation

Under review/ validation

Once validated, issue closed

The screenshot shows the GitLab interface with a sidebar on the left containing navigation options like 'Project information', 'Repository', 'Issues', 'Boards', 'Service Desk', 'Milestones', 'Merge requests', 'CI/CD', 'Security & Compliance', 'Deployments', 'Packages and registries', 'Infrastructure', 'Monitor', 'Analytics', 'Wiki', 'Snippets', and 'Settings'. The main content area displays a list of 'Open' issues, including titles like 'Intégration continue pour reconstruire le document "Mati"', 'Positionnement de cette traduction par rapport à la précédente ISO 21127:2014', and 'OmbME.dataforhistory.org'.

The screenshot displays a Kanban board with five columns representing the workflow stages: 'To be translated', 'Translation ongoing', 'Translated, waiting for validation', 'Under review/ validation', and 'Once validated, issue closed'. Each column contains a list of issues with their titles and IDs, such as 'P93 took out of existence (was taken out of existence by)', 'P136 was based on', 'P98 brought into life (was born)', and 'E24 Physical Human made Thing'. A callout box at the bottom center states: '1 issue = 1 element to translate (ie. text of a class or a property)'. The board is color-coded with blue and purple borders.

1 issue = 1 element to translate (ie. text of a class or a property)

Example issue #55: translation of E25

Bertrand DAVID-JACQUOT > Doc fr CIDOC CRM > Issues > #55

Open Issue created 1 year ago by Guillem Anais (Maintainer)

Close issue

E25 Human Made Feature

Document à traduire/traduit E25 ← link to the file in repo

Edited 1 year ago by Raphaëlle Krummeich

0 thumbs up 0 thumbs down

Create merge request

Drag your designs here or click to upload.

Tasks 0

Add

No tasks are currently assigned. Use tasks to break down this issue into smaller parts.

Linked items 0

Add

Link issues together to show that they're related. Learn more.

Activity

Sort or filter

Guillem Anais added 3-Traduit en attente de révision label 1 year ago ← current status

Guillem Anais changed title from E25- Human Made Feature to E25- Human Made Feature (AG, Février 2021) 1 year ago

Bulle Tuil Leonetti @btuilleonetti · 1 year ago
Traduction de feature dans E25 Antériorités/sources existantes : #issue 17 - index des termes les plus utilisés #issue 58 - feature #issue 74 - initiative canadienne #issue 76 - cohérence entre 25 & 26
La proposition de traduction de E25 Human-Made Feature est E25 Trait caractéristique fabriqué
Plusieurs propositions (facture remarquable, trait distinctif, caractéristique) sont discutée dans #issue76, ainsi qu'une autre (fait - contribution des archéologues) dans #issue 58
Considérant que : fait s'inscrit dans une logique argumentaire de la trace archéologique #issue58 caractéristique non féminin correspond à un E55 type, en accord avec l'argument typologique de #issue58 fonctionnalité (hors sujet ici)
la proposition est d'ajouter le substantif trait à caractéristique fabriquée.

Bertrand DAVID-JACQUOT mentioned in issue #76 (closed) 1 year ago

Bertrand DAVID-JACQUOT mentioned in issue #58 (closed) 1 year ago

Raphaëlle Krummeich added 4-En cours de révision label and removed 3-Traduit en attente de révision label 1 year ago

Raphaëlle Krummeich changed title from E25 Human Made Feature (AG, Février 2021) to E25 Human Made Feature 1 year ago

labels

Add a to do

Assignee Guillem Anais

Labels

4-En cours de révision Entité

v7.1.1

Milestone None

Due date None

Time tracking No estimate or time spent

Confidentiality Not confidential

Lock issue Unlocked

Notifications

4 participants

Reference: bdauid/doc-fr-cidoc-cr...

