Guide to the Qualitative Content Analysis of Spontaneous Word Associations Related to Artificial Intelligence

I. Introduction and General Coding Guide

Unit of coding: The basic unit of analysis is a single, already cleaned and lemmatized (brought to dictionary form) word or short phrase.

The coder's task: To each word or phrase, an appropriate deductive or inductive category or categories must be assigned. You will find the description of deductive and inductive categories and the difference between them in detail later in this document.

In this research, the terms waiting for coding originate from educators teaching in Catholic high schools in Hungary, but the codebook is suitable for the examination of other samples as well.

Researchers asked the educators to write at most five such concepts that describe their relationship to artificial intelligence. Accordingly, in every row of the database, the associations provided by a specific teacher can be found. Not every teacher provided five concepts; the number of expressions in the rows of the database can range from one to five.

During coding, the most important is to grasp the underlying, contextual meaning of the expression. The goal is to explore which religious-spiritual and ethical dimensions appear in the conceptual space of the teachers – with particular regard to the Catholic context – in connection with artificial intelligence.

In order for the analysis to simultaneously grasp the dominant meaning of the expressions and their potential additional meanings, the coding applies the system of "primary and secondary codes" according to the following:

a) Primary code (mandatory to provide): To every expression, a primary code must be assigned. The primary code is the category representing the dominant, central, primary meaning of the expression. This code will form the basis of statistical analyses. *Example:* In the case of the word "stupefies", the primary meaning refers to cognitive decline, the inhibition of thinking. Therefore, the primary code is: Category 3 (Contemplation and technological forcing out).

b) Secondary code (provision is optional): If an expression has a clear, but less emphasized secondary meaning than the primary meaning, the loss of which would result in information loss, there is a possibility to assign at most one secondary code. *Example:* In the case of the word "stupefies" (*elbutít*), cognitive decline also formulates a serious pedagogical-ethical concern. Therefore, the secondary code could be Category 5 (Ethical dimensions and moral attention).

General principles of coding:

a) Base every one of your decisions on the definitions and criteria assigned to the categories!
b) Choose the category that best captures the primary, dominant meaning of the word!
c) If a word could fit into multiple categories, consider which dimension is more emphasized. If the interpretation is uncertain, mark the most likely category, and record the ambiguity.
d) The work of human coders participating in the coding process serves a dual goal in this research.

The primary task is the execution of traditional qualitative coding. The secondary, but at least as important role is that the results of human coding serve as a validation set for checking the reliability of machine coding. Consequently, for the coding, it is explicitly forbidden to use large language models (artificial intelligence) or any automated assistant, or help. Human coding must be based on the researchers' own, expert interpretation! The methodological reason for this strict demarcation is that without it, the researchers would not be able to validly measure the agreement between human and machine coding. The latest research also proves that the validity of LLMs' coding can only be reliably established if we compare it to a reference sample based on pure, exclusively human interpretation.

II. Detailed description and example database of the coding categories determined via the deductive path

Deductive coding means that concepts must be classified into pre-given categories. In the current research, there are five deductive categories according to the following:

1. Disembodiment and dissociation
2. Anthropomorphization and human uniqueness
3. Contemplation and technological forcing out
4. Transcendence and immanence
5. Ethical dimensions and moral attention

In the interest of reliable coding, an operative definition (what to look for in the text?), theoretical background (why does it belong here?), as well as positive and negative examples (what to code here, and what not) belong to every category. In the following, I present the categories one by one.

**

1. Disembodiment and dissociation (Code: 1)

Such expressions belong here that point to the lack of physical presence, touch, living human relationship, or emotional warmth. Included here is the “cold,” “mechanical,” “sterile” nature of technology, as well as the feeling of detachment from reality, of escaping into the virtual world.
What to look for?

- Words pointing to physical/bodily lack (no presence, impersonal).
- Words indicating emotional distance (cold, rigid).
- Words pointing to the loss of reality (virtual, simulation, alienation).

Examples (If you see these -> Code: 1): “impersonal”, “mechanical”, “cold”, “no eye contact”, “robotic”, “emotionally empty”, “isolates”, “loneliness”, “virtual bubble”, “faceless”, “artificial”, “unnatural”, “aversion”, “alienation”, “automates”, “aloof” (*távolságtartó*), “aloofness” (*távolságtartás*), “reservation” (*tartózkodás*), “reserved” (*tartózkodó*), “emotionless”, “emotionlessness”, “artificial”, “isolation”, “sterile”.

