

# 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](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](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

# Method and workflow

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

automated update of  
a static webpage with  
the actual translation

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

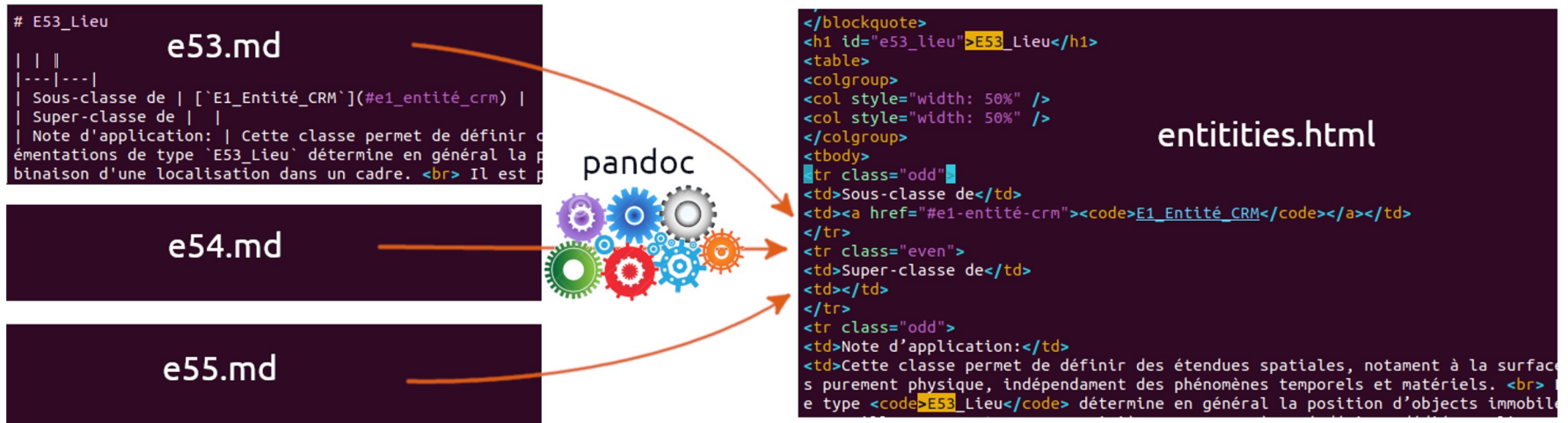
- translations (non git or git users)
- technical issues (git users)

# 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 shows the GitLab interface for the project 'Doc fr CIDOC CRM'. The left sidebar 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 displays project statistics (761 Commits, 2 Branches, 0 Tags, 164.1 MB Project Storage) and a commit by 'màj graph groupes logiques' (74a99a0d). Below this is a table of files and folders with their last commit details.

Name	Last commit	Last update
assets	amélioration mise en forme (CSS)	4 months ago
entities	Merge branch 'translate' of <a href="https://gitlab.hu...">https://gitlab.hu...</a>	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 displays the GitLab interface with a Kanban board. The board is divided into several columns representing different stages of the workflow. The first column, 'Open', contains a list of general issues. The subsequent columns are part of a Kanban board with the following labels: '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 respective titles and IDs. The board is color-coded: blue for 'To be translated', purple for 'Translation ongoing', and purple for 'Translated, waiting for validation'. The 'Under review/validation' and 'Once validated, issue closed' columns are also purple. A callout box at the bottom center states: '1 issue = 1 element to translate (ie. text of a class or a property)'. The left sidebar shows the GitLab navigation menu with 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'.

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.