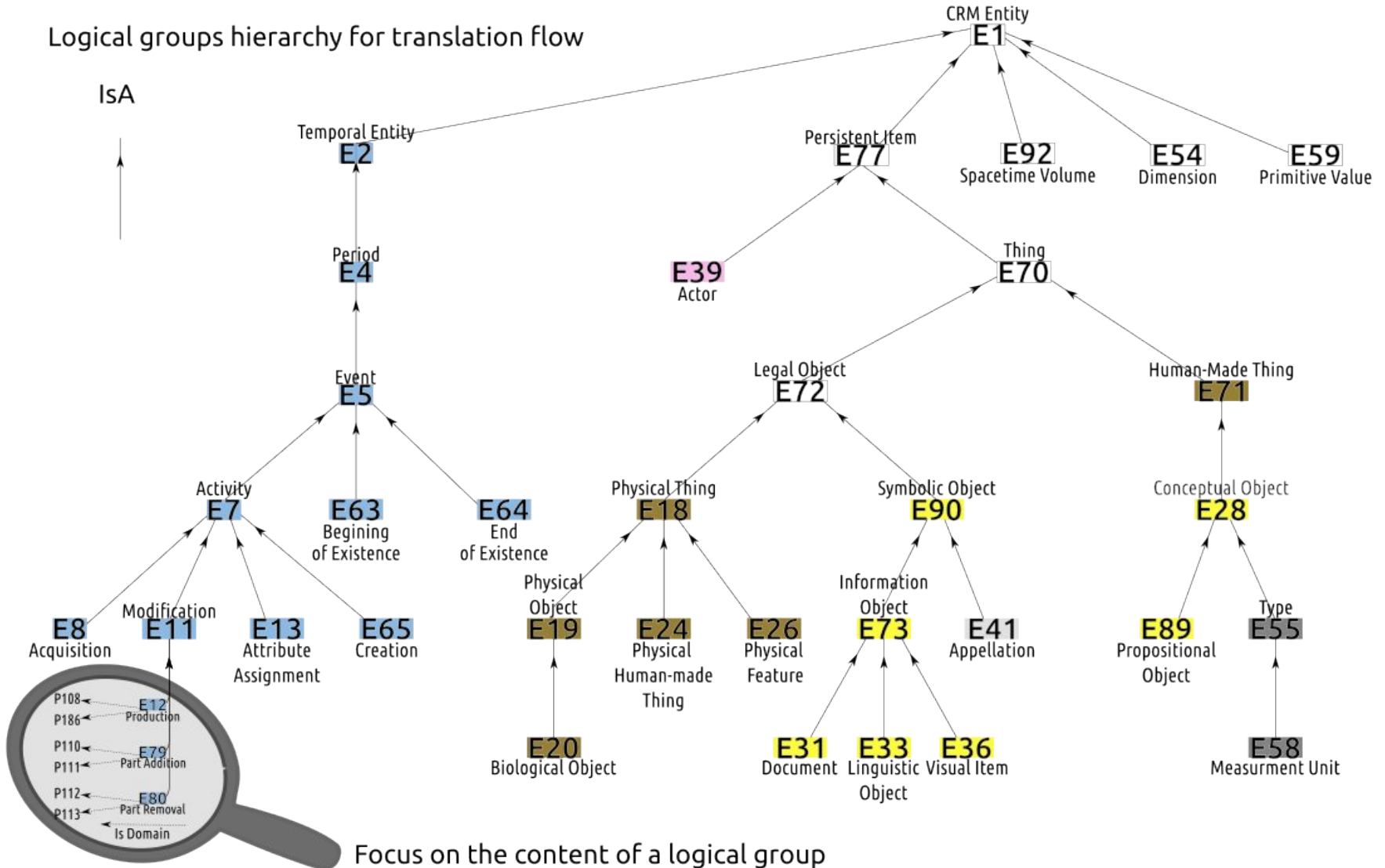
Move issue

translation notes added as comments

Milestones' construction : logical groups

Sources : class hierarchy & property-class hierarchy tables of Cidoc CRM reference Manual

