Material 2

Provisional Translation

# Direction of Provision Method of Data of Laws and Enhancement of Utilization (proposal)

November 9, 2022

The Seventh Meeting of the Working Team on Digitalization of Legislative Affairs

Digital Agency

# Progress of the meetings

## Special Commission on Digital Administrative Reform

Chairperson: Prime Minister
Vice chairpersons: Minister for Digital Transformation and Chief Cabinet Secretary

- 2021.11.16: First meeting
- 2021.12.22: Second meeting → Formulation of "Digital Principles for Structural Reforms"
  → "Priority Policy Program for Realizing Digital Society" (cabinet decision in 2021.12)

## Working Group

Chairperson: Senior Vice-Minister for Digital Transformation

- 2022.2.10: First meeting

## Working Team on Digitalization of Legislative Affairs

Chairperson: Senior Vice-Minister for Digital Transformation

- 2.17: First meeting
  - Matters to be deliberated, etc.
- 2.28: Second meeting
  - Existing systems of legislative affairs, etc.
  - Conformity check process/system
  - Draft items for research on other countries
- 3.16: Third meeting
  - Conformity check process/system
  - Digitalization of legislative affairs; BPR; sharing of roles between the public and the private sectors

Points in question, etc. → 3.23 Seventh meeting → 3.30 Third meeting

- 4.13 Fourth meeting
  - Timeline for digitalization of legislative affairs
  - Progress of research on other countries
- 5.13 Fifth meeting
  - Progress of research on other countries
  - Timeline for digitalization of legislative affairs
  - Timeline for conformity check process

Timeline, etc. → 5.20 Eleventh meeting → 6.3 Fourth meeting

- 9.12 Sixth meeting
  - Preliminary draft of requirements, data structure, etc.
  - Trial conformity check, etc.
- 11.9 Seventh meeting (held this time)

→ Formulation of "Plan for Overall Review of Regulations in the light of the Digital Principles"

### Agenda of the Working Team on Digitalization of Legislative Affairs

**(1) Process/system to check the conformity of new laws to the Digital Principles**

**(2) Digitalization of legislative affairs, BPR, and ideal sharing of roles between the private and public sectors**

- Development and delivery of a base registry of data of laws (digital authenticated copy)
- Sharing of roles between the public and private sectors for the use of legal documents, etc.
- Utilization of Legal Tech in legislative affairs

# Purpose of legislative affairs system

(Reposted from Material 1 for the sixth meeting of the Working Team held on September 12, 2022)

- **One of the significant challenges toward structural reforms for attaining a digital society is: Legislative affairs themselves do not conform to Digital Principles.**
- **We need to design a legislative affairs system to address this challenge with the help of digitalization, ensure fundamental information for attaining a conformity check cycle, promote the strategic provision of open data of laws, etc.**

## Current challenges

### Goals for attainment of a digital society

Attainment of a society where the government gets closer to citizens and local communities, and individuals and businesses can show their abilities to the maximum extent possible

⇒ The "analog" structures remain due to regulations such as laws. It is essential to realize economic growth by reviewing these "analog" structures so that the power of digitalization is fulfilled to the maximum extent possible.

### One of the challenges toward structural reforms: Legislative affairs

■ **Enormous volume of manual work and large-scale mobilization of the workforce:** Inspection of existing laws; check of the conformity of new laws to the Digital Principles; review of existing laws adapted to the technological evolution

■ **Unclear big picture of "rules":** It is difficult to capture the big picture because laws are not society's only rules or norms.

■ **There is no authenticated digital data of "rules":** There are no legal grounds as to how to promulgate a digital authenticated copy of laws and its timing.

## Purpose of legislative affairs system

### Fundamental information to attain conformity check cycles

To maintain reliable data of laws that can be utilized efficiently as preconditions for checking the conformity to the Digital Principles and for continuously and efficiently reviewing and improving laws, etc.

### Strategic provision of open data of laws, etc.

To provide machine-readable, timely, and verifiable open data of laws, etc., to promote the wide distribution of proper understanding of legal texts, ensure sound development of industries, and appropriately protect the interests of all people.

