

2020 Formosa Grand Challenge – Talk to AI Competition Prospectus

I. Competition purpose

Artificial Intelligence (AI) is accelerating the change of global development trends and patterns. In order to achieve the goal of human-machine interaction in the future, it is necessary for AI to equip machines with linguistic competence. Therefore, AI needs diversified language training, from simple to complex. However, understanding languages requires knowledge of vocabulary, semantics and grammatical structure. In addition, it is sometimes easy to mistake some Chinese words for other similar sounding words. By hosting the 2020 “Formosa Grand Challenge-Talk to AI”, the Ministry of Science and Technology encourages innovators to utilize the potential, technology and creativity of AI development to address the challenges of voice applications and looks forward to all possible progress and ideas.

II. Competition theme

In this competition, “Band C” of the “Test of Chinese as a Foreign Language” (hereinafter referred to as TOCFL) is set as the level of difficulty as in the previous competition. The design of the reading comprehension test refers to the TOCFL, and the questions include short-answer questions and dialogue. It is hoped that the AI program (model) can convert a large amount of data into applicable knowledge through the process of reading comprehension, and finally achieve the goal of simulating human self-thinking.

The competition questions are divided into two categories: “short-answer questions” and task-oriented “multi-round dialogue”. “Short-answer questions” are based on textual content for reasoning; “Multi-round dialogue” focus on task-specific natural language dialogues, such as ordering travel tickets, hotel room types, car rental services, etc., which require step-by-step inquiry and back-and-forth dialogue between the customer and the customer service personnel to complete the order (task).

In this competition, the organizer entrusts the team led by Dr. Wen-Lian Hsu, distinguished research fellow at the Institute of Information Science, Academia Sinica, to design these questions. The aim is to test the technical level of the participating teams at all aspects of natural language processing, including but not limited to information retrieval and extraction, unknown word processing, especially named entity recognition and new word recognition, and inference, such as synonym and entity linking, and even distinguishing between semantic role relationships. An example of the competition question is as follows:

An example of a short-answer question

No.	Article	Time	Source	Portal	Content	Words	Question
D002	Taiwan Green Gold	2019/04/15	News	Liberty Times	<p>The domestic export value of domestic edamame reached a new 28-year high, with an external sales volume of 37,520 metric tons and a trade volume of US\$81.18 million (about NT\$ 2.45 billion). It also helped 7 frozen food companies to build new plants or expand factories in Taiwan. With an investment of NT\$ 3 billion and thousands of job opportunities, it is called “Taiwan Green Gold”!</p> <p>“Edamame” is the immature and greenish-colored edible soybean. When the fresh pods of the whole plant are 80% full, the pods are green with fuzz, so it is called “edamame”, also known as “vegetable soybean”, and it is called “edamame bean” in Japan. Generally, beans contain cottonseed sugar, which may cause flatulence. However, edamame has a low amount of cottonseed sugar, but is rich in potassium, which can improve exhaustion and loss of appetite caused by the lack of potassium ions. It is often used as an appetizer; it has more high-quality protein than ordinary beans, and is called plant meat, which is suitable for vegetarians to supplement nutrition.</p> <p>Shun-fa Tai, director of the Kaohsiung District agricultural improvement plant of the Council of Agriculture, points out that in the past 10 years, under the cooperation of Mr. Chou, Kuo-lung and the farmers from Kaohsiung and Pingtung areas in the Taiwanese edamame industry, except for 85.4% sold to Japan, it is also sold to 24 countries including the United States and Canada. Tai, Shun-fa explains that the output value of Taiwanese frozen edamame products has reached 44.8% in Japan, which is 1.63 and 1.91 times of competitors from Thailand and China, respectively. The average price per kilogram is 250 yen, which is 31.9% higher than China’s price of 189 yen.</p> <p>Tai, Shun-fa emphasizes that the Kaohsiung plant has outstanding advantages in variety research and development, and launches high-yield</p>	576	<ol style="list-style-type: none"> 1. Is edamame a Taiwanese cash crop? 2. Which plant is called “Taiwan Green Gold”? 3. From what year did the edamame export start? 4. What other names are there for edamame? 5. What is the cause of flatulence from eating beans? 6. In the text, which nutrients are found in edamame? 7. Is eating edamame healthy? 8. To which countries are Taiwanese edamame exported? 9. Where is the production area of Taiwanese edamame concentrated? 10. What are the two varieties of edamame that have applied for variety patents? 11. What are the advantages of edamame as the “Taiwan Green Gold”?