translation notes added as comments

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

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

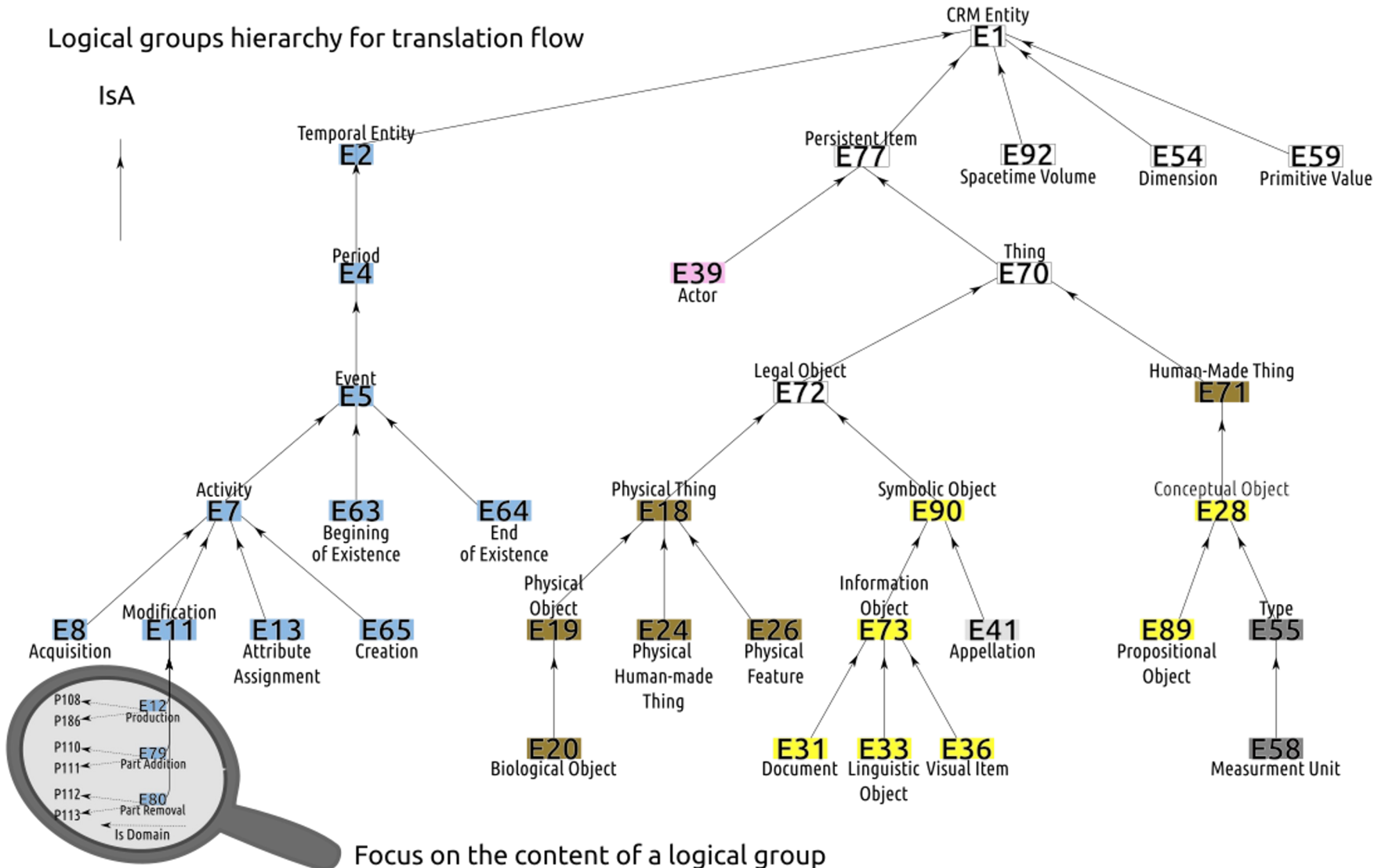
Reference: bldavid/doc-fr-cidoc-cr...

