

# The Possibilities of eCourts Data for Advancing Research on Law Implementation

Experiences of using Case Metadata to understand the Child and Adolescent Labour (Prohibition and Regulation) Act, 1986



# The Possibilities of eCourts Data for Advancing Research on Law Implementation

*Experiences of using Case Metadata to understand the Child and Adolescent Labour  
(Prohibition and Regulation) Act, 1986*

**by Enfold Proactive Health Trust**



*March 2025*

*Technical  
Assistance from*



*With  
Support from*



© Enfold Proactive Health Trust, 2025

Design and Cover Illustration by: *that studios*

Cite as: Enfold Proactive Health Trust, *The Possibilities of eCourts Data for Advancing Research on Law Implementation (2025)*

# Contents

<b>Acknowledgments</b>	<b>6</b>
<b>Abbreviations</b>	<b>7</b>
<b>Introduction</b>	<b>8</b>
<b>I. The eCourts Portal: Possibilities and Limitations</b>	<b>9</b>
<b>II. Areas of Reform in Maintenance of eCourts Data</b>	<b>14</b>
<b>III. Conclusion and Suggestions</b>	<b>22</b>

# Acknowledgments

This report has come to fruition thanks to the tireless efforts of many individuals working towards this common goal. Foremost, this exploratory study would not have been possible without the generous support of the Patrick J. McGovern Foundation through its *Data to Safeguard Human Rights Accelerator Program*.

We are grateful to the CivicDataLab team comprising Gaurav Godhwani - Executive Director & Co-founder of CivicDataLab and Deepthi Chand - Technology Director & Co-founder of CivicDataLab. We would also like to thank the data team from CDL comprising of Apoorv Anand, Manjunath Hegde and Sai Krishna for their able technical assistance in data curation of metadata for the analysis of the Child and Adolescent Labour (Prohibition and Regulation) Act, 1986.

The core research team at Enfold, comprising Shruthi Ramakrishnan, Shivangi Puri, Anindita Pattanayak, Gopika Nangia and Vibha Angre, were responsible for data cleaning, data analysis, data validation and data visualisation that formed the experiences on the basis of which this report was conceptualised.

Anindita Pattanayak and Shivangi Puri identified the types of gaps encountered in eCourts data and collected instances of these gaps which formed the basis of this report. The report was written by Anindita Pattanayak. Further restructuring, writing and comprehensive review of the draft was done by Swagata Raha.

We wish to acknowledge Mr. Ravi K. and Ms. Aishwarya T. from *that studios* for designing the report and bringing it to life with data visualisation.

Our gratitude extends to the administrative and finance team at Enfold comprising Geetha Nair, Bhavana D., Bharath H.M., and Amrutha M. for their unwavering support.

Invaluable insights and guidance provided by our external reviewers, Apoorv Anand (CivicDataLab), Sandhya P.R., Smita Mutt (DAKSH) and Urmi Chudgar Vohra, helped shape the final report.

# Abbreviations

AHTU	Anti-Human Trafficking Unit
BNS	Bharatiya Nyaya Sanhita, 2023
BNSS	Bharatiya Nagarik Suraksha Sanhita, 2023
CALPRA	Child and Adolescent Labour (Prohibition and Regulation) Act, 1986
CDL	CivicDataLab
CJM	Chief Judicial Magistrate
CNR	Case Number Record
CPOLA	Children (Pledging of Labour) Act, 1933
CrPC	Code of Criminal Procedure, 1973
FIR	First Information Report
FTC	Fast Track Court
IPC	Indian Penal Code, 1860
JJ Act, 2000	Juvenile Justice (Care and Protection of Children) Act, 2000
JJ Act, 2015	Juvenile Justice (Care and Protection of Children) Act, 2015
JMFC	Judicial Magistrate of the First Class
NCRB	National Crime Records Bureau
PJMF	Patrick J. McGovern Foundation
POCSO Act	Protection of Children from Sexual Offences Act, 2012

# Introduction

The [eCourts portal](#), managed by the [e-Committee under the Supreme Court of India](#), offers information about ongoing and disposed judicial proceedings. Apart from copies of judicial orders, the portal provides *metadata* such as provisions and legislations under which the case is filed, case type, judge designation, judicial forum, dates of registration, hearings and disposal, and pendency and disposal nature. Although the portal is designed to help lawyers and litigants find details of their cases online and monitor them, the platform has made it possible to view court data at an aggregate level, enabling analysis of the implementation of specific legislations in an unprecedented manner.