					edamame varieties such as “Kaohsiung No. 9 - Green Crystal” and “Kaohsiung No. 11 - Xiangmi Tea Bean” and authorizes industrial applications. At the same time, in order to protect intellectual property rights, in addition to applying for domestic variety rights, it also applies for variety rights in Japan, so that Taiwanese edamame can jump ahead of market price competition, and gain advantages in innovation and development progress. It also proves that the continuous promotion of industrial upgrading is the policy of sustainable management.	
--	--	--	--	--	---	--

An example of multi-round dialogue

Basic information needed to answer the task-oriented multi-round dialogue is provided to assist the participants in this task. Take the “Hotel Room Reservation” task as an example. Information such as “Hotel Basic Information”, “Room Type and Price”, “Available Date and Room Number” and “Customer Basic Information”, etc., are provided. The participant of this competition assumes the role of the Agent to start a dialogue, respond to the inquiry of the User (the competition organizers), and closes the conversation with common customer service greetings at the end. The entire log of a dialogue task is shown in the Table “Customer Demand Dialogue Content”.

“Hotel Basic Information”

Hotel Name	Summer Business Inn
Address	No. 16, Sec. 7, Zhongxiao E. Rd., Nangang Dist., Taipei City
Contact number	02-2600-0000
URL	https://summer.taipeihotels.website/
Check-in time	14:00
Check-out time	12:00

“Room Type and Price”

Room type	Square metre	Bed type
Deluxe Room	35	King size bed
Skyline Room	35	King size bed
Deluxe Suite	40	King size bed + sofa bed

Mini Suite	33	Double bed
------------	----	------------

“Available Date and Room Number”

Date	Week	Deluxe Room	Skyline Room	Deluxe Suite	Mini Suite
2020.9.1	Sunday	5	2	3	3
2020.9.2	Monday	5	3	5	5
2020.9.3	Tuesday	3	2	1	3
2020.9.4	Wednesday	2	5	1	1
2020.9.5	Thursday	1	2	2	2
2020.9.6	Friday	4	3	3	1
2020.9.7	Saturday	2	2	3	1
2020.9.8	Sunday	1	4	4	3
2020.9.9	Monday	3	2	3	2
2020.9.10	Tuesday	5	3	5	4
2020.9.11	Wednesday	4	2	2	3

“Basic Customer Information”

Customer name	Miss Lin
Customer gender	Female
Contact number	0912-345-678

“Customer Demand Dialogue Content”

Dialogue number	Dialogue role	Content
1	Agent	Hello, this is the Summer Business Inn. May I help you?
2	User	I want to make a reservation
3	Agent	OK. May I have your name, please?

Appendix included in the Prospectus: Award Description

4	User	My surname is Lin
5	Agent	Hello Miss Lin, when do you expect to check in?
6	User	In the evening on September 9
7	Agent	How long do you plan to stay?
8	User	2 nights
9	Agent	How many guests are there?
10	User	Five
11	Agent	What kind of rooms would you like to have?
12	User	One deluxe suite and one deluxe room
13	Agent	Alright, your check-in date is Monday, September 9 and check-out date is Wednesday, September 11. Could you please confirm your reservation?
14	User	Ok
15	Agent	Great, check-in on Monday, September 9 and check-out on Wednesday, September 11. The room type booked is one deluxe suite and one deluxe room for two nights.
16	User	That's right
17	Agent	Could you please tell me the name of the contact person?
18	User	Lin Chi-ling
19	Agent	May I have your phone number
20	User	0912-345-678
21	Agent	Let me confirm the reservation information with you as follows. Contact person: Lin Chi-ling Number of guests: 5 Contact number: 0913-345-678 Number of different room types: Deluxe Suite, Deluxe Room Check-in date: September 9 Check-out date: September 11 Is the information correct?
22	User	Yes, thank you.

23	Agent	Please be noted that our check-in time is 14:00 and check-out time is 12:00. Thank you for choosing the Summer Business Inn. Is there anything else that I can help you with ?
24	User	No, thank you.
25	Agent	It's my pleasure to serve you. If you have any questions, please feel free to contact us! I wish you all the best.

III. Competition system

The competition is divided into a preliminary and a final round. The top 10 teams will advance to the final round according to the scores in the preliminary round, and the teams will receive awards according to the final score.

