# ISO/IEC JTC 1 SC 42 Artificial Intelligence – Working Group 4

**Use Case Submission Form**

# The quality of use case submissions will be evaluated for inclusion in the Working Group’s Technical Report based the application area, relevant AI technologies, credible reference sources (see References section), and the following characteristics:

* Data Focus & Learning: Use cases for AI system which utilizes Machine Learning, and those that use a fixed *a priori* knowledge base.
* Level of Autonomy: Use cases demonstrating several degrees (dependent, autonomous, human/critic in the loop, etc.) of AI system autonomy.
* Verifiability & Transparency: Use cases demonstrating several types and levels of verifiability and transparency, including approaches for explainable AI, accountability, etc.
* Impact: Use cases demonstrating the impact of AI systems to society, environment, etc.
* Architecture: Use cases demonstrating several architectural paradigms for AI systems (e.g., cloud, distributed AI, crowdsourcing, swarm intelligence, etc.)

1. **General**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ID | (leave blank, for internal use) | | | |
| Use case name | Deep Learning Based User Intent Recognition | | | |
| Application domain | Retail | | | |
| Deployment  model | On-premise systems | | | |
| Status | In operation | | | |
| Scope[[1]](#footnote-1) | Recognizing users’ intent to solve their problems in e-commerce fields | | | |
| Objective(s)[[2]](#footnote-2) | To recognize and understand users’ intent by AI and deep learning technologies and apply such technologies to build chat bot systems to further reduce labor cost and to be applied in various fields. | | | |
| Narrative | Short description (not more than 150 words) | Intelligent customer service chat bot is mainly used to categorize users’ questions, recognize users’ intents and answer users’ questions intelligently for different business jobs. Currently, this chat bot has been used to handle 90% of online customer service and has enabled JD.com to save over 100 million labor costs every year. | | |
| Complete description | JD.com has been committed to using technology to drive business growth and improve user experience in all customer service fields. Based on the improvement of customer consulting experience and the developing trend of artificial intelligence technology, as early as 2012, JD had decided to develop intelligent chat bots to fulfill the needs of continuous expansion of business, to save customer service costs and increase service capability. Intent recognition is a key and core technology to build such an intelligent customer service chat bot. By applying natural language processing technologies, deep learning technologies, traditional machine learning algorithms, intent recognition accuracy has reached to 95%. Based on accurate intents, and a series of solution finding algorithms, our chat bot can solve the user’s problems to a great extent and give the user a high quality consulting experience. Finally, in order to provide diversified and personalized customer services, we are continuously improving the accuracy of intent recognition, personalized solution generation, sentiment recognition, and image recognition. So far, intelligent customer service has revolutionized the traditional customer service consulting business. | | |
| Stakeholders[[3]](#footnote-3) | users | | | |
| Stakeholders’  assets, values[[4]](#footnote-4) | Users’ experience | | | |
| System’s threats & vulnerabilities[[5]](#footnote-5) | high semantic ambiguity, Multiple language expressions in one sentence | | | |
| Key performance indicators (KPIs) | ID | Name | Description | Reference to mentioned use case objectives |
| 1 | Accuracy | The number of correctly recognized users’ intent over total number of users. Currently, accuracy reaches 95%. | Improve accuracy of recognizing users’ intent |
| 2 | Resolution | The number of answers solved over total number of questions asked | Improve the resolution of questions from users |
| 3 | Satisfaction | The number of users who are satisfied with customer service over total number of users | Improve user experience |
|  |  |  |  |
| AI features | Task(s) | Natural language processing | | |
| Method(s)[[6]](#footnote-6) | Machine learning and deep learning | | |
| Hardware[[7]](#footnote-7) | GPU and CPU | | |
| Topology[[8]](#footnote-8) | TensorFlow | | |
| Terms and concepts used[[9]](#footnote-9) | Natural language processing, deep learning, CNN, HAN, logistic regression | | |
| Standardization  opportunities/ requirements | Process Standardization will Improve Quality and Productivity | | | |
| Challenges and issues | Current challenges of deep leaning and intent recognition:  1. high semantic ambiguity, similar sentences can deliver different meanings.  2. Unclear classification rules caused by complicated business logics  3. Hard to answer reasoning questions | | | |
| Societal  Concerns[[10]](#footnote-10) | Description | 1. Solve problems intelligently to increase efficiency 2. Free labors from repetitive work to save large amount of resources for the society | | |
| SDGs[[11]](#footnote-11) to be achieved | Decent work and economic growth | | |

**Data (optional)**

|  |  |
| --- | --- |
| Data characteristics | |
| Description | Question answering data from the JD.com online dialogue log |
| Source[[12]](#footnote-12) | Customer's dialogue log at JD.com |
| Type[[13]](#footnote-13) | Text |
| Volume (size) | Millions |
| Velocity[[14]](#footnote-14) | Real time |
| Variety[[15]](#footnote-15) | various scenarios, various business, various categories of products |
| Variability  (rate of change)[[16]](#footnote-16) | Non-linear |
| Quality[[17]](#footnote-17) | good |