Move issue

# Milestones' construction : logical groups

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

Logical groups hierarchy for translation flow

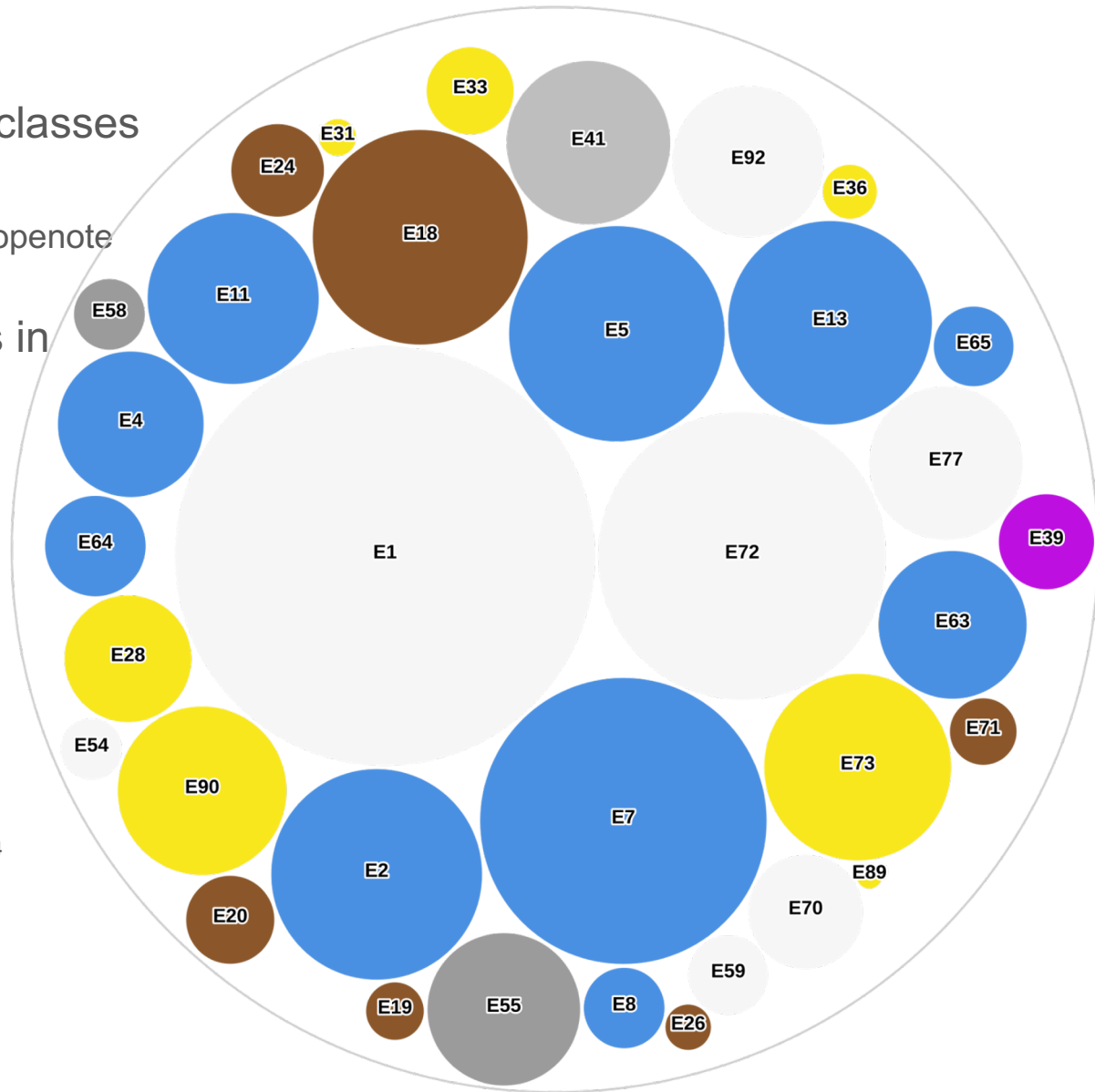
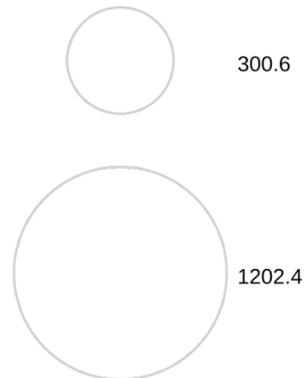