1. The short-answer questions are used in the preliminary round tests, and the final round includes the short-answer questions and the dialogue tasks. Detailed question types, input format and content are shown in the table below. For multi-round dialogue tasks, 5 of the 20 task-oriented topics will be selected by the organizers. The speech recognition is all performed by the “Ya Ting” speech recognition system created by the Taiwan AI Labs. The input and answer output of the short-answer questions and multi-round dialogue tasks are in text format, and the competition is graded in text format. For details of the competition format, please refer to Section VII “Competition environment” of this prospectus.

Competition type	Question types	Question text	Proportion	Score	Number of questions	Question description
Preliminary	Elementary	Short-answer question text	40%	40	20	Require basic information retrieval techniques or additional semantic resources such as “synonyms”, “words with appositive concept”, “antonyms”, “negation words”, etc., to answer these questions.
		Ya Ting voice recognition result	10%	10	5	Similar to the above, except that the text comes from speech recognized by the “Ya Ting” system, which may contain typos.
	Advanced	Short-answer question text	40%	40	20	In addition to basic understanding, it is necessary to further add inferential techniques to enable advanced understanding such as “kinship”, “entailment”, “temporal relation”, “spatial relation”, “causal relation”, “enumeration relation” and other inferences.
		Ya Ting voice recognition result	10%	10	5	Similar to the above, except that the text comes from speech recognized by the “Ya Ting” system, which may contain typos.

Competition type	Question types	Question text	Proportion	Score	Number of questions	Question description
Final	Elementary	Short-answer question text	30%	30	19	Require basic information retrieval techniques or additional semantic resources such as “synonyms”, “words with appositive concept”, “antonyms”, “negation words”, etc., to answer these questions.
		Ya Ting voice recognition result	6%	6	4	Similar to the above, except that the text comes from speech recognized by the “Ya Ting” system, which may contain typos.
	Advanced	Short-answer question text	30%	30	19	In addition to basic understanding, it is necessary to further add inferential techniques to enable advanced understanding such as “kinship”, “entailment”, “temporal relation”, “spatial relation”, “causal relation”, “enumeration relation” and other inferences.
		Ya Ting voice recognition result	6%	6	4	Similar to the above, except that the text comes from speech recognized by the “Ya Ting” system, which may contain typos.
	Argumentation	Short-answer question text	4%	4	2	Combine basic and advanced understanding techniques. Also, summarization, information integration and text generation technologies are needed.
		Ya Ting voice recognition result	4%	4	2	Similar to the above, except that the text comes from speech recognized by the “Ya Ting” system, which may contain typos.
	Task-oriented multi-round dialogue		20%	20	5	According to the “Task Goal” and the “Basic Information of Dialogue” given by the question, a fluent dialogue should be generated and the goal achieved.

2. The answer must be completed within the time limit. In case of special circumstances, the organizer may extend or shorten the competition time.
3. The organizer will hold the final at the designated time and place. Each team must send more than two members (including the representatives of each team) to the competition site. If the number of participants does not meet the designated attendance or the team representative is absent, the team will be considered to have forfeited. If the total number of team members is one, it is not subject to this limit.
4. The competition test design covers textbooks, modern Chinese writing texts, news corpus, Wikipedia (version of August 20, 2019) and other resources.

Wikipedia version download URL:
<https://dumps.wikimedia.org/enwiki/20190820/>

5. Training materials will be published after September 2019, as detailed in the official website announcement. The participating teams may use any self-acquired learning materials, not limited to the materials provided by the organizer.
6. The input and output of the program model shall comply with the requirements of the organizer and shall not affect the competition result. No replacement or compensation may be required in any way.

IV. Competition schedule

Schedule	Item	Content
2019.09	Online registration	Time: From now on to 2019.10.07 23:59 (UTC + 8 hours) Registration URL: https://fgc.stpi.narl.org.tw/activity/2020_Talk2AI , Registration information is processed in accordance with the Personal Data Protection Act.
2019.10	Formosa Grand Challenge orientation meeting	Before the competition, the Formosa Grand Challenge orientation meeting will be held to help participants understand the test design method, rules, goals and scoring standard. Participants are required to complete the registration before they can attend <u>free of charge</u> . The workshop will be videotaped for use by participants who are not able to attend. The workshop agenda is available on the official website. Workshop time: 2019.10.07 (Monday) 14:00-17:00 Workshop location: Academia Sinica Institute of Information Science (No. 128, Sec. 2, Academia Rd., Nangang Dist., Taipei City) Auditorium106 at IIS new Building
2019.12	Preliminary	The online competition method is adopted, and the results are sorted according to the scores. The preliminary results and qualifications are reviewed by the organizer, and the teams advance from the preliminary round to the final.