**Process scenario (optional)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Scenario conditions | | | | | |
| No. | Scenario name | Scenario description | Triggering event | Pre-condition[[18]](#footnote-18) | Post-condition[[19]](#footnote-19) |
| 1 | Training | Based on millions of labeled streaming data, train a model using diversified algorithms, such as a deep learning neural network or a traditional machine learning algorithm | The training sample is ready |  |  |
| 2 | Evaluation | Evaluate the performance of the model on online dialogue data | The training procedure has been finished |  | Each requirement must be satisfied or exceeded to reach the condition of 'success' (e.g. the accuracy  should be more than 95%) |
| 3 | Execution | Apply the trained model to predict user’s intent | Require user’s query |  |  |
| 4 | Retraining | Take a training sample from online dialogue to retrain the model and compare it with the old one by AB test | bad cases are feed back to update the training dataset |  | The requirement is that the new model must be better than the old one |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

**Training (optional)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Scenario name | Training | | | | |
| Step No. | Event[[20]](#footnote-20) | Name of process/Activity[[21]](#footnote-21) | Primary actor | Description of process/activity | Requirement |
| 1 | Raw data stored in the database | Data extraction | Database engineer | Extract related data from the database to generate the raw dataset |  |
| 2 | Completion of Step 1 | Generating training samples | Data labeling team | Label the raw dataset of step one with 300 categories |  |
| 3 | Completion of Step 2 | Pre-process | AI engineer | Segment the sentence into words and convert those words into vectors |  |
| 4 | Completion of Step 3 | Model training | AI engineer | Based on vectors generated on step 3 to train a model using diversified algorithms, such as a deep learning neural network or a traditional machine learning algorithm |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
| Specification of training data | | After manual verifying, the accuracy of labelling should be more than 95% | | | |

**Evaluation (optional)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Scenario name | Evaluation | | | | |
| Step No. | Event[[22]](#footnote-22) | Name of process/Activity[[23]](#footnote-23) | Primary actor | Description of process/activity | Requirement |
| 1 | Certain period of time has passed since the last training/retraining | Data Extraction | Database engineer | Randomly take a sample from streaming data to form a test sample |  |
| 2 | Completion of Step 1 | Prediction | AI engineer | Predict the test sample in step 1 by the trained model |  |
| 3 | Completion of Step 2 | Evaluation | Data labeling team | Compare the result of predicted with the result of labeling |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
| Input of evaluation | | the result of labeling and the result of prediction | | | |
| Output of evaluation | | The accuracy and recall rate | | | |

**Execution (optional)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Scenario name | Execution | | | | |
| Step No. | Event[[24]](#footnote-24) | Name of process/Activity[[25]](#footnote-25) | Primary actor | Description of process/activity | Requirement |
| 1 | Acquire the user’s query | pre-process | AI engineer | Segment the sentence into words and convert those words into vectors | The trained model has been in operation |
| 2 | Completion of Step 1 | Text classification | AI engineer | Predict the label of user’s query |  |
| 3 | Completion of Step 2 | Response | AI trainer | Answer the query based on the result of intent classification |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
| Input of Execution | |  | | | |
| Output of Execution | |  | | | |

**Retraining (optional)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Scenario name | Retraining | | | | |
| Step No. | Event[[26]](#footnote-26) | Name of process/Activity[[27]](#footnote-27) | Primary actor | Description of process/activity | Requirement |
| 1 | Certain period of time has passed since the last training/retraining | Data extraction | Database engineer | Randomly take a sample from streaming data to from a training sample |  |
| 2 | Completion of Step 1 | Labeling the sample | Data labeling team | Manually label the sample data |  |
| 3 | Completion of Step 2 | Model training | AI engineer | Combine the new training sample with the old and train a model (deep learning and machine learning) |  |
| 4 | Completion of Step 3 | AB Test | AI engineer | Compare the predicted results of the new model with the results of the old one | The performance of the new model is better than results of the old one |
| 5 | Completion of Step 4 | Online active of new model | AI engineer | The new model is been active online at JD.com |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
| Specification of retraining data | | After the calibration, the accuracy of labelling should be more than 95% | | | |

**References**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| References | | | | | | |
| No. | Type | Reference | Status | Impact on use case | Originator/organization | Link |
| 1 | Paper | Convolutional Neural Networks for Sentence Classification |  |  | New York University | https://arxiv.org/abs/1408.5882 |
| 2 | Paper | Hierarchical Attention Networks for Document Classification |  |  | Carnegie Mellon University, Microsoft Research, Redmond | http://www.aclweb.org/anthology/N16-1174 |
| 3 | Paper | LIBLINEAR: A library for large linear classification Journal of Machine Learning Research |  |  | National Taiwan University | http://www.jmlr.org/papers/volume9/fan08a/fan08a.pdf |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

