

Annotation Codebook

A Supervised Machine Learning Approach for Assessing Grant Peer Review Reports

The codebook contains a description and coding instruction for each category that was annotated, including those categories that were later omitted from further analyses due to annotation difficulty, low inter-coder agreement or later poor classification performance.

For some categories more examples and longer descriptions are provided than for others. This reflects the discussions and recalibration of the four annotators concerning certain categories.

Some categories in the codebook have been renamed in the manuscript for clarity and consistency, but they measure the same aspects of the data as before. These changes do not alter the annotation process or the interpretation of the data. We have applied the updated category names while maintaining the original purpose and categories behind the classification.

General instructions

- Please code each target sentence according to the categories described in the following chapters.
- We use a binary classification, where 1 indicates “yes”, meaning that the category applies to the sentence, and 0 (or an empty column) indicates “no”, meaning that the category does not apply to the sentence.
- Only select 1 (“yes”) if a sentence unambiguously covers the respective category. We will train machine learning classifiers and must identify sentences that clearly address the category.
- When in doubt, we code 0/empty column (“no”). If you are not sure, do not understand the sentence and would not be able to clearly explain why you coded it as “yes”/1 **leave it blank**.
- For each target sentence the preceding and the following sentences are provided. The surrounding sentences can be used to better understand the context and topic of the target sentence. The surrounding sentences **should not** be used to assign categories to the target sentence; the exception is the category “**rationale context**”.
- One sentence can be coded for more than one category.

Overview of the categories

- 1. Criteria: Which criteria, defined by the SNSF, does the review apply?**
 - Track record and expertise
 - Relevance, originality and topicality
 - Suitability
 - Feasibility
- 2. Focus: Which parts of the SNSF application documents does the review refer to?¹**
 - Candidate_Other
 - Candidate_Quantity
 - Proposal
 - Proposal_Method
- 3. Statement type: Is the review making a positive or negative statement? Does it provide suggestions?**
 - Positive statement
 - Negative statement
 - Suggestion
- 4. Reasoning: Is a rationale for the positive and negative statements provided?**
 - Rationale
 - Rationale_Context
- 5. Does the review mention an impact beyond academia?**
 - Impact beyond academia

¹ The three categories "Candidate_Other", "Candidate_Quantity" and "Proposal_Methods" were renamed into "Applicant", "Applicant: Quantity" and "Methods" respectively for publication in line with the SNSF use of terms.

Detailed coding instructions for each category

1. Which criteria, defined by the SNSF, does the review apply?

Definition/ question prompting coding:

Has the reviewer applied the criterion “track record/ expertise”, “relevance”, “suitability” or “feasibility” in the sentence?

Relevance and informative value of the category:

Get understanding of if and how often the different criteria are applied and whether they are applied in the “right” text boxes of the reviewer form.

Get understanding if a certain criterion is often mentioned together with a strength/ weakness/ neutral and whether it relates to the applicant or proposal.

Basis: [SNSF regulations on Project Funding, Art 15.](#)

1.1 Criterion_track_record

Definition/ question prompting coding:

Code 1 if a sentence addresses the scientific qualifications of the applicant, in particular the track record and the expertise to carry out the research project. Also, general mentions of the track record should be coded as track record. This can contain:

- scientific qualifications of applicants
- track record of applicants
- scientific publications (characteristics of journals, author positions, any metrics and/or citation counts)
- other research output
- supervision
- experience in leading a research team
- teaching activities
- memberships in panels, boards, etc., and individual scientific reviewing activities
- grants, prizes, awards, fellowships
- patents and/or licenses
- outreach activities (e.g. public engagement in science, technology and knowledge transfer activities, scientific art performances, presentations etc.)
- general contributions to science (e.g. spokesperson for experiments, leader of expeditions, founder of networks and training programmes, etc.)
- other artefacts with documented use (e.g. maps, methods, prototype demos, software, databases/datasets, design, arXiv-articles, contributions to big data collaborations, etc.)
- competences with regard to the ability to carry out the project, links between past activities and proposed research described (relevance of past activities for the proposed research)
- the applicants' independence/lack of independence
- descriptions of applicants' contributions to the field

Example (“yes” do code with “1”):

- *His CV is well-aligned with the proposed project, and his publications over the past five years reflect an ongoing interest and activity in the relevant field.*

1.2 Criterion_relevance

Definition/ question prompting coding:

Code 1, if sentence addresses:

- the scientific relevance of the proposed research project
- the scientific impact of the proposed research project
- whether the state of the art/work of third parties has been described accurately
- the topicality of the proposed research project (e.g. "research questions are timely", "this research topic is important right now")
- the originality of the proposed research project (e.g. novel, novelty)
- whether the proposed research project/approach/method is innovative.