2020.04	Final	<p>The on-site competition method is adopted, and the awards are issued to teams that pass the threshold scores. The final ranking and the qualifications for the awards are reviewed by the organizer, and the awards can be vacant when necessary. For details, please refer to Section VI “Scoring method” of this prospectus and Section VIII “Awards” and the Appendix “Award Description”.</p> <p>The time and location of the competition are undecided and will be announced on the official website.</p>
---------	-------	--

V. Entry qualification

1. The competition can be registered by teams, and the number of members is limited to 20. At least one of the team members shall have the nationality of Taiwan. Members whose nationality is not Taiwan shall hold a Taiwan work permit or student permit. Relevant permit certificates (Note 1) must be attached at the time of registration, and the team registration will take effect. The above qualification should be obtained on or before the registration deadline date and shall be valid for the period of the competition.
2. In order to ensure the fairness of the competition, each entrant can only participate in one team, and no team members can be changed during the competition. If the above rule is violated, the entrant’s qualification shall be cancelled immediately. After the affected team removes the member, the team composition shall still meet the eligibility requirements. The team that does not meet the requirements shall be disqualified.
3. Each team should designate a member as the representative. For the rights and obligations of the competition (including but not limited to the notices sent by the organizer and the distribution of awards, etc.), this representative is deemed as the contact for notification and acceptance.
4. The registered name will be used during the competition and no inappropriate words should be used.
5. From the appointment date of the question-setting team by the organizer (August 15, 2019) to the end of the competition, if the entrants and any member of the question-setting team have one of the following circumstances, they should recuse themselves and may not participate in this competition.
 - 5.1 Self, or spouse, ex-spouse, blood relation or relation by marriage within three generations
 - 5.2 Teachers and students with the relationship of doctoral or master thesis guidance

5.3 Joint plan implementation, paper or research result publishing

5.4 There is an employment or appointment relationship

6. If the entrant violates the above regulations, the organizer may cancel the entry qualification of the team or forfeit the awards at any time. The remaining members shall not raise any objection.

Note 1: Certificate of employment: Work permit or residence certificate issued by the Ministry of Labor (the residence reason must be stated as employment).

Student status certificate: Student ID card (with semester registration stamp). If the student ID card is exempt from the registration stamp, the studying certificate of that semester should be provided; if the applicant is temporarily absent from school (but still retains student status), the studying certificate should be provided. The above certificates are applied by the applicants to the schools directly; self-study students are requested to provide the “non-school type experimental education” approval documents.

VI. Scoring method

The competition is scored in text format. The short-answer questions and multi-round dialogue tasks are reviewed and scored by the jury. The answering criteria are as detailed below. Details of the answering content, format and scoring method will be announced at the Formosa Grand Challenge orientation meeting and on the official website.

Scoring judgment	Scoring criteria
Incorrect	Irrelevant answer, incomplete sentences, wrong sentence structure, unclear and wrong meaning
Correct	Appropriate description, fluent sentence, high quality content information, precise words, or with extended of knowledge.

*The jury consists of at least five experts and scholars in the fields of AI, Chinese language and reading comprehension.

VII. Competition environment

1. The organizer provides competition environment with CUDA toolkit on Ubuntu. The participating teams are limited to deploying all the programs and completing the test questions of the official competition (including the preliminary and final) in this environment. The detailed environment specifications are announced on the official website.
2. The team must complete the system installation within the agreed time stated by the organizer, and complete the competition in the context of disconnecting the external; the answering method of the competition is formulated by the organizer, and the simulation test shall be conducted within the agreed time.
3. The following example shows how to use JSON to store information related to team based on their question and answer.

```
{
  "id": "D003-1",
  "subject": "Taiwan Green Gold",
  "time": "2019/04/15",
  "source": "Liberty Times",
  "type": "Life",
  "document": "Domestic edamame last year... (omitted)",
  "questions": [
    {
      "q_id": "D003-1-Q01",
      "q_text": "Whether edamame... (omitted)",
      "answer": "Yes"
    },
    {
      "q_id": "D003-1-Q02",
      "q_text": "There is \"Taiwan... (omitted)\"",
      "answer": "Edamame "
    },
    ...
  ]
}
```

The JSON below illustrates a team's answer.

```
{
  "q_id": "D003-1-Q01",
  "answer": "Yes"
}
```

VIII. Awards

The requirement for the final awards is that the AI program (algorithmic model) must score higher than the human pilot test score. The pilot test method for the final is discussed by the Steering Committee for the Test of Proficiency-Huayu (SC-TOP) and the result of the test is completed and announced by the executive unit before the final. The final score is calculated as the total score of the short answer questions and the dialogue tasks.