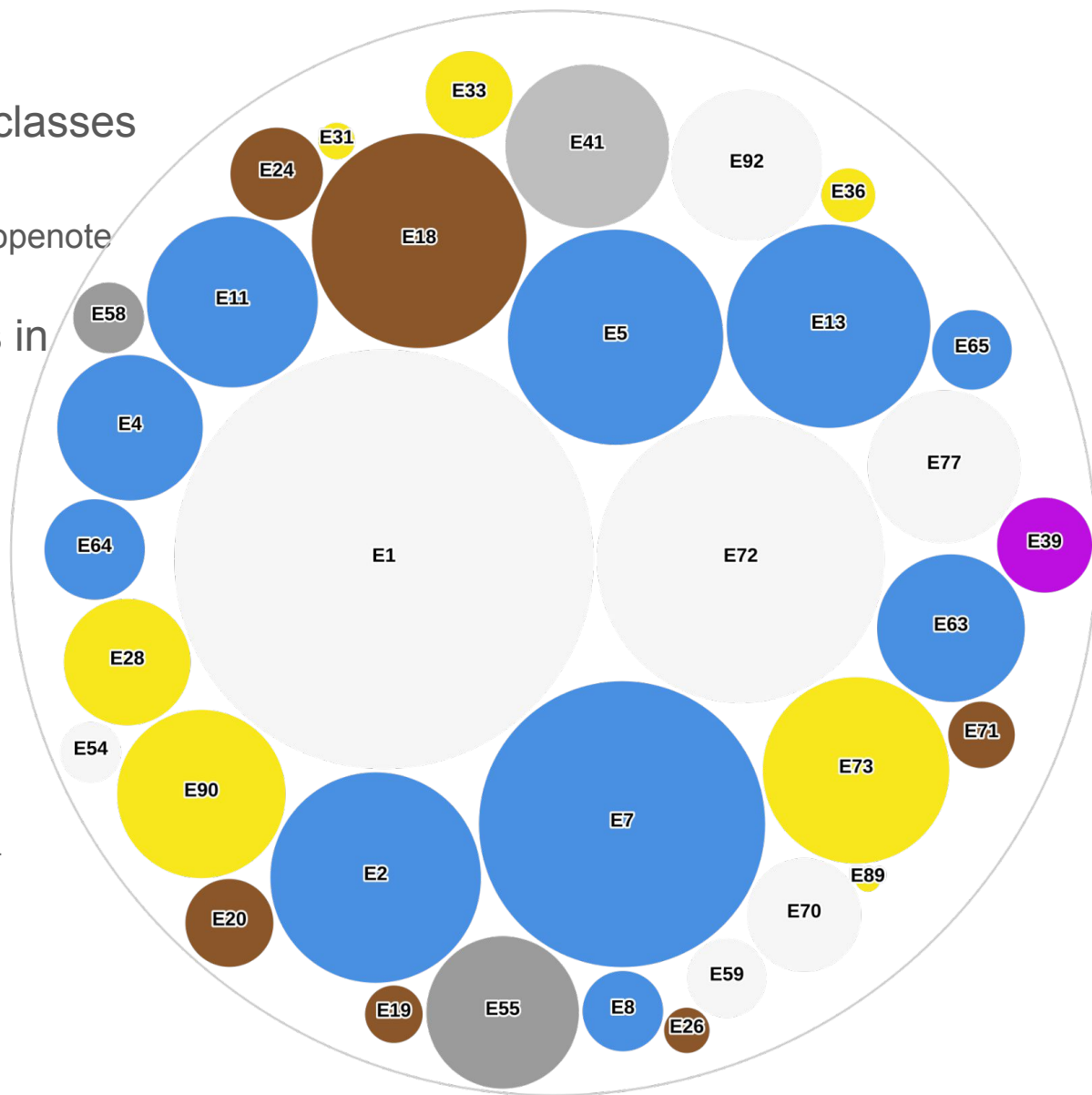
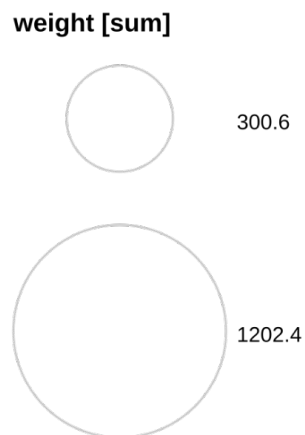
Logical groups hierarchy for translation flow