Example ("yes" do code with "1"):

- *In my opinion the main topic of the proposed project is scientifically very relevant.*

1.3 Criterion_suitability

Definition/ question prompting coding:

Code 1 if the sentence addresses: the suitability of the methods to be used within the proposed research project, such as the methods chosen, their combination and the research plan.

Example ("yes" do code with "1"):

- *The randomized design is a key strength of this proposal, and the proposal does a fine job of laying out the empirical strategy.*

1.4 Criterion_feasibility

Definition/ question prompting coding:

Code 1 if the sentence addresses: the feasibility of the proposed research project such as targets or milestones, available resources, the scope of the project (workload) proportionate to the planned duration of the project, timing and logical sequence of steps.

Example ("yes" do code with "1"):

- *The proposed study is certainly feasible and has the potential to offer a very significant contribution to the academic community.*

2. Which parts of the SNSF application documents does the review refer to?

Definition/ question prompting coding:

Which parts of the application documents does the sentence address: the applicant and/or the proposal (in general) and/or the methods of the proposal?

We code for mentions of proposal and applicant and not the main focus of the sentence.

Relevance and informative value of the category:

Get understanding of if and how often a certain part of the application document is addressed and on which part a review is focusing on and in which context this is done, e.g. in combination with which criteria and as a strength, weakness or neutral statement.

Basis: [SNSF regulations on Project Funding, Art 15.](#)

2.1 Candidate_other

Code 1 if the applicant(s) or the team, or their qualifications are addressed. Do **NOT** code if only statements about quantity of certain achievements etc. of the candidate are made, for this code candidate_quantity.

Example ("yes" do code with "1"):

- *Their combined expertise has led them to propose a synthetic program of research that will have a much higher probability of novel contributions.*

2.2 Candidate_quantity

Code 1, if the sentence mentions the quantity of publications, or productivity of the applicant(s) with regard to publications. Also, code 1, to capture a broader idea of "quantity", i.e. mentions such as "a lot", "a big number", and similar expressions (not only numbers and indexes).

Relevance/ informative value:

Referring to purely quantitative indicators when assessing a candidate is a practice the SNSF wants to discourage. By identifying references to quantitative indicators, we gain better understanding of their prevalence and contexts they are used in (through combination of criteria, in relation to positive/negative statements etc.)

Examples ("yes" do code with "1"):

- *His main publications from the postdoctoral research (5 for the PNAS 201 and 2 for the JCB 2014 and his PhD work are cited well (15 for Cancer Research 2006 and 9 for Blood 2009).*
- *He has **published a lot**.*
- *He has shown **excellent productivity** in the form of publications over the past years.*
- *The applicant shows to have **many well-cited papers published** in topics related to the topic proposed here.*

Following the same sentence structure and characteristics, the references can be "negative". The sense of the sentence is not the key focus of the category but a clear mention of quantity.

- *However, that **single publication** is very important and relevant to the proposed research.*
- *They have both published papers at a **constant rate** over the last five years, even though **not at a high frequency**.*

Negative examples ("no" do code with "0"):

- *XXX is an original, wide-ranging, **extremely productive**, and unusually well-read scholar.*
- *In short, **the rate and level of his productivity is quite impressive** at this stage of his career.*
- *She has **a lot of output**.*

2.2 Proposal

Code 1 if the sentence addresses the proposed project either generally or specific parts of it. If the sentence addresses ONLY the methods, do not code proposal, but proposal_method. If both the proposal in general, or another non-method part of the proposal is addressed, code 1.

Also, references such as “why does it matter” / “it is built around two axes”, since it is unclear from the sentences alone what they refer to, be conservative and do not code as “proposal”.

Example (“yes” do code with “1”):

- *The proposal not only has a strong rationale, it also incorporates original ideas and questions that will be tested using modern approaches.*

2.3 Proposal_method

Code 1 if the methods are addressed.

Relevance/ informative value:

This category identifies sentences that are merely addressing the methods without applying the criteria “suitability” or “feasibility” but by giving context or summarizing what the applicants are planning to do.

