

# Patent obtainment in Japan: analysis of AI-based drug discovery

Masato Iida of [Shiga International Patent Office](#) explains the patent eligibility of AI-related inventions in Japan

---

**D**rug discovery requires long-term research and myriads of money from pharmaceutical companies. As of 2020, the Pharmaceuticals and Medical Devices Agency (PDMA) approved 123 new drug products in Japan. Drug approval takes about 10.6 months on average from submission of application for marketing approval.

Today, artificial intelligence (AI) is expected to be a promising tool even for drug development. AI reduces time-consuming efforts of searching and identifying the most preferable candidate drugs from enormous quantities of experimental data by pharmaceutical companies. The pharmaceutical industry has hastened to implement digital transformation of drug discovery and development in response to a dire need to reform the research process and improve production capability.

For example, Chugai Pharmaceutical created a self-made AI-based antibody discovery technology named 'Malexa-Li' which optimises the antibody discovery process with training data of a machine learning model that was provided by analysing sequence information accumulated from the company's research conducted over many years. Sumitomo Dainippon Pharma completed an exploratory phase of drug discovery within 12 months together with Exscientia, a UK-based AI-drug discovery company. That exploratory research often takes four and a half years on average.

What if a company wishes to file a patent application related to a novel drug derived from AI in Japan? Several case examples are analysed below and some practical tips are given to consider when filing patent applications related to a combination of AI and pharmaceuticals. Generally, an AI-related invention, regardless of drug discovery or computer programs, is deemed to be a computer-related invention in Japan.

The first part of this article explains patent eligibility of AI-related inventions in Japan, and in the latter part, specific case examples are explored.

### Patent eligibility of AI-related inventions

Computer/software inventions are categorised into the following four types of inventions in the Japanese Patent Office (JPO) Examination Guidelines.

- Method;
- Computer readable storage medium;
- Computer program;
  - Module, library;
  - Neural network, support vector machine, trained model; and
- Information equivalent to a computer program;
  - Data structure, data having structures

Below are examples of ineligible subject matters:

- A law of nature as such;
- Mere discoveries and not creations;
- Technology in which a law of nature is not utilised; and
- A mere presentation of information

In Japan, the scope of an allowable subject matter, such as a program and its equivalents (data structure, etc.), is broader than in other jurisdictions. To protect such equivalents, specific recitations are required in the claims such that a cooperative process with hardware according to the purpose of use is realised.

According to the JPO, a trained model can be protected as a program at this moment. In research and



**Masato Iida**

Vice president

Shiga International Patent Office

T: +81 3 5288 5811

E: info@shigapatent.com

Masato Iida is vice president of Shiga International Patent Office. As a patent attorney and technical expert, he is qualified to represent parties in court in IP litigation. He specialises in biochemistry, molecular biology, and pharmaceuticals.

Masato has experience constructing target screening systems and evaluation systems for anti-cancer agents (both cell-based and cell-free systems). He also has detailed knowledge of genome editing and gave a lecture on notes of caution for patent maps and strategies for patent applications on genome editing at a seminar organised by Nikkei Biotechnology & Business.

Masato is a graduate of the University of Tokyo. He has contributed articles to medical magazines based on his wide knowledge and experience in both biotechnology and the IP field.

development (R&D) and drug discovery using AIs, specific characteristics, patterns or rules are detected by AI from massive data stocked in research. If there may be certain patterns which could leverage drug discovery, the relevant algorithm creates and predicts the most favourable research results.

With regard to an invention of a pharmaceutical product which is presumed to have a certain characteristic by using a learned model, at least a description of production and evaluation of an actual product in the specification is required. Otherwise, the application does not satisfy the enablement requirements even if a certificate of experimental results is submitted later.

### Case examples of AI-related pharmaceutical applications

There are several patent applications in which AI has been used in drug discovery

and predicting research results. These have been sorted out in this article with keywords: drug discovery, medical compounds, machine learning, neural networks, and other relevant technical terms. Below are three case examples of AI-based pharmaceutical applications.