Exclusion rule (What NOT to code here): If the word “mechanical” refers to children’s thinking becoming mechanical (not the lack of relationship, but the atrophy of thinking) -> that is 3 (Contemplation and technological forcing out). If the expression “soulless” or “has no soul” is used in a theological sense (i.e.: not human, the divine spark is missing from it), then it is code 2 (Human uniqueness).

**

2. **Category: Anthropomorphization and human uniqueness (Code: 2)**

Such expressions belong here that endow artificial intelligence with human qualities, personify it, OR specifically fear for human uniqueness (soul, creativity, *Imago Dei*) because of it. Included here is the blurring of the boundary between machine and human.
What to look for?

- Words presenting artificial intelligence as a partner, companion, or rival.
- Such qualities that, from a theological perspective, would only be due to humans (soul, conscience, true creation).
- References to the replaceability or irreplaceability of humans.

Examples (If you see these -> Code: 2): “smart companion”, “creative”, “new colleague”, “rival”, “has no soul”, “not human”, “soulless”, “imitation”, “playing God”, “human-replacer”, “outshines us”, “támasz” (support/crutch), “human redundancy”, “nyit” (opens), “autistic behavior”, “consumer”, “funny”, “rational”, “objective”, “sensible”.

Exclusion rule (What NOT to code here): If the word “smart” or “intelligent” only refers to the fact that artificial intelligence is an effective tool (for example “smart help for lesson plans”) -> that is 6 (Pragmatic tool, inductive). It is 2 only if the adjective “smart” suggests human-likeness. The same applies to the words “helper” and “assistance.” If the word “help” and “assistance” point to the fact that it helps as a companion or partner, brainstorming with the user, then the code is: 2. If as a tool it “helps,” speeds up the work, then the code is: 6 (Pragmatic tool). In the case of the words “reservation” and “reserved,” if “fear,” “danger,” or “distrustful” is also found, then the reservation can also stem from fear, therefore in these cases the primary code is: 1, and the secondary code is: 2.

**

3. **Category: Contemplation and technological forcing out (Code: 3)**

Such expressions belong here that refer to cognitive processes, learning, and the quality of thinking. One must think primarily of negative effects: the atrophy of thinking, laziness, superficiality due to speed, the “copy-paste” culture. Technology suppresses deep, slow thinking.
What to look for?

- Words pointing to the lack of mental effort.
- Devaluation of knowledge, superficiality.

- Cognitive laziness, “it thinks instead of me”.

Examples (If you see these → Code: 3): “stupefaction”, “laziness”, “lack of thinking”, “copy-paste”, “superficial”, “too fast”, “eased path”, “mental atrophy”, “creativity-killer” (if it refers to the skill), “convenience”, “makes one comfortable/lazy”.

Exclusion rule (What NOT to code here): If “speed” appears in a positive sense (efficiency), it will be code 6 (Pragmatic tool). If “copying” appears as an ethical sin (cheating), it is code 5 (Ethical dimensions and moral attention). If it is a cognitive problem (not learning to write), it is code 3.
If “harmful effects” or “unknown effects” are paired with fear, it is code 5. If these also refer to fear of stupefaction or laziness, it is secondary code 3. If the word “unnecessary” and/or “useless”/“superfluous” is paired with “stupefies,” then it is code 5, and it also belongs to the Contemplation and technological forcing out secondary category 3. If “convenient” does not appear in the sense of making one lazy or stupefying but refers to the speeding up or facilitating of work, then it is code 6 (Pragmatic tool).

**

4. **Category: Transcendence and immanence (Code: 4)**

Such expressions belong here that reflect the placement of artificial intelligence into a metaphysical, life-altering, or religious dimension. When the teacher looks at artificial intelligence not as a tool, but as a “higher power,” a “new era,” an “idol,” or “fate.” The relationship between technology and faith/God characterizes the words of the category.
What to look for?

- Words related to the future, fate, the end of the world.
- Religious terminology for AI (idol, false prophet, salvation, omnipotent).
- The “omnipotence” or inevitability of technology.

Examples (If you see these → Code: 4): “future”, “new world”, “inevitable”, “idolatry”, “new god”, “apocalypse”, “turning point”, “historical moment”, “omnipotent knowledge”, “modern”, “modernization”, “opportunity”.

Exclusion rule (What NOT to code here): Simple development, if it is of a technical nature, is code 6 (Pragmatic tool) or 9 (Personal growth, inductive). For code 4, a kind of “perspective” or “destiny-like quality” is needed.

The word “challenge” is 4 if “future” or “opportunity” is next to it, but “challenge” is 5 if it is paired with “fear” or “concern,” and it is 7 (motivation) if it can be brought into connection with cognitive motivation, “interest,” or curiosity. “Peak” (*hullámhegy*) and “trough” (*hullámvölgy*) can also refer to its life-altering nature if paired with the words “development” and/or “flow,” and if they refer to the fact that it hides matchless opportunities within itself.