The competition award totals NT\$ 25 million. The winners of the final will be

presented with the award and the honor certificate by the organizer. The top two winners will be determined based on the score ranking of the team AI programs that score higher than the human pilot test score. If the team score is equal to or lower than the human pilot test score, 80% (rounded to the whole number) of the score is used as the benchmark, and the highest score team of this benchmark can receive the special award of NT\$ 5 million. If no team reaches 80% of the human pilot test score, the top three teams will receive special awards according to the ranking of the scores.

The winning team shall cooperate with the organizer for public demonstration (such as post-match technical exchanges, etc.). Please refer to the Appendix “**Award Description**”. The award distribution method is as follows:

Final award benchmark	Competition award
The highest score of the team is greater than ($>$) the human pilot test score	1st place: NT\$ 20 million 2nd place: NT\$ 5 million
The highest score of the team is equal to or less than (\leq) the human pilot test score, but greater than or equal to (\geq) 80% of the human pilot test score	1st place: NT\$ 5 million (special award)
The highest score of the team is less than (\leq) 80% of the human pilot test score	1st place: NT\$ 2 million (special award) 2nd place: NT\$ 500,000 (special award) 3rd place: NT\$ 300,000 (special award)

* If only one team scores higher than the human pilot test score, only first place will be awarded.

* In the case of a tied score, the teams with the same score will receive an average of the total available award. For example, if there are two teams in first place, they will receive an average of the total top two awards.

IX. Important Notices

1. The organizer reserves the right to change the relevant rules and schedule of the competition, and to explain the rules of the competition. If necessary, each award can remain vacant.
2. It is forbidden to answer questions manually during the competition. If there are fraudulent situations such as fabrication, falsification or plagiarism, the organizer may disqualify the team and cancel all its awards. The participating team should also bear all the relevant legal responsibilities on its own.

3. The materials, techniques and program codes used in the entries must be original or have been legally authorized, or the open source code is publicly available. If any third party claims infringement of intellectual property rights or other illegal matters, the participating team shall bear all the relevant legal responsibilities on its own, and the organizer may disqualify the team.
4. Those who have been selected for the competition awards (excluding special awards) are required to submit their award-winning program code and their derivative works (including the program model) copyright free and authorize use by the National Experimental Research Institute Laboratories (hereinafter referred to as NARLabs) for non-profit purposes. Its scope of authority includes the right to test, verify, reproduce and publicly demonstrate the finished products. For the subject of the foregoing authorization, including the self-developed program model, if the program uses the subject of legal authorization by a third party, the winner shall ensure that two years from the time of the announcement of the award, NARLabs can legally and freely demonstrate its finished products, and no other third party has the right to make any claims against NARLabs. NARLabs can reproduce the products for the purpose of public disclosure. The winning team is also required to assist the organizer to check and reproduce the competition result. If the team cannot provide the environment to reproduce the competition result, all of its awards will be cancelled.
5. The organizer has the right to request the team to provide the program codes and related materials for verification, and to entrust a credible third party to carry out code review and reproduction. If it is impossible to reproduce the same answer as in the competition or there are fraudulent situations such as fabrication, plagiarism or fraud, the organizer may disqualify the team and cancel all its awards. The participating team shall bear all the relevant legal responsibilities on its own.

X. Contact Information

1. Science & Technology Policy Research and Information Center
2. 1F, No. 106, Sec. 2, Heping E. Rd., Taipei City
3. The official website of the Formosa Grand Challenge: <https://fgc.stpi.narl.org.tw/>
4. E-mail: grandchallenge@stpi.narl.org.tw
5. FB Fan Group: Formosa Grand Challenge

Organizer: Ministry of Science and Technology

Executive unit: Science & Technology Policy Research and Information Center (STPI)