### Conformity of legislative affairs to the Digital Principles

To streamline the processing of legislative affairs to the maximum extent possible and offer attractive usability as a part of the reform toward a sound and sustainable workflow which can ensure the reliability of laws, and control risks such as errors

## Expected effects

### Whole society

Current state: Lack of fundamental legal information for structural reforms

Promotion of structural reforms for attaining a digital society by ensuring fundamental information on a conformity check cycle

### Users of laws, etc.

Current state: Difficulty in immediately obtaining the latest articles

It becomes possible to timely obtain accurate data of laws, and more advanced services are created through Legal Tech

### Government

Current state: Enormous volume of stressful manual work

Efficient, reliable, and user-friendly legislative affairs which allow the staff to concentrate on the content of texts they ought to work on

# Requirements for legislative affairs system (working draft)

(Marked parts were changed based on Material 1 for the sixth meeting of the Working Team held on September 12, 2022)

- **Requirements for attaining the purpose are sorted out to be used as the foundation for architecture design and PoC plan.**
- **Not only technological feasibility and performance goals but also usability and consensus-building process are taken into account at all times.**

## A. Requirements for data structure

1. Machine-readable
2. Compatibility of data formats with international standards and existing systems
3. ★ (Timely) immediate provision after the promulgation
4. ★ Including past versions of documents
5. ★ Provision of reliable reference texts prior to enforcement
6. ★ Availability of consolidated texts as of a specific point in time
7. Capability of editing not amendment texts but consolidated texts
8. Coexistence with amendment texts for the time being
9. Capability with complex enforcement patterns caused by amendment of unenforced laws by interrupting enforcement and uncertainty in the enforcement dates due to cabinet order on the enforcement date
10. Keeping the consistency of not only specific points of time but also the whole time series (validity of amendment texts, etc.)
11. Support for version management during modification by law editors (a different axis from the time series)
12. ★ Supports laws that are included in official gazettes (that is, public notices and higher)

★: To be explained in detail in the following pages

## B. Requirements for workflow & API

1. Distributed information management among departments at the time of deliberating draft laws
2. Smooth data sharing and collaborative editing
3. Expecting changes in participating players and procedures, such as examination by the Cabinet Legislation Bureau and talks among different ministries
4. Efficient workflow with the scope of responsibility/interests of each player in mind
5. Flow in anticipation of busy season and emergency response
6. Mechanical collaboration with official gazettes and websites
7. ★ Provision of data required for external services with useful API and identifiers
8. Timely and appropriate involvement of assist tools such as error checks and conformity checks (CI/CD)
9. Protocols/flows highly compatible with widely-used management tools
10. Design allowing flexible development and extension by internally utilizing API connections
11. Layers can be shared for works that have some common workflows with legislative affairs (accessible with the same API and work in a seamless flow)
12. Usability for similar workflows other than ministries and agencies

## C. Requirements for UI

1. Usability equivalent to popular services provided in the private sector to promote the use of systems
2. Update to prevent obsolescence following services provided in the private sector and technological advancement
3. The layout work is separated in the article editor so that users can concentrate on the deliberation of contents