# Logical groups as “time-weighted” milestones

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

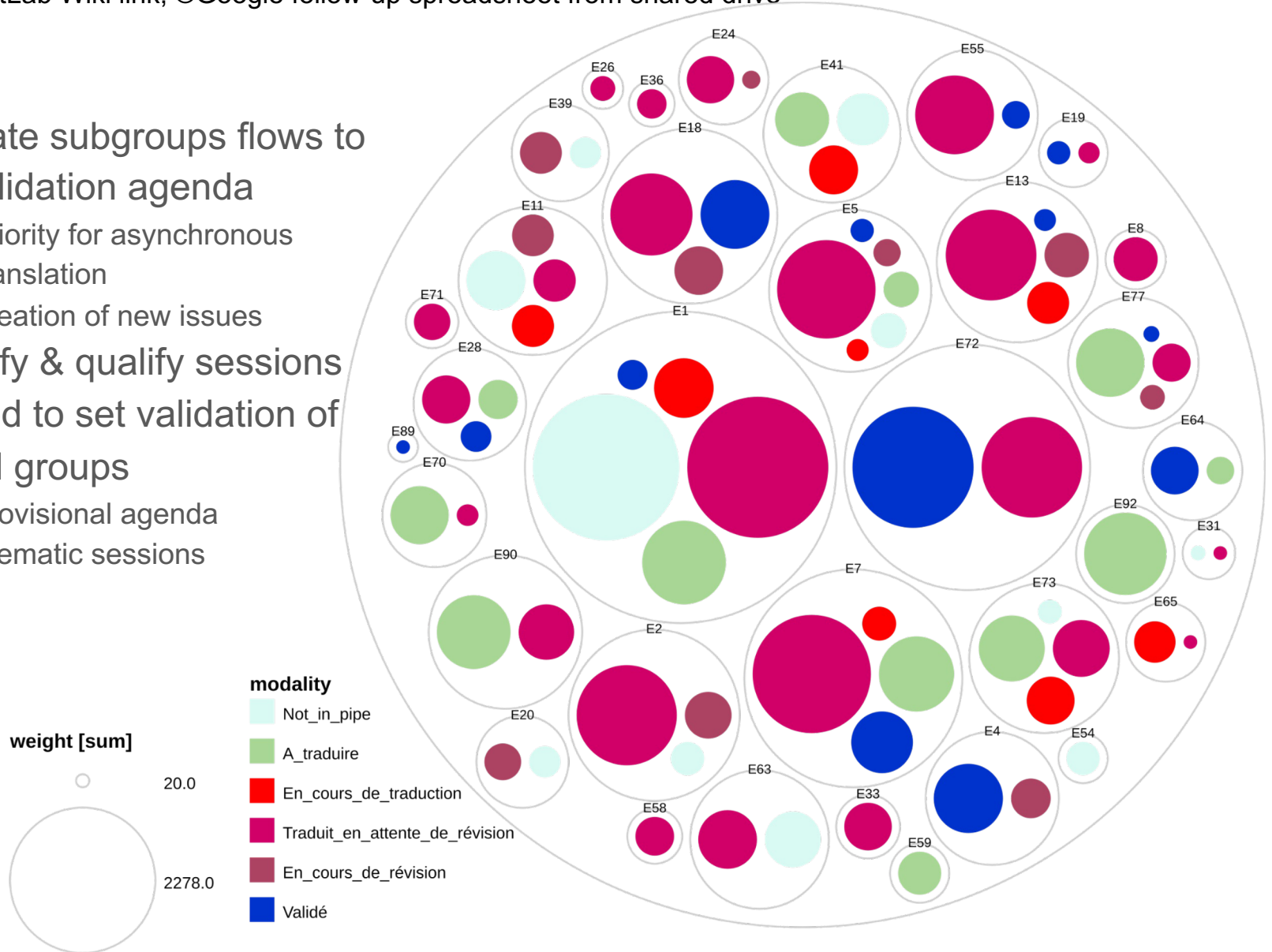
weight [sum]



# 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



# 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	9	0	44% complete	Close Milestone
Groupe E71 Human-Made Thing	1	0	0% complete	Close Milestone
Groupe E72 Legal Object	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 shows the GitLab interface for a milestone titled "Groupe E72 Legal Object". At the top, a progress bar indicates "70% complete". Below this, the milestone is marked as "Open" and started on "Sep 23, 2022". The main content area 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 P190, E90, P196, P156, and P128, all with a status of "3-Traduit en attente de revision".

**Completed Issues (closed):** 12 items. These include issues like P106, P198, P59, P53, P144, P52, P51, P50, P49, P46, P45, P44, and E18, all with a status of "5-Valide".

Annotations on the image include:

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

## 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