Logical groups as “time-weighted” milestones

Source : XML (interchange format) file of Cidoc CRM reference Manual

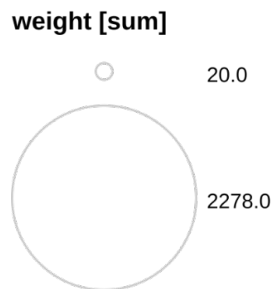
- use of XML file to “weigh” classes and properties
 - XML automated parse of scopenote
 - Count words in scopenote
- group “weighted” concepts in logical groups



Translation flow labels inside logical groups

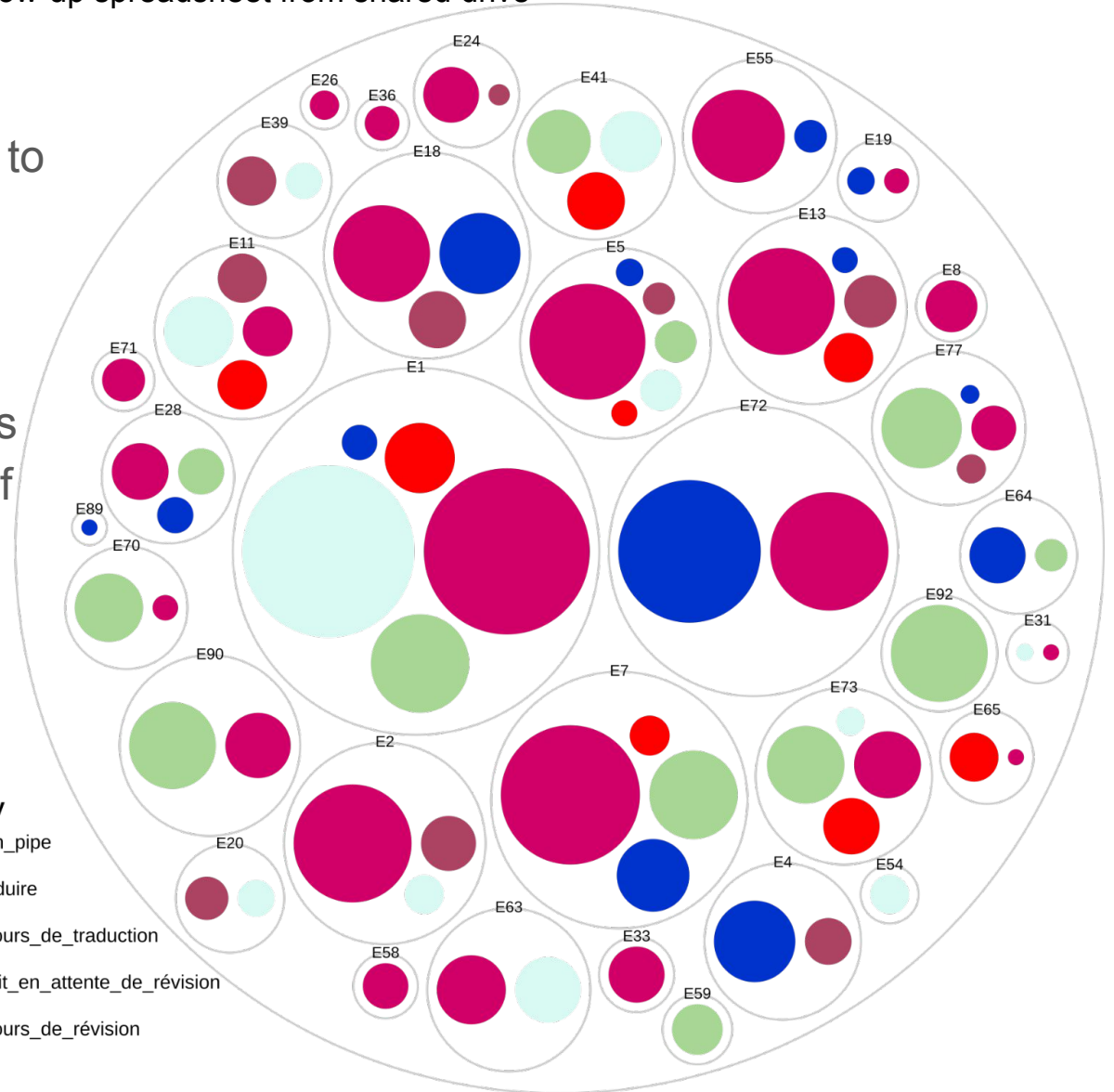
Source : GitLab Wiki link, ©Google follow-up spreadsheet from shared drive