## D. Other architectural requirements

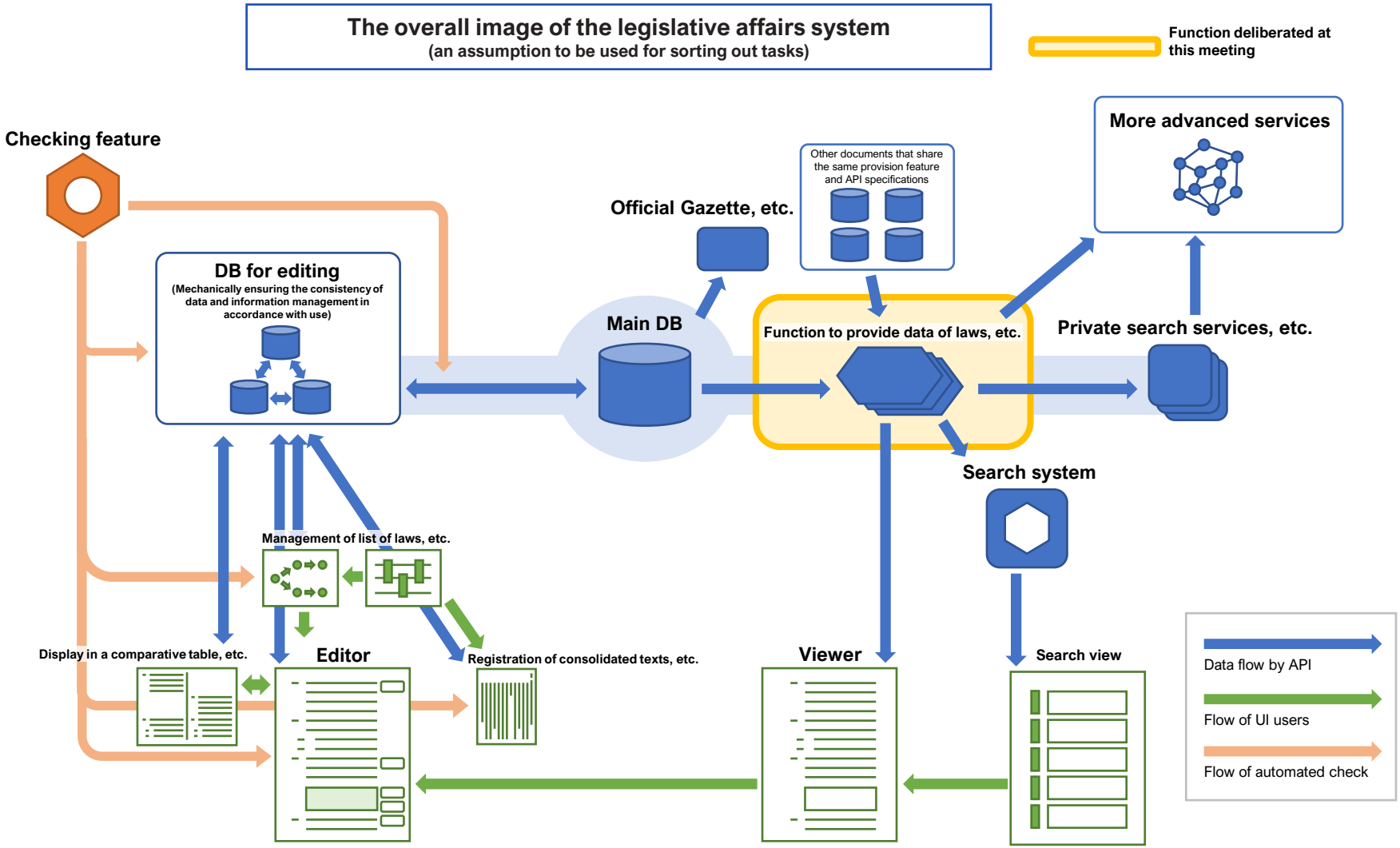
1. Utilization of existing systems (such as e-LAWS and legal examination support system), existing PoC, and existing product assets
2. Design should be automated to the extent possible, taking into account parts that need to be checked or modified manually
3. Capability of quantification and tracking of effects
4. Smooth migration from existing flows

# Provision method of data of laws, etc.

Function deliberated at this meeting:

# Provisioning function of data of laws, etc.

- A method of provision of data of laws, etc., and principles for the design of API to be provided are proposed at this meeting.



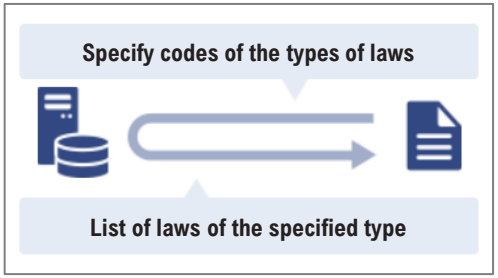
# Outline of provision of data of laws using current API

- **The current API for laws only provides current provisions of laws (ministerial ordinances and higher). In addition, laws that have not been enforced but promulgated are provided on e-Gov Law Search in anticipation of a particular enforcement order.**
- **We will deliberate the API development to include past data and other documents, such as public notices.**

**Features of current API for laws**

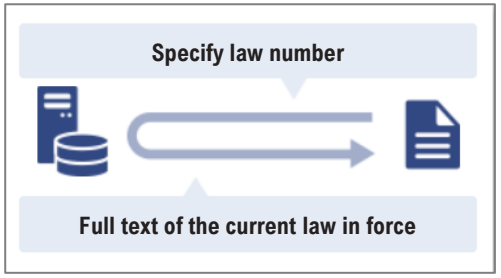
(Explanation of API for laws: https://elaws.e-gov.go.jp/apitop/ )

### ● API for retrieval of a list of laws



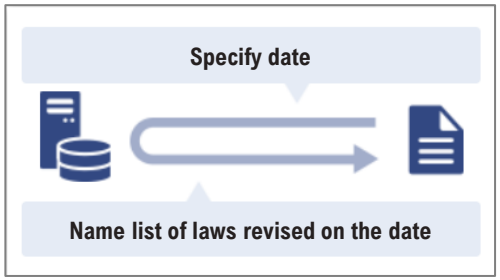

Retrieve law ID, names, numbers, and dates of promulgation of promulgated laws

### ● API for retrieval of laws



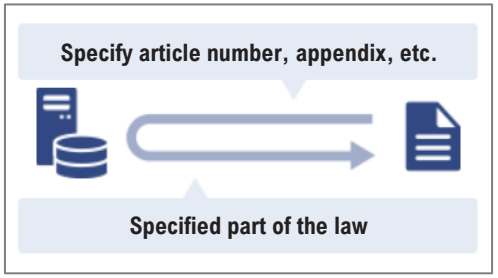

Retrieve full text of current law in force (including images in figures)