# Acceptable Reference Sources of Use Cases

* Peer-reviewed scientific/technical publications on AI applications (e.g. [1]).
* Patent documents describing AI solutions (e.g. [2], [3]).
* Technical reports or presentations by renowned AI experts (e.g. [4])
* High quality company whitepapers and presentations
* Publicly accessible sources with sufficient detail

***This list is not exhaustive. Other credible sources may be acceptable as well.***

## Examples of credible sources:

* [1] B. Du Boulay. "Artificial Intelligence as an Effective Classroom Assistant". IEEE Intelligent Systems, V 31, p.76–81. 2016.
* [2] S. Hong. "Artificial intelligence audio apparatus and operation method thereof". N US 9,948,764, Available at: [https://patents.google.com/patent/US20150120618A1/en. 2018](https://patents.google.com/patent/US20150120618A1/en.%202018).
* [3] M.R. Sumner, B.J. Newendorp and R.M. Orr. "Structured dictation using intelligent automated assistants". N US 9,865,280, 2018.
* [4] J. Hendler, S. Ellis, K. McGuire, N. Negedley, A. Weinstock, M. Klawonn and D. Burns. "WATSON@RPI, Technical Project Review".

URL: [https://www.slideshare.net/jahendler/watson-summer-review82013final. 2013](https://www.slideshare.net/jahendler/watson-summer-review82013final.%202013).

1. The scope defines the intended area of applicability, limits, and audience. [↑](#footnote-ref-1)
2. The intention of the system; what is to be accomplished?; who/what will benefit?. [↑](#footnote-ref-2)
3. Stakeholder are those that can affect or be affected by the AI system in the scenario; e.g., organizations, customers, 3rd parties, end users, community, environment, negative influencers, bad actors, etc. [↑](#footnote-ref-3)
4. Stakeholders’ assets and values that are at stake with potential risk of being compromised by the AI system deployment – e.g., competitiveness, reputation, trustworthiness, fair treatment, safety, privacy, stability, etc. [↑](#footnote-ref-4)
5. Threats and vulnerabilities can compromise the assets and values above - e.g., different sources of bias, incorrect AI system use, new security threats, challenges to accountability, new privacy threats (hidden patterns), etc. [↑](#footnote-ref-5)
6. AI method(s)/framework(s) used in development. [↑](#footnote-ref-6)
7. Hardware system used in development and deployment. [↑](#footnote-ref-7)
8. Topology of the deployment network architecture. [↑](#footnote-ref-8)
9. Terms and concepts used here should be consistent with those defined by Working Group 1 (AWI 22989 and AWI 23053) or to be recommended for inclusion. [↑](#footnote-ref-9)
10. To be inserted. [↑](#footnote-ref-10)
11. The Sustainable Development Goals (SDGs), also known as the Global Goals, are a collection of 17 global goals set by the United Nations General Assembly. SDGs are a universal call to action to end poverty, protect the planet and ensure that all people enjoy peace and prosperity.

    URL: <http://www.undp.org/content/undp/en/home/sustainable-development-goals.html> [↑](#footnote-ref-11)
12. Origin of data, which could be from customers, instruments, IoT, web, surveys, commercial activity, simulations, etc. [↑](#footnote-ref-12)
13. Structured/unstructured text, images, voices, gene sequences, numbers, composite: time-series, graph-structures, etc. [↑](#footnote-ref-13)
14. The rate of flow at which the data is created, stored, analysed, or visualized. Could be in real time. [↑](#footnote-ref-14)
15. Domains and types of data employed including formats, logical models, timescales, and semantics. Could be from multiple databases. [↑](#footnote-ref-15)
16. Changes in data rate, format/structure, semantics, and/or quality. [↑](#footnote-ref-16)
17. Completeness and accuracy of the data with respect to semantic content as well as syntax of the data (such as presence of missing fields or incorrect values). [↑](#footnote-ref-17)
18. Describes which condition(s) should have been met before this scenario happens. [↑](#footnote-ref-18)
19. Describes which condition(s) should prevail after this scenario happens. The post-condition may also define "success" or "failure" conditions [↑](#footnote-ref-19)
20. The event that triggers the step. This might be completion of the previous event. [↑](#footnote-ref-20)
21. Action verbs should be used when naming activity. [↑](#footnote-ref-21)
22. The event that triggers the step. This might be completion of the previous event. [↑](#footnote-ref-22)
23. Action verbs should be used when naming activity. [↑](#footnote-ref-23)
24. The event that triggers the step. This might be completion of the previous event. [↑](#footnote-ref-24)
25. Action verbs should be used when naming activity. [↑](#footnote-ref-25)
26. The event that triggers the step. This might be completion of the previous event. [↑](#footnote-ref-26)
27. Action verbs should be used when naming activity. [↑](#footnote-ref-27)