Examples (“yes” do code with “1”):

- *To overcome the limitation of sustained expression of the BE system, the applicant proposes delivery of Cas9-BE by lipid nanoparticles that is not a novel approach but is attractive.*

To ensure this category also captures “methods” in the humanities, we added some examples coming from this discipline. Be aware that “methods” in the humanities may have different, potentially less explicit wording than fields in MINT, SS or LS. Some examples from methods related sentences in the humanities:

- *The research project intends to answer the research questions by means of ethnographic enquiries to be conducted in the three CR sites identified above.*
- *There would be one case study in Bologna - a finished memorial building, and one in Genève that will still be under construction while the research is happening.*
- *Furthermore, there is no information on how the annotated database should be used by researchers (as an interactive platform, repository, etc.?).*
- *Where are fieldwork locales to provide insights into financing (not funding) and the big ticket financialization schemes?*
- *Furthermore, the application does not explain under which conditions the stated knowledge production of artists is taking place or what the applicants understand by the term.*

3. Is the review making a positive, or negative statement? Does it provide suggestions?

Definition/ question prompting coding:

Is the sentence a positive statement or does it contain a positive statement on the research proposal or the applicant OR is the sentence a negative statement or does it contain a negative statement about the research proposal or the applicant?

Grammatically and semantically neutral sentences should NOT be coded as positive/negative even if they might be interpreted as positively or negatively by additional knowledge or bias.

Relevance/ informative value of the category:

This category indicates if a review is generally positive, negative or neutral in its tone. Combined with the other categories it indicates what entities the positive or negative statement refers to and if the application has been judged positive or negatively based on a certain criterion. Combined with the category “rationale” it indicates if the review discusses strengths or weaknesses of the proposal or the applicant.

Basis: [SNSF regulations on Project Funding, Art 15.](#)

3.1 Positive statement

Definition/ question prompting coding:

Code 1, if the sentence itself is a positive statement or if it contains a positive statement. Do not code positive if merely general statements.

Examples (“yes” do code with “1”):

- *This is an outstanding proposal on an important subject. [not substantiated positive statement about the proposal in general and the criterion “relevance, originality, topicality”]*
- *This is a nice proposal, and I am strongly recommending the support due to the following reasons. [Positive statement about the proposal in general, substantiated in the following sentences; no criterion applied]*
- *This application has therefore several highly innovative aspects. [positive statement about criterion “relevance, originality, topicality” that is substantiated in the previous sentence]*
- *She is a lecturer at a high-ranked university.*

Negative examples (“no” do code with “0”):

- *Dr. XXX is a lecturer at the University of Lausanne. [candidate; track record]*
- *In addition to her research activities, the applicant is also engaged in clinical work, teaching/mentoring and other institutional responsibilities. [candidate, track record]*
- *She has published in Nature. [candidate, track record]*

3.2 Negative statement

Code 1, if the sentence is a negative statement or contains a negative statement.

Examples (“yes” do code with “1”):

- *Although the proposal was strongly rooted in the literature, there was not a very strong linkage to fundamental research questions and hypotheses.*
- *My main reservation in this area is that the proposal could have said more on international aspects: domestic saving is not the only source of capital, and at least some goods are traded even in low-income countries, so the approach to this seems worth some thought.*

3.3 Suggestion

Code 1, if the sentence is suggesting how to improve the proposal. No coding for rationale needed.

Relevance/ informative value:

It captures not clearly stated weaknesses of the proposal, which we would otherwise miss.

Example (“yes” do code with “1”):

- *For Task 1 some preliminary data would be helpful, as well as some knowledge of what actually constitutes a positive signal that will be pursued.*

3.4 Is a clear rationale provided for a positive or negative statement?

Definition/ question prompting coding:

This category indicates whether there is a clear, nameable rationale for the positive or negative statement provided. (Test questions: why is this positive/negative, or how so?)

This is a subcategory ONLY applied if a positive or a negative statement has been identified in the target sentence.

The rationale can be given either in the target sentence itself, OR in the previous or following sentences. If it is NOT provided in the target sentence, but in one or more of the surrounding sentences, the category rationale_context should be coded.

If the rationale is provided within the to-be-coded sentence AND in one of the surrounding sentences both categories (e.g. _Rationale AND _Rationale_Context) need to be coded 1 (“yes”).

Previous sentence	To-be-coded sentence	Following sentence
<i>Rationale can be here</i>	<u>positive/ negative statement</u> <i>Rationale can be here</i>	<i>Rationale can be here</i>

The category rationale is NOT applied if the surrounding sentences contain the positive or negative statement and the target sentence itself is neutral but containing the rationale for the surrounding sentence, see the example:

Previous sentence	To-be-coded sentence	Following sentence
Positive/negative statement	<i>Neutral statement substantiating the positive/ negative statement</i>	Positive/negative statement

Relevance/ informative value:

This category informs if positive and negative statements are explicitly substantiated, e.g. if there is a clear reason (rationale) provided that explains the reviewers’ assessment.