### ● API for retrieval of a list of revised laws




Retrieve a list of laws revised on a specified date

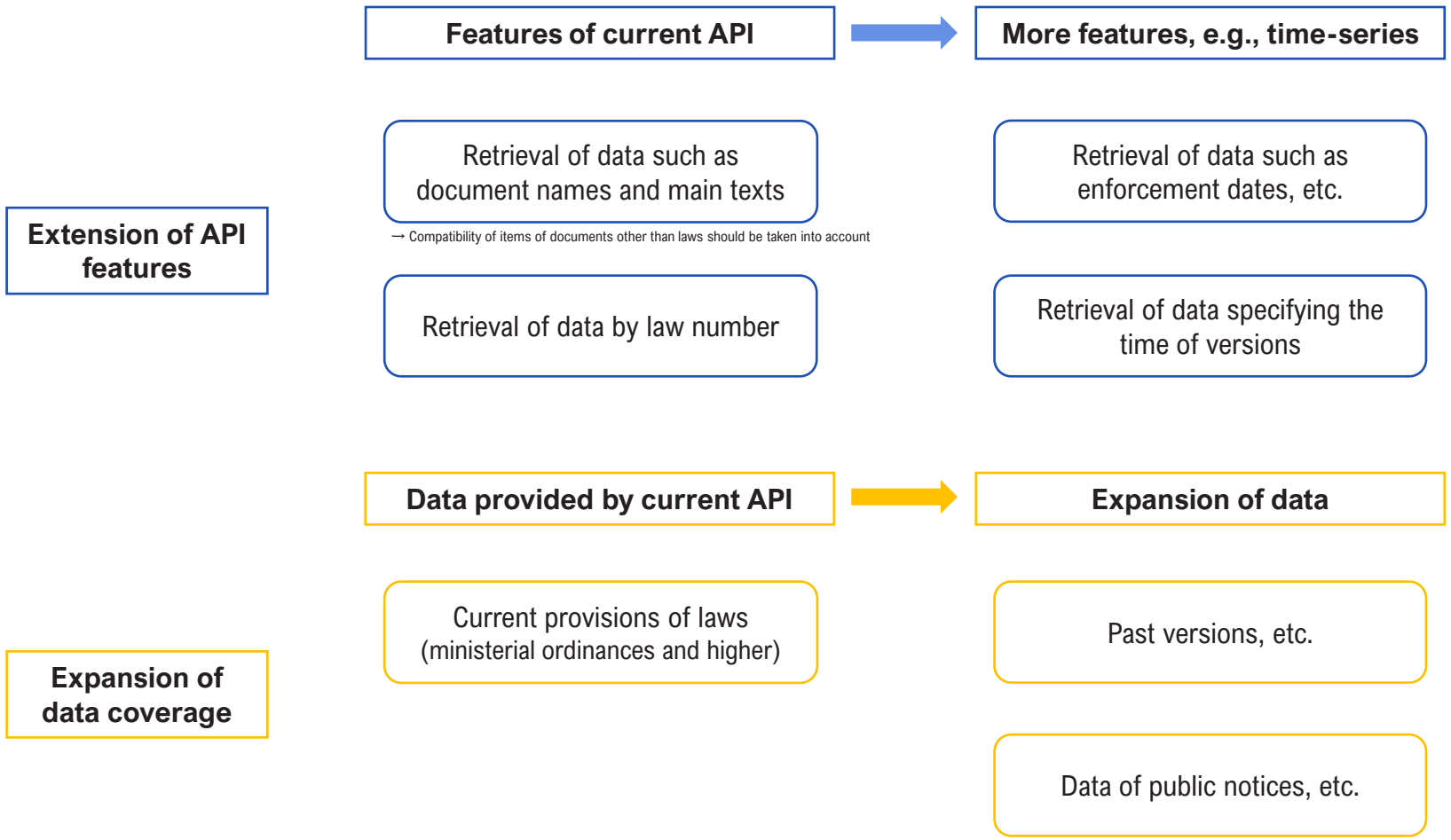
### ● API for retrieval of articles




Retrieve the content of current laws that match specified conditions (law number or ID and article, paragraph, or appendix)