Data from eCourts has been used extensively to conduct studies on the functioning of courts, court management, pendency and judicial reform, sometimes in combination with other sources of data collection<sup>1</sup>. Some studies have relied on eCourts data to study specific legislations such as the Protection of Children from Sexual Offences Act, 2012 (POCSO Act)<sup>2</sup>.

In 2023-24, with the support of PJMF's *Data to Safeguard Human Rights Accelerator*, Enfold Proactive Health Trust undertook an exploratory study on the implementation of the Child and Adolescent Labour (Prohibition and Regulation) Act 1986 (the CALPRA) with technical support from CivicDataLab. Advanced data tools were used to extract district court data on the eCourts platform, the analysis of which culminated in the report [Decoding Data on Implementation of the Child and Adolescent Labour \(Prohibition and Regulation\) Act, 1986](#) (the CALPRA report). The CALPRA report involved analysis of the metadata of 10,800 cases from six States (metadata analysis) along with the detailed analysis of 142 select judgments from three States wherein the judgments were identified using the metadata (judgment analysis).

For metadata analysis, metadata was extracted from all cases under statutes with the terms “child” or “children” and “labor” or “labour” between 1 January 2015 and early 2023. As data was scraped over a two-month period, the end-date for data collection was different for each state. A total of 24,723 cases were retrieved from the entire country. For the purpose of an in-depth analysis of the metadata based on select variables, Assam, Bihar, Jharkhand, Maharashtra, Tamil Nadu and Uttar Pradesh were selected based on a consideration of the total number of the CALPRA cases registered as per data from eCourts, number of cases registered under the CALPRA as per *Crime in India*, incidence of child labour as per the *Census of India*, availability of judgments in English language and geographical diversity. This data was then cleaned to remove false positives, using information available from other fields (e.g., removing civil case types as proceedings under the CALPRA are criminal in nature), resulting in a dataset of 10,800 cases. The metadata offered a glimpse into the interplay between various legislations applicable in the context of child labour and district and State-wide trends in disposal and pendency, among others.

This report examines the potential of eCourts metadata in advancing analyses of specific legislations and describes gaps in data quality that persist, based on the experience of legal researchers and data analysts curating a dataset, cleaning and analysing it to track trends in cases under the CALPRA. It concludes with suggestions on improvements in eCourts data that can enable effective data analysis to understand the implementation of a legislation and interplay of laws.

1 Vidhi and DAKSH, '[Litigation Landscape of Bengaluru](#)' (July 2019); Devendra Damle and others, '[Characterising cheque dishonour cases in India: Causes for delays and policy implications](#)' (16 April 2022); Manaswini Rao, '[Front-line Courts As State Capacity: Evidence From India](#)' (12 October 2023); Manaswini Rao, '[Judicial Capacity Increases Firm Growth Through Credit Access: Evidence from Clogged Courts of India](#)' (2020); Elliott Ash and others, '[Measuring Gender and Religious Bias in the Indian Judiciary](#)' (2021) 3 Center for Law & Economics Working Paper Series; Manaswini Rao, '[Whither Justice?: Judicial Capacity Constraints Worsens Trial and Litigants' Outcomes](#)' (29 December 2020).

2 Vidhi, '[A Decade of POCSO: Developments, Challenges and Insights from Judicial Data](#)' (17 November 2022); HAQ Centre for Child Rights and CDL, '[Unpacking Judicial Data to Track Implementation of the POCSO Act in Assam, Delhi & Haryana](#)' (2021)

## I. The eCourts Portal: Possibilities and Limitations

### 1. The eCourts Platform and Metadata

The Supreme Court initiated the [National Policy and Action Plan for Implementation of Information and Communication Technology](#) in the Indian Judiciary in 2005 to digitise court records and systems (eCourts Project). The aim was to promote transparency and citizen centric justice delivery, specifically the generation and public availability of data “to enable evidence-based policy making”<sup>3</sup>. Case records of pending and disposed cases at the Supreme Court, High Court and district court level, have been digitised to some extent and made accessible to the public through the [eCourts](#) platform. The eCourts Project is now in Phase III, with a focus on moving towards digital, online and paperless courts.<sup>4</sup>

