

5問 データベースの利用者における利用の方法や指定法人が整備するシステムの具体的な運用イメージについて、法務当局に問う。

[利用の方法について]

- 本制度においては、民事・行政事件の判決書等が広く指定法人のデータベースに収録され、指定法人から民事裁判情報の提供を受けた一次的な利用者が、様々な価値を付加して製品やサービスを開発・提供し、それが二次的な利用者に提供されて活用されることを想定している。

- 具体的には、例えば、判例データベース事業者などの一次利用者において、民事裁判情報につき、
 - ・ 裁判例の体系化
 - ・ 解説や英訳の付与を行うことのほか、デジタル技術を活用し
 - ・ 裁判例の横断的分析
 - ・ より精緻な統計的分析
 - ・ 機械学習の素材にしてA Iの研究開発を行うこと等も考えられる。

[指定法人のシステム]

- 指定法人については公募を経て指定することから、その整備するシステムについて現時点で確定的なお答えをするのは困難であるが、有識者検討会(民事判決情報データベース化検討会)においては、
 - ・ A Iを用いるなどして機械的に仮名処理を行うためのシステム
 - ・ 最高裁判所から民事裁判情報を取得し、機械的な仮名処理

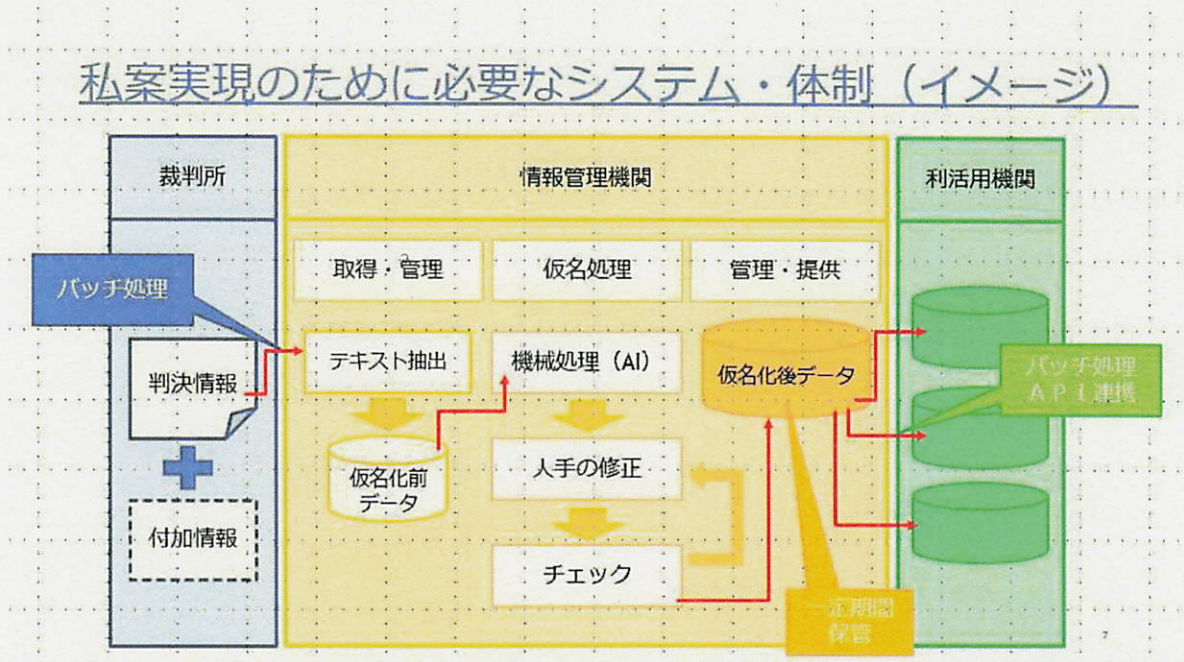
の正確性を確認し、確認後の情報を利用者に提供し、この間において情報を適切に管理するためのシステム
等が必要になるとされたところ。

- （先ほど述べたように）一次的な利用者においては、デジタル技術を活用した民事裁判情報の分析等が行われることが想定されることから、有識者検討会においては、指定法人が利用者に提供するデータの形式はC S VやXML等の機械判読に適したものとすべきであるとされており、こうした形式でデータを提供するためのシステムも必要になると考えられる。

（参考1）有識者検討会におけるヒアリング（民事判決情報データベース化検討会第5回〔令和5年2月22日実施〕議事録抜粋）

資料の6ページは「私案実現のために必要なシステム・体制」です。情報管理機関としては、主として次のような業務を行うということを予定しております。まず、裁判所から仮名処理前の民事判決情報を取得して、それを管理するということです。次に、取得した民事判決をAIによって仮名処理、そしてAIによって仮名処理したものを人手で修正・ダブルチェック、そして仮名処理した後の民事判決情報を利活用機関に提供、最後に、現在の紙の判決書やそれをテキスト化したデータにはない、付加的な情報についても従来の紙の判決書の情報に付加して裁判所から提供を受け、これについても利活用機関に提供するということを予定しているところです。この付加的な情報としましては、どのようなニーズがあるかについて、事業の在り方に関するワーキンググループに参加しているデータベース会社等にヒアリングをしたところ、判決のIDとか判決の言渡し日、裁判所のID、事件番号、審級関係、原審裁判所の判決のID、口頭弁論の終結の日、閲覧等の制限決定があった場合、更正決定があった場合等の情報について、付加的に提供いただきたいというニーズがあるということが分かっております。

(参考2) 民事判決情報データベース化検討会第5回資料 (抜粋)



(参考3) 機械判読に適した形式

C S VやXML等の形式を想定している。C S Vとは、Comma Separated Values の略称で、値や項目をカンマで区切ったテキストファイル・データのこと。Microsoft Excel 等様々なソフトで取り扱うことができる。XMLとは、Extensible Markup Language の略称であり、「タグ」と呼ばれるマークアップ記号を利用してテキストに情報を付加することができる (例えば、文中に「…〈主文〉1 被告は、原告に対し…。〈/主文〉」などと記号を挿入することで、記号間のテキストが主文であるという情報を付加することができる。)