# Image of extended API to provide laws (preliminary draft)

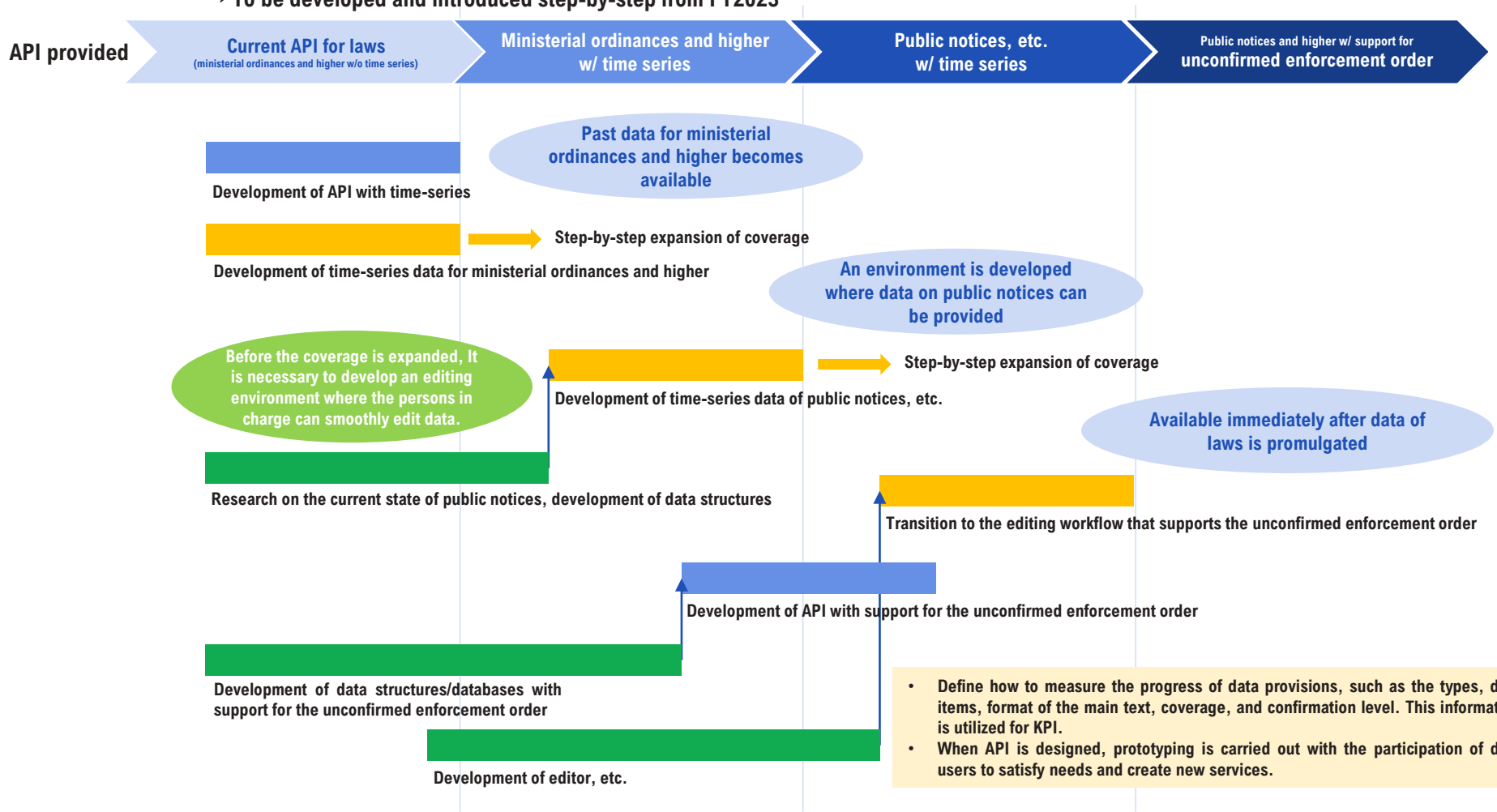
- **When the extension of API to provide laws is to be deliberated, compatibility and scalability should be taken into account so that the extension of features of API and data preparation advance step-by-step.**
- **When the data coverage is expanded, we continue to deliberate how we deal with a degree of confirmation of reliability, etc. and what kind of workflow is used for enhancement.**



# Mid-term Roadmap for API to provide laws (preliminary draft)

- Since data structures have already been defined for ministerial ordinances and higher, it might be easier to accelerate the development of API for time-series data of ministerial ordinances and higher.
- Enhanced data is provided step-by-step so that services are created in the private sector, and data is utilized in the government sector from the early stage.