### 2. What is “metadata” from eCourts and how is it analysed?

For the purposes of this report, “metadata” signifies the fields and entries that are gathered from the landing page of the eCourts platform when the relevant case number is entered. “Fields” such as case-type, filing date and disposal type and their respective “entries” for each case, combinedly constitute the metadata (as shown in Image 1). While orders and judgments in the case that are available as links on the landing page are also sources of information about the case, this was not considered as “metadata” for the purpose of this report.

The metadata was collected using computer programming scripts, creating an aggregated dataset for analysis. This compiled dataset was then cleaned and analysed to glean insights as shown in Image 2.

<sup>3</sup> eCommittee, Supreme Court of India, [Digital Courts Vision & Roadmap e-Courts Project Phase III](#) (Vision & Roadmap, 2022) 34; See eCommittee, Supreme Court of India, [National Policy and Action Plan for Implementation of Information and Communication Technology in the Indian Judiciary](#) (1 August 2005)

<sup>4</sup> eCommittee 2022 (n 3); Ministry of Law and Justice, [‘Third Phase of eCourts’](#) Press Information Bureau (7 December 2023).



### 3. Significant Data Insights Obtainable from eCourts

On the basis of the metadata analysis of 10,800 the CALPRA case records, the following insights could be gleaned, which can be replicated in other studies pertaining to implementation of legislations or specific provisions:

- a. **Data on actual number of criminal trials under a legislation:** [Crime in India](#) published annually by the National Crime Records Bureau (NCRB) is the main source for the number of reported incidences of criminal offences in India. However, it follows the “Principal Offence Rule” as per which among many offences registered in a single FIR case, only the most heinous crime (maximum punishment) is considered as a counting unit.<sup>5</sup> On the basis of this rule, crimes that constitute minor offences such as under child labour laws when clubbed with serious offences are not reflected in the data on offences under laws regulating child labour. The extraction of the metadata from eCourts has made it possible to gauge the volume of cases of child labour in which the criminal law is set into motion.

*Data on the the CALPRA cases/incidences for Assam, Bihar, Jharkhand, Maharashtra, Tamil Nadu & Uttar Pradesh (between 2015 and 2022)*

NCRB, Crime in India	Metadata <sup>6</sup>	
1,329 incidences under the the CALPRA	9,193 - Criminal Trials in metadata which mention the the CALPRA.	7,792 Criminal Trials in metadata which mention <i>only</i> the the CALPRA and no other legislation.

- b. **Expanded dataset available for analysis:** Metadata also offers some advantages over analysis based solely on judgments, as it offers information about pending cases or cases with judgments in the vernacular. It substantially expands the data available for analysis and also enables district and State-wise comparisons and identification of geographies with a high number of prosecutions of certain types of offences.
- c. **Insights into the interplay of legislations:** Metadata allows for a detailed analysis of charges applicable in criminal cases to track how offences under different legislations are applied together, the choice of judicial forum, the duration of the trial and other such disaggregated granular analysis. For instance, in the context of the the CALPRA analysis, it was seen that provisions of the Indian Penal Code, 1860 (IPC) and the JJ Act, 2000 and JJ Act, 2015, were also applied. Further, the inclusion of some legislations along with the the CALPRA such as the Arms Act, the Explosives Act and the Explosive Substances Act, Jharkhand Minor Minerals Concession Rules, 2004 and The Bombay Prevention of Begging Act, 1959 raises concerns about the processes and occupations in which children and adolescents are being allegedly employed. Such data may also be used to understand region specificity that may then guide prevention and response-related efforts to address child and adolescent labour.
- d. **Identification of trends and practices that may require attention or re-examination:** Analysis based on metadata helps identify trends that require further research or exploration. Trends in disposal and pendency can help identify districts that warrant attention and plan interventions, such as the establishment of additional courts. A significant finding was the use of Lok Adalats for reaching settlements in the CALPRA cases, which was observed in 21.1% of disposed criminal trials.<sup>7</sup> While Lok Adalats may result in speedy disposal, it calls for an examination as to whether it is in the best interest of children and is legally permissible. Another finding was that convictions in the CALPRA offences varied widely between States. It was 9.8% in Maharashtra, whereas in Uttar Pradesh and Tamil Nadu it was 99.1% and 91.7%, respectively.<sup>8</sup> This wide disparity indicates the need for a deeper examination of the factors leading to convictions in these cases, such as the proportion of cases in which the accused pleaded guilty, and whether victims’ testimony was considered, among other factors. These trends also call for examination of data entry practices in different States to ascertain the accuracy and granular nature of the data.