#### Case one

**Title of invention:** Cytotoxic T-cell, cytotoxic T-cell inducer, and pharmaceutical composition and vaccine using thereof

**Japanese patent registration number:** 5633825

**Patentees:** NEC Corporation, Kochi University, and Ehime University

**Filing date:** August 6 2012

**Registration date:** October 24 2014

#### Claim 1

A cytotoxic T-cell inducer comprising a peptide,

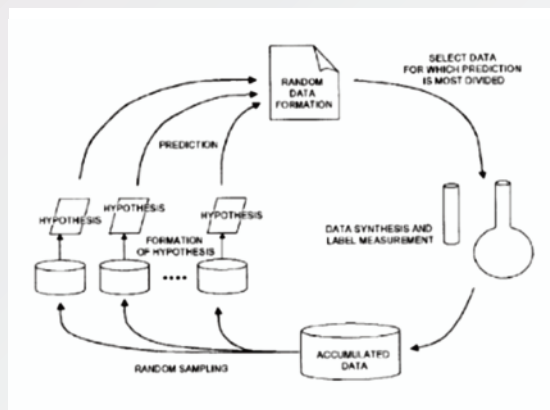
wherein the peptide is represented by an amino acid sequence consisting of nine amino acid residues of SEQ ID NO: 5, and binds to an HLA-A24 type class-I molecule on the surface of a cell that is a target of a cytotoxic T-cell.

**Claim 2**

A pharmaceutical composition for treating a disease that is caused by hepatitis C virus, the composition comprising a peptide, wherein the peptide is represented by an amino acid sequence consisting of nine amino acid residues of SEQ ID NO: 5, and binds to an HLA-A24 type class-I molecule on the surface of a cell that is a target of a cytotoxic T-cell.

**Claim 3**

A vaccine for use in preventing or treating a disease that is caused by hepatitis C virus, the vaccine comprising a peptide, wherein the peptide is represented by an amino acid sequence consisting of nine amino acid residues of SEQ ID NO: 5, and binds to an HLA-A24 type class-I molecule on the surface of a cell that is a target of a cytotoxic T-cell.



**Figure 1: Random data formation**

In this invention, new amino acid sequences were determined using the trained model mentioned in the examples in the specification. The experimental results are predicted based upon hypotheses generated by an algorithm and determined amino acid sequences. Example 1 describes a method to identify the experimental data results as follows:

**Example 1**

[0055]

... Specifically, procedures of prediction, experiment, and evaluation in the present Examples were carried out based on an active learning experiment design described in WO2006/004182, and in general the following steps were repeated, thus creating rules.

[0056]

(1) A trial of a lower-order learning algorithm, which will be described later, was carried out once. That is, a plurality of hypotheses were generated by random sampling from accumulated data...

[0058]

According to such an active learning method, the number of examinations of the binding experiment for peptides consisting of nine amino acid residues, which would otherwise have to be carried out for 500,000,000,000 (=209) or more combination of all the candidates for HLA-binding peptides, could be reduced.

As shown above, although the scope of the claims do not describe a peptide found by an algorithm, the examples describe, in detail, how precise data for the new drug is generated with the active learning model. Furthermore, this application was jointly filed by a Japanese IT giant and universities, and exemplifies a quintessential case of corporate-academic technical collaboration surpassing beyond technical boundaries.

**Case two**

**Title of invention:** Generative machine learning systems for drug design

**Japanese patent registration number:** 6866375

**Patentee:** Preferred Networks Inc

**Filing date:** December 2 2016

**Registration date:** April 9 2021

**Claim 1**

Computer system comprising a decoder that is realised as a neural network and generates information about compounds from potential representations; wherein the decoder is trained by decoding a potential representation generated by encoding information and labels about a compound with an encoder realised as a neural network and generating a reconstruction of the information about the compound; and wherein training of the decoder is constrained by reconstruction error.

**Claim 3**

The computer system of claim 1 or 2, wherein the training labels comprise one or more label elements having predetermined values.