→ To be developed and introduced step-by-step from FY2023




- Define how to measure the progress of data provisions, such as the types, data items, format of the main text, coverage, and confirmation level. This information is utilized for KPI.
- When API is designed, prototyping is carried out with the participation of data users to satisfy needs and create new services.

# Direction and future image of enhancement of utilization of data of laws

# Image of the utilization of data of laws

- **Since they lay a foundation for future enhancement, it is required to deliberate mid-term features for digitalizing legislative affairs so that it addresses future needs.**
- **The more effective use of the data of laws is expected to lead to the development of various value-added services and the utilization to enhance fair and efficient administrative management. Therefore, needs for use will be investigated toward further development of data and future feature enhancement.**
- **To satisfy the needs, since it is challenging to handle information such as meanings of legal terms with current technologies, it is expected that there be items for which technological development needs to be promoted.**

## Examples of expected needs for utilization

- **A service that allows users to timely update information by receiving an alert when any laws are modified or to be enforced through specifying IDs to refer to clauses from commentaries or written contracts.**
  - **Required data: Immediate, up-to-date data of laws, clause ID, updated information**
- **A service to predict laws that have a higher need for revision in the next years and trends that might affect own business, in association with past news and tweets, the progress of talks on regulations, and data on histories of revisions**
  - **Required data: Past data of laws, minutes of the Diet, various reports, commentaries, public comment, and information on legal terms/related words**
- **A service to research clauses that are more likely to have an impact on regulatory reforms and departments in charge and stakeholders, taking into account the relationship of delegation between laws and keywords of different fields**
  - **Required data: Clause ID, cross-reference information, information on departments in charge and stakeholders by keyword based on public information, information on legal terms/related words**
- **A service to verify the compatibility of products and terms and conditions by comparing corresponding clauses on trade, intellectual property, and technical regulations in different countries in a list so that they can be traced on a continuous basis.**
  - **Required data: Clause ID, data of laws, and clause ID of different countries**
- **A service to search related legal documents and websites based on keywords when creating new services or conducting research in less-familiar fields, taking into account the relationship of delegation and to deliberate about an initial legal reaction.**
  - **Required data: Cross-reference information, information on legal terms/related words**
- **A service to smoothly link, classify and search laws and related information by describing not only clause names but also clause IDs in public documents such as reports and sentences.**
  - **Required data: Clause ID, public documents with clause ID, etc.**

# Researches and cases of advanced policy planning utilizing data of laws

- **Research on regulatory reforms and policy planning utilizing advanced data of laws beyond the development of the base registry have been conducted in Japan and other countries, and some of their results have been put into practice. There are concepts in these deliberations, such as "Rules as Code", an effort to efficiently implement systems by expressing laws in a machine-executable format.**
- **Moreover, research on Legal Tech applying machine learning technology has progressed in recent years. Their results are expected to be utilized for advanced technological development, the creation of services, and efficient administrative management thanks to the accumulation of the data of laws.**
- **In order to realize advanced policy planning, etc., utilizing data of laws, it is required to establish reliable data of laws, develop a technology and a data infrastructure to analyze relationships between documents related to laws and meanings of legal terms with machines.**

## "Better Rules" (Government of New Zealand, 2018)

https://www.digital.govt.nz/dmsdocument/95/html

- **The intention of laws that were prepared in a non-digital format may be detached from implementation in our time when administrative services are provided digitally.**
- **A methodology was proposed to raise the effect of systems so that laws can be easily executed and simulated with digital technologies by describing laws in a machine-executable format. Challenges are sorted out by a service design method.**
- **The term "Rules as Code" is pointed out to be invented by Better Rules Team (OECD, 2020, described on the right-hand side)**

## "Rules as code: Seven levels of digitisation" (Wong, 2020)

https://ink.library.smu.edu.sg/sol_research/3093/

- **Advancement levels of digital tools such as rule engines, ontology, and automated tools are classified in anticipation of laws whose conditional branch is clear.**
- **This concept aims to cite levels when tools are utilized. For example, if someone proposes to "implement a level 3.0 Rules as Code prototype", the system can respond, "O.K., here are the required time, resources, and procedures".**

## "Cracking the Code - Rulemaking for humans and machines" (OECD, 2020)

https://doi.org/10.1787/3afe6ba5-en

- **In addition to laws written in conventional natural languages, machine-executable codes are created as "Rules as Code" when formulating laws. In this way, they are edited simultaneously so that codes can be utilized directly for system operations.**
- **Willemstein & Ross (2021)* pointed out that this methodology is a test-driven development in system design.**

* https://www.brcommunity.com/articles.php?id=c059

## "OpenFisca" (Government of France, 2011~)

https://openfisca.org/en/

- **A system mainly focuses on taxes and benefits to calculate the amount of tax by entering systems as machine executable codes.**