- evaluate subgroups flows to set validation agenda
 - priority for asynchronous translation
 - creation of new issues
- quantify & qualify sessions needed to set validation of logical groups
 - provisional agenda
 - thematic sessions



modality

- Not_in_pipe
- A_traduire
- En_cours_de_traduction
- Traduit_en_attente_de_révision
- En_cours_de_révision
- Validé



Focus on logical groups and translation workflow

The screenshot displays the GitLab interface for the 'Doc fr CIDOC CRM' project. The left sidebar shows navigation options, with 'Milestones' selected. The main content area shows a list of milestones, each with a progress bar and completion percentage. A purple box highlights the 'Groupe E72 Legal Object' milestone, which is 70% complete. An arrow points from the text 'focus on validation workflow' to this milestone.

Milestone Name	Issues	Merge requests	Completion	Action
Groupe E65 Creation	3	0	0% complete	Close Milestone
Groupe E73 Information Object	10	0	0% complete	Close Milestone
Groupe E7 Activity	33	0	9% complete	Close Milestone
Groupe E70 Thing	6	0	0% complete	Close Milestone
Groupe E18 Physical Thing started on Sep 9, 2022	9	0	44% complete	Close Milestone
Groupe E71 Human-Made Thing	1	0	0% complete	Close Milestone
Groupe E72 Legal Object started on Sep 23, 2022	17	0	70% complete	Close Milestone
Groupe E77 Persistent Item	8	0	12% complete	Close Milestone
Groupe E8 Acquisition	2	0	0% complete	Close Milestone
Groupe E90 Symbolic Object	4	0	0% complete	Close Milestone
Groupe E92 Spacetime Volume	6	0	0% complete	Close Milestone

Focus on logical group E72 and validation workflow

The screenshot displays the GitLab interface for a milestone titled "Groupe E72 Legal Object". The page shows a progress bar at 70% completion, a start date of Sep 23, 2022, and a due date of "No due date (71 days elapsed)". The milestone is categorized as "Open" and started on Sep 23, 2022. The main content area is titled "Validation du groupe logique E72 Legal Object" and is divided into three columns: "Unstarted Issues (open and unassigned)", "Ongoing Issues (open and assigned)", and "Completed Issues (closed)".

Unstarted Issues (open and unassigned): 0 items.

Ongoing Issues (open and assigned): 5 items. These include issues like #237, #236, #148, #147, and #146, all with labels such as "Traduit en attente de revision", "Propriété", "Entité", and "v7.1.1".

Completed Issues (closed): 12 items. These include issues like #238, #149, #145, #144, #142, #143, #140, #139, #138, #137, #136, and #92, all with labels such as "S-Valide", "7.1.2", "v7.1.1", and "Propriété".

Annotations on the screenshot include:

- A callout box pointing to the "Unstarted Issues" header: "link to the calculated logical group graph md file in repo".
- A callout box pointing to the "Ongoing Issues" header: "Open issues: ongoing translation, or waiting for review".
- A callout box pointing to the "Completed Issues" header: "Closed issues: validated translations".
- A callout box pointing to the "70% complete" progress bar: "Percentage of completion".

Conclusion: a mean to build a community of CRM users

- CIDOC CRM ontology aims at facilitating Linked Open Data workflows. The translation in French is a **mean** to facilitate its intake in French speaking communities, not an **end**
- The translation project is an **opportunity** for:
 - building a community of CIDOC CRM users, building on existing expertise and projects
 - open discussions on data integration among cultural heritage experts (from different institutions) through translation questions
 - share an understanding on CIDOC CRM ontology and (best) practices
- The translation project as **experimental space** in DH:
 - experimenting on open source tools, methods, workflows
 - (re)inventing how to do a translation: collective, peer-review, hybrid, open