**Claim 9**

The computer system of claim 3, wherein the label comprises one or more label elements selected from the group consisting of bioassay results, toxicity, cross-reactivity, pharmacokinetics, pharmacodynamics, bioavailability, and solubility.

**Claim 12**

A training method in which the computer is running, wherein in the training method, an encoder realised as a neural network is configured so that a decoder realised as a neural network decodes a latent representation



encoded based on both information about a compound and a label, and wherein the training is constrained by reconstruction error.

This invention is related to generative machine learning systems for drug design. In prior technologies, exploration of lead compounds used to be slow, costly, and less effective when conducting high throughput or virtual screening.

This invention is a method for drug design that directly generates candidate compounds comprising a group of predetermined attributes by encoding information about chemical compounds in neural networks. Training run by computers can predict candidate chemical compounds. The aforementioned claims clearly describe that the method is for predicting drugs utilising AI.

### Case three

**Title of invention:** Effective clustering of immunological entities

**Japanese patent registration number:** 6500144

**Patentee:** KOTAI Biotechnologies Inc

**Filing date:** March 28 2018

**Registration date:** March 22 2019

#### Claim 1

A method of analysing a collection of immunological entities, comprising:

- A step providing feature of at least two immunological entities, wherein the step eliminates calculating feature amounts from a three-dimensional model of at least two immunological entities;
- A step making a computer machine learning the analysis of antigen specificity or the binding mode of the immunological entities without specifying the antigen specificity or the binding mode based on the features; and
- A step classifying or determining the difference of the antigen specificity or the binding mode.

#### Claim 3

A method of analysing a collection of immunological entities, the method comprising:

- A step extracting features for each of at least one pair of sequences of members of the set of immunological entities, wherein the step eliminates calculating feature amounts from a three-dimensional model of at least two immunological entities;
- A step projecting the features into a high-dimensional vector space, wherein the distance of the members in space reflects the functional similarity of the members;
- A step clustering the set of immunological entities based on the distance; and

- A step analysing based on classification by the clustering as necessary.

#### Claim 9

The method of claim 3, wherein the feature includes at least one selected from sequence information, CDR1-3 sequence length, sequence coincidence, framework region sequence coincidence, total charge of molecules / hydrophilic / hydrophobic / number of aromatic amino acids, numbers of CDRs, framework region charge / hydrophilicity / hydrophobicity / number of aromatic amino acids, number of amino acids, heavy chain-light chain combination, number of somatic cell mutations, positions of mutations, presence / matching of amino acid motif, rarity relative to a reference sequence set, and an odds ratio of bound HLA by reference sequence.

#### Claim 10

The method according to claim 3 or 9, wherein the immunological entity is a cell comprising any one or more of an antibody, an antigen-binding fragment of an antibody, a B cell receptor, a fragment of a B cell receptor, a T cell receptor, a fragment of a T cell receptor, and a chimeric antigen receptor (CAR).

This invention provides a method for classifying immunological entities based on epitopes, and is related to generating epitope clusters and their application. The claims cover the process of clustering conducted by AI.

## Summary

It is assumed that digital transformation of drug development will further accelerate in the years to come. As mentioned above, at present, AI may serve as a tool assisting in drug discovery to predict a combination of peptides that could become a medical compound.

Medical compounds identified and generated by AI are rarely specified in patent applications, but the fact that experimental data is led by AI is described in examples. Under Japanese patent practice, because AI is not eligible as an inventorship, most of the titles of inventions tend to be 'computer system for drug discovery for...'. Even still, it could be said that some applications do not dare describe in the examples that AI discovers medical compounds.

The description of experimental data is a mandatory requirement when filing bio/pharma applications. Hypothetically, if the scope of the claims covering a medical compound of a certain peptide sequence is granted, does it mean that an applicant no longer needs to submit experimental data? Would enablement requirements supported by AI be admitted in the future? Or will process patents of a certain medicine derived from AI be granted? So many questions have come up – an eye has to be kept on the trends surrounding AI-related medical inventions.