**

5. Category: Ethical dimensions and moral attention (Code: 5)

Such expressions belong here that are about good and bad, about danger, about regulation, about responsibility, or about truth. Included here is the fear of losing control, as well as cheating and plagiarism as moral offenses.

What to look for?

- Concrete ethical offenses (cheating, lying, theft).
- Sense of danger, fear (if it is not specified what they fear, then generally existential/ethical fear).
- Questions of truth, authenticity, source criticism.
- The need for control, regulation.

Examples (If you see these → Code: 5): “danger”, “fear”, “cheating”, “plagiarism”, “lying”, “fake news”, “loss of control”, “ethics”, “responsibility”, “regulation is needed”, “unreliable”, “ambivalent”, “ambivalence”, “reservation”, “disappointment”, “error”, “limited”, “anger”, “disgust”.

Exclusion rule (What NOT to code here): If the context of “plagiarism” is that the student does not learn to formulate, it is code 3. If the context is dishonesty, it is code 5. If “ignorance” is paired with the words “concern” and “uncertainty,” and refers more strongly to ethical concern instead of stupefaction, then the primary category is 5 and the secondary category is 3. If next to the words “alienation” or “aversion” belonging to Category 1, any of the words belonging to code 5 appear among the answers, for example: “fear,” “scary,” “vulnerability,” “uncertainty,” “concern,” etc., then code 5 should also appear as a secondary category.

III. Further (inductive) categories identified during the empirical research

In the previous phase of the research (during human coding), the coders identified four further (inductive) categories emerging from the data. In the current coding phase, you must handle these categories (Categories 6-9) as equal, pre-determined categories, just like the deductive categories 1-5. The categories are the following:

6. Pragmatic tool
7. Cognitive and extrinsic motivation
8. Lack of experiential relationship
9. Personal growth

Coding rule:

a) First, examine whether the expression fits into one of the deductive categories 1-5.
b) If not, examine whether it fits into one of the inductive categories 6-9.
c) Create a new (10th or higher) category only if the meaning of the expression cannot be fitted into any of the categories 1-9 in any way.
d) We almost never assign a secondary code to inductive categories, even if the expression could be connected to one of the deductive categories as well. EXCEPTION: Category 6 (Pragmatic

tool), next to which a deductive secondary code (e.g., 2) may appear if the context justifies it (see the description of Category 6). This ensures the simplicity and consistency of the coding process, and preserves the conceptual purity of the inductive categories. Correct example: “efficiency” → Primary: Pragmatic tool (6), Secondary: empty. Incorrect example: “efficiency” → Primary: Pragmatic tool (6), Secondary: ethics (5).

The inductive codes ensure that conceptual dimensions important to the teachers but not covered by the theoretical framework are not lost during the analysis.

**

6. **Category: Pragmatic tool (Code: 6)**

Every word or expression belongs here that refers to the practical use or helpful nature of artificial intelligence. Such functions with which it makes the performance of work easier.
What to look for?

- If the expression refers to artificial intelligence as a tool.
- It speeds up, simplifies, or makes work performance more effective.
- If “help” appears in the sense of time-saving or usefulness.
- If it can be applied as a helping tool for language learning or scientific research.

Examples (If you see these → Code: 6): “fast”, “useful”, “tool”, “practicality”, “saves time”, “helping tool”, “work-acceleration”, “note-making”, “facilitation”, “searchability”, “aid”, “language”, “information”, “scientific research”, “systematization”, “varied editing”.

Exclusion rule (What NOT to code here): If “learning,” “help” refers to personal development, then it is code 9 (Personal growth). If development refers to the future, or that there is no stopping, then it is code 4 (Transcendence and immanence). If “idea” and “source of ideas” refer to a tool because it speeds up or facilitates work, then it is code 6 (Pragmatic tool), but if it refers to human qualities, “creative” and “resourceful,” then it is code 2 (Anthropomorphization and human uniqueness).

Note-making is basically a Pragmatic tool (6). However, if the answer refers to the fact that the machine performs this as an intelligent co-creator or human assistant (thus imitating a human role), then code 2 (Anthropomorphization) can also be given as a secondary code.

**

7. **Category: Cognitive and extrinsic motivation (Code: 7)**

Such expressions belong here that express that the relationship to artificial intelligence is determined by some kind of motivation. Cognitive motivation develops in the individual if they have an internal urge, if they find joy in cognition, are curious, or full of expectation toward the use of the new tool. Motivation can also be under the effect of external pressure, if it is a must,

inevitable, or necessary for work performance or learning. In such cases, extrinsic motivation develops. The lack of motivation, lack of interest, and indifference also belong here.
What to look for?