Basis: “Substantiation” has been identified as an important part of peer review reports in the SNSF compliance report 2019/2020 to support transparency and scientific quality of funding decisions. “Lengthy texts are not necessary; key words that are understandable by other specialists of the respective field are sufficient”.

3.4.1 Rationale

Code 1, if the why and how of an identified positive or negative statement can be clearly stated. We do NOT judge the “quality” of the argument to avoid introducing our own interpretation or bias.

Example (“yes” do code with “1”):

- Given the comments above, this reviewer considers that the proposed methods are well suited to provide key answers to the proposed hypotheses, particularly because these methods and models have provided the background knowledge justifying the "reverse engineering" approach.

3.4.2 Rationale_Context

Code 1 if the rationale referring to the target sentence is provided in the previous or following sentence.

Examples ("yes" do code with "1"):

Previous sentence	To-be-coded sentence	Following sentence
	<i>This application has therefore several highly innovative aspects.</i> [positive statement]	<i>This is the sRNA_3 itself, which apparently is conserved in other gram-negative bacteria and may therefore also have a conserved role in stationary phase transitions.</i> [rationale provided]
<i>She has supervised about fifteen PhD students and most of them are now recognized researchers.</i> [rationale provided]	<i>She is undoubtedly one of the best international specialists in the themes addressed in this project.</i> [positive statement]	

Negative example ("no" do code with "0"):

Previous sentence	To-be-coded sentence	Following sentence
	<i>Overall, I feel this is an excellent proposal that will address timely and exciting questions on how the brain generates predictions.</i> [positive, proposal, relevance, not substantiated]	<i>The proposed study uses a combination of experimental and theoretic techniques, and will capitalise on an interdisciplinary expertise of the Applicant.</i> [proposal_method, neutral, candidate]

4. Is there a mention to an impact beyond academia?

4.1 Impact_beyond

Definition/ question prompting coding:

Code with 1, "yes", if a sentence refers to the effects that the proposed research may or may not have on society, policy, industry, environment, culture, health, services, quality of life etc. (e.g. the effects of research in the real world). This includes (non-evaluative) sentences, that merely mention these effects without making a positive or negative statement. We still want to capture them, as they indicate that the reviewer thought of the criterion impact_beyond.

Examples ("yes" do code with "1"):

- *If the conclusion of the study indicates that a thorough information of pelvic floor anatomy and individualized physiotherapy considering the reinforcement of muscle required during efforts the study **might contribute to change the paradigm of how this treatment option is offered to SUI women.***
- *Moreover, it will **contribute to critically evaluate, update and re-write the art education curriculum**, which has been long overdue in view of the contemporary deep currents in art and culture, technology and society.*
- ***To maximize opportunities for transfer of results to practitioners and decision-makers**, those stakeholders should be involved from the very start and be intimately involved in the choice of indicators and methods, driven by their needs for analysis.*
- *I don't have enough background knowledge **to judge the industrial impact** since I do not know the details of the state of the art methods of removing metal ions.*
- *Other potential impacts: **I would want to see more on societal impact of the project.***
- *On the contrary I'm convinced that it is **relevant for any city (not only european)** aiming at reconsidering its scopes and their coherence to its societal patterns.*
- *In addition to the broad theoretical impact I foresee for this research, there will **no doubt be major policy and management impacts.***
- *Transferability of results: To what degree can research results be **put into practice?***
- *It has to be noticed that **several broader impacts** are mentioned in the project.*

Negative examples ("no" do code with "0"):

- *I strongly suggest the team to emphasize how these activities can be generalized and applied to a broader scientific application especially for low-resource and endangered language analysis with more concrete research plans.*
- *Understanding the immunological response of dogs to infection, and the plausibility of vaccination will be an important proof of concept.*
- *Therefore, this sentence does not seem to hold: 'Therefore, any advances in restoring motor control in disease models and ultimately human patients will need to achieve equal restoration of function in the proprioceptive system.'*
- *The proposed intervention is short and low-cost, and thus may be implemented on larger scales relatively easily.*
- *The project aims to have a broader impact by way of creating an operationalizable theory for how to explain the work of ML systems to those using them or affected by them.*

Contributors in alphabetical order:

Matthias Egger
Anne Jorstad
Alberto de León
Katrin Milzow
Stefan Müller
Gabriel Okasa
Michaela Strinzel