## "DataLex" (AustLII (Australia))

https://datalex.org/

- **The original project started around 1987*. This system can write machine-executable codes using a rule description language similar to English and display applicable parts in response to questions.**

* https://dl.acm.org/doi/10.1145/41735.41737

## "Xelen" (Kakuta, 2016~)

https://xelen.jp/secom_pj/

- **The relationship of rights and obligations and their transfer are simulated based on rules described in Python. It is equipped with a law code editor and a conversion feature into Japanese clauses.**

# Digital Legislation Roadmap (proposal for discussions)

- We propose a long-term digital legislation roadmap aiming at realizing advanced private businesses based on advanced technology such as Legal Tech, flexible and effective policy design based on digital technologies and data of laws, and fair and efficient administrative operations. It aims at the continuous implementation of advanced policy/service design by stepwise progression cycles of developing technologies, establishing data infrastructure utilizing the technologies, and policy implementation.

(Equivalent to "Public notices and higher w/ support for unconfirmed enforcement order" in the mid-term roadmap)

## Phase 0 (as-is)

- Lack of reliable open data of laws, inefficient manual investigation of clauses

→ Realization of timely provision of reliable and machine-readable data of laws (incl. acts and ordinances)

Dev Use

## Phase 1

Keyword:
### Law base registry

- Reliable and machine-readable data of laws is available and continuously updated

→ Cross-reference/coordination of data in public and private sectors will progress, and services will emerge

Dev Use

## Phase 2

Keyword:
### Connected data

- Delegation and references are compiled as data, and documents are coordinated through the unified API

→ Analysis of data, including peripheral information on laws, and interpretation of terms, will progress

Dev Use

Together with the creation of services and utilization in legislative affairs using technologies developed, data analysis is promoted to achieve the next phase.

## Phase 3

Keyword:
### Law ontology

- Basic semantic information is developed, such as meanings of legal terms and logical relationships

→ Analysis of the application and structure of laws will progress. "Analog" regulations will be identified automatically

Dev Use

## Phase 4

Keyword:
### Law static analysis

- Logical structures of laws are available as data. Relationships between the structure and operations of laws are compiled

→ Inconsistency and redundancy of laws will be automatically identified, and optimal digital operations will be suggested

Dev Use

## Phase 5

Keyword:
### System digital twin

- Laws are simulated in virtual spaces, including meanings

→ Effects of laws are analyzed automatically, enabling fair, effective, and flexible policy design with feedback of actual changes in real world

Dev Use

Rather than trying to progress the phases for all fields at once, the progression is carried out on a trial basis from areas where it is easier to implement (e.g., rules on numbers).

# Digital Legislation Roadmap (details): Phases 0 & 1

- **The legislative affairs system to be designed aims to reach Phase 1 to create services and research which have been difficult to be attained in the past and to promote the infrastructural development toward the next phase.**

## Phase 0
(as-is)

**Available technology and data infrastructure**

- The latest texts of laws (ministerial ordinances and higher), which are reliable to a certain extent, are provided in XML format after a short time lag from their promulgation.

**Services and research to be attained**

- The reliable data of the latest laws is not always available. There was no choice but to check articles with large-scale workforce mobilization when regulations were reviewed at the Special Commission on Digital Administrative Reform.
- The legal documents lower than public notices, which govern a large part of operations, and past versions of laws required for analyzing rules at a particular time of the past are not available as machine-readable data.

**Efforts for the next phase**

- Provision of reliable law base registry (public notices and higher) on a timely basis and collection of past versions.
- Design of machine-readable data format and API easier for secondary use and cross-reference.
- Establishment of environments for editing and provision, organizations for management and development, and systems to achieve the above goals.

## Phase 1

Keyword:
**Law base registry**

(Equivalent to "Public notices and higher considering unconfirmed enforcement order" in the mid-term roadmap)

**Available technology and data infrastructure**

- The latest texts of laws (public notices and higher) are provided in XML format and on a website. Some past data becomes available. Cross-reference is possible with ID, which refers to any part of the documents.

**Services and research to be attained**

- It becomes possible to analyze dependencies of legal documents using reliable data of laws (public notices and higher).
- Services and data analysis related to legislative updates and events at a particular time will be realized.
- Documents other than legal documents are cross-referenced to accumulate connected data.

**Efforts for the next phase**

- Development of methods and technologies to handle related documents other than laws as connected data, organizations for management and development of data, and systems and infrastructure.
- Development of reliable databases of dependencies and delegation of laws.
- Accumulation of machine-readable data of related documents other than laws (notifications, public comments, commentaries, etc.)