<sup>5</sup> See National Crime Records Bureau, [Crime in India 2022](#), Vol. 1, v, “Limitations”.

<sup>6</sup> Note that in the metadata, it is possible that more than one case can arise out of a single incident reported if the criminal trial is transferred from one court to the other in some cases.

<sup>7</sup> Enfold Proactive Health Trust, [‘Decoding data on Implementation of the Child and Adolescent Labour \(Prohibition and Regulation\) Act, 1986’](#) (2024) 26.

<sup>8</sup> *Ibid.*

Overall, such insights can impact strategic interventions to promote the monitoring of the implementation of all legislations, by providing valuable evidence to inform advocacy, policy formulation and capacity-building training of institutional and civil society actors.

## What eCourts metadata can indicate about the implementation of a legislation

### *Extent and Type of Cases*



- Volume of cases filed under a specific legislation in a given time period
- Distribution of case load under this legislation across different States, districts and judge categories
- Types of cases: civil, criminal, trials, bail, anticipatory bail, etc.

### *Disaggregated Data on Specific Provisions and Interplay of Laws*



- Interaction between the selected legislation and other legislations applied with it
- Specific provisions being applied under different legislations

### *Forum of Adjudication*



- Courts dealing with cases under the legislation
- Where a special court is designated to deal with cases, whether the trial is being conducted by them
- Whether the case was decided in a Lok Adalat

### *Status of Disposal*



- Proportion of cases pending and age of pending cases
- Districts and courts with low, medium, and high pendency
- Stage-wise duration of time spent on a case

### *Duration of Cases*



- How soon after filing is a case heard or listed by the court for the first time?
- How frequently is a case being heard by the court? Does this frequency change with the type of court or the type of case?
- Which stage of the case has the maximum number of hearings?

### *Nature of Outcomes*



- Convictions, acquittals, transfer, committal, discharge and settlements
- Outcomes in bail cases
- Nature of civil relief granted

### *Cross-analysis of Different Data-points*



- How long does a disposal take based on outcomes, i.e., disposal time for acquittals vis-à-vis disposal time for convictions?
- How long does a criminal trial under the legislation typically take? Is the time taken different based on the provision applied?
- How long does a civil case under the legislation typically take? Is the time taken different based on the provision applied?
- Disposal time with respect to the forum of adjudication
- Duration between the year in which the FIR was lodged and the start of the criminal trial

#### 4. Limitations of Metadata

Despite the wealth of information it provides, the quality of the data on the eCourts portal constrains its potential. The [India Justice Report](#), a ranking of the capacity of States to deliver justice vide the police, judiciary, prisons and legal aid relied on eCourts data to conduct the assessment of the judiciary, noting that, “problems like missing data points, discrepancies in different datasets as well as limited access to archival data restrict the full use of such portals”.<sup>9</sup>

Specifically in the context of using eCourts data as a primary data source, studies have largely looked at full legislations or a limited set of fields for specific provisions.<sup>10</sup> Studies and research endeavours that tried to carry out a more extensive analysis of legislations such as the POCSO Act<sup>11</sup>, Negotiable Instruments Act, 1881<sup>12</sup> and legislations relating to land acquisition by the State<sup>13</sup> have found the quality of data on eCourts to be a limitation. For instance, HAQ Centre for Child Rights and CDL, in their 2021 report ‘[Unpacking Judicial Data to Track Implementation of the POCSO Act in Assam, Delhi & Haryana](#)’ (HAQ-CDL Report) stated that “significant challenges were faced during data analysis due to lack of uniformity and standardisation in the use of terminologies, the manner in which information is written and uploaded on the e-Courts portal”.<sup>14</sup>

The errors in entry and lack of standardisation form an obstacle in