- Cognitive motivation, internal urge (curiosity, expectation, interest, joy).
- Extrinsic motivation (external pressure, necessity, inevitability).
- Lack of motivation (lack of interest, indifference).
-

Examples (If you see these → Code: 7): “curious”, “curiosity”, “interesting”, “interest”, “excitement”, “exciting”, “expectation”, “joy”, “necessary”, “must”, “need to”, “to be learned”, “to be taught”, “pressure” (*kényszer*), “indispensable”, “indifferent acceptance”, “indifference”, “don’t care”, “lack of interest”.

Exclusion rule (What NOT to code here): If “expectation” is paired with fear or uncertainty, it is not motivation, therefore it is code 5 (Ethical dimensions and moral attention).

**

8. **Category: Lack of experiential relationship (Code: 8)**

If someone cannot form an opinion because they have no experience of any kind with artificial intelligence, or have only little or indirect knowledge of it, then the word / expression does not appear as a qualitative evaluation, but refers to the deficiency of experience.
What to look for?

- Lack of experience or knowledge.
- Little, initial experience (getting acquainted).
- Knowledge acquired from second-hand (indirect knowledge).
-

Examples (If you see these → Code: 8): “none”, “I have no relationship with it”, “lack of experience”, “inexperienced”, “lack of knowledge”, “I am getting acquainted with it”, “getting acquainted”, “unknown”, “known from second-hand”, “rarely used”.

Exclusion rule (What NOT to code here): If the word refers to indifference or neutrality, that is not the lack of experience, but the lack of motivation, therefore it is code 7 (Cognitive and extrinsic motivation).

9. **Category: Personal growth (Code: 9)**

The expressions belonging here do not describe the relationship to artificial intelligence, but rather what kind of effect it has on the respondent’s personality, development, or learning.
What to look for?

- The word related to development describes not the quality of the machine, but its effect on the human, on the development of the human.

Examples (If you see these → Code: 9): “learning”, “broadening of knowledge”, “insight”, “reinforcement”, “opportunity for self-development”, “it will help me”, “personality-shaping effect”, “utilization”, “self-development”, “practice”, “critical thinking”, “preparedness”.

Exclusion rule (What NOT to code here): If in learning artificial intelligence is a tool and is not connected to personal growth, then it is code 6 (Pragmatic tool).

IV. The practical course of coding (use of the excel file)

During qualitative coding, the interpretation of context is indispensable. Therefore, the following rule must be strictly followed:

Before you code a single expression, read all the expressions belonging to the given respondent found in the same row, which the given respondent provided.

The words together form a conceptual web, which helps to more precisely understand the dominant meaning of the individual expressions. For example, the meaning of the word “challenge” is different if the words “fear” and “danger” stand next to it (likely code 5), and it means something entirely different if it appears in the environment of the words “development” and “learning” (likely the inductive code Personal growth).

You will receive an Excel file, which contains the 25% sample to be coded. Find the sheet designed for coding (Human Coder 1 and Human Coder 2). The structure of the sheet is the following:

Every row represents a single respondent. In one row, at most 5 expressions (Expression_1 to Expression_5) and the columns to be filled, Primary_Code (P) and Secondary_Code (SZ), belonging to them are found.

Proceed row by row. For every expression column, fill in the Primary_Code and (if necessary) Secondary_Code columns next to it with the numerical identifiers of the categories detailed below in point 5, or with a new, inductive code.

V. Numerical identifiers of the categories

To every single expression, assign the appropriate category or categories according to the rules of primary and secondary coding. This means that to the concepts provided by a given teacher (the number of which can range from 1 to 5), a total of 1–10 codes must be assigned, depending on whether only a primary, or both a primary and secondary code belongs to the individual expressions.

During coding, the following numerical identifiers must be used for the pre-determined categories:

1 = Disembodiment and dissociation
2 = Anthropomorphization and human uniqueness
3 = Contemplation and technological forcing out
4 = Transcendence and immanence

5 = Ethical dimensions and moral attention
6 = Pragmatic tool
7 = Cognitive and extrinsic motivation
8 = Lack of experiential relationship
9 = Personal growth

Table 1. Example for coding in a data row

Expression	Primary Code	Secondary Code
stupefies	3	5
danger	5	
partner	2	
tool	6	
future	4	5

If a respondent provided fewer than five expressions, leave the cells belonging to the missing expressions (Expression_X, Primary_Code_X, Secondary_Code_X) completely empty. Do not write anything, for example “N/A” or “null” text, into them.