# Digital Legislation Roadmap (details): Phases 2 & 3

- **Phase 2 is attained by utilizing accumulated data, R&D results, and infrastructure established in Phase 1. Subsequently, phase 3 is achieved by promoting more advanced data accumulation and R&D, and semantic infrastructures of legal terms will be established.**

## Phase 2

Keyword:
**Connected data**

### Available technology and data infrastructure

- Machine-readable data of related documents other than laws (notifications, public comments, commentaries, etc.) can be obtained from unified APIs. Reliable connected data on dependencies and delegation of laws is established.

### Services and research to be attained

- At least at the paragraph level, the affected extent of changes made (e.g., propagation of editorial changes such as renumbering) becomes automatically evident, so it will no longer be necessary to investigate such changes manually.
- Services for analyzing trends in legislation and related institutions using keywords and associated fields will be available.
- It will be possible to comprehensively analyze the background of the formulation of articles and related documents, including subsequent interpretations, so that essential information for semantic analysis will be accumulated.

### Efforts for the next phase

- Development of methods and technologies to analyze laws based on natural and formal languages and understand legal vocabulary and fundamental logical structures.
- Analysis and data accumulation of the background of legal terms and their interpretations, logical structures, and contexts of laws.
- Development of the validity of data of legal terms and related data in law interpretation and establishment of organizations and systems for management and development of such data.

## Phase 3

Keyword:
**Law ontology**

### Available technology and data infrastructure

- Reliable databases of building blocks of laws are established, such as legal terms, their meanings, and fundamental logical structures.

### Services and research to be attained

- It will be possible to apply mutatis mutandis with implicit word replacements at a fair precision, providing a foundation for semantic structure analysis of law application. Automatic generation of customized commentaries will also be possible.
- It will be possible to analyze laws based on abstract correspondence beyond literal keywords, enabling automatic detection of existence and application of regulations to a certain extent. E.g., “analog” regulations are identified almost automatically at this point.
- It will be possible to classify and compare laws with similar structures, providing a foundation for analysis of the implementation and operation of systems.

### Efforts for the next phase

- Development of methods and technologies to express and analyze logical structures of application of laws in a machine-readable way; and accumulation of data.
- Development of the validity of logical structures of application of laws in law interpretation and establishment of organizations and systems for management and development of such data.
- Development of methods and technologies to analyze the relationship between structures of laws and implementation and operation of systems, and their efficiency from the administration and technology sides, considering the background of enactment, amendment, and implementation; and accumulation of data.

# Digital Legislation Roadmap (details): Phases 4 & 5

- **When Phase 4 is attained, system improvements using digital technologies are brought to realization. Moreover, in phase 5, simulations using a digital twin are achieved, realizing effective, fair, and flexible system implementation, even under highly uncertain situations.**

## Phase 4

Keyword:
**Law static analysis**

### Available technology and data infrastructure

- Technologies for analyzing logical structures of the application of laws and for converting them into a machine-readable format are established. Moreover, reliable databases on the relationship between structures of laws and the implementation and operation of systems are developed.

### Services and research to be attained

- It becomes possible to simulate static structures of laws. Inconsistency and redundancy in regulations are automatically identified, and draft modifications to convert them into cleaner provisions can be presented.
- An optimal operational flow and technologies can be automatically suggested according to institutional or regulatory patterns. Draft improvements of “analog” regulations, including operations, can be presented.

### Efforts for the next phase

- Development and implementation of models and processing systems that expresses the meanings of legal terms and logical structures in virtual spaces.
- Development and implementation of theories and technologies that enable simulations of laws’ meanings, effects, and changes and economic activities in virtual spaces; and establishment of organizations and systems for management and development of such technologies.
- Research on the relationship between machine-executable codes and laws and regulations; implementation of related systems.

## Phase 5

Keyword:
**System digital twin**

### Available technology and data infrastructure

- laws and their meanings are described in a machine-executable format. Moreover, technologies for simulating laws in virtual spaces are established.

### Services and research to be attained

- It becomes possible to simulate dynamic structures of laws. The impact of regulations can be analyzed automatically to design optimal policies through social experiments in virtual spaces.
- Effective, fair, and flexible system implementation is realized even in uncertain situations by constantly reviewing systems with feedback on changes in society and the nature and effects of system changes, to simulations in virtual spaces.

### Efforts for the next phase

- Data on the performance of digital legislation and trends of changes in society and nature where laws are applied will accumulate, and its analysis will advance. As a result, even higher-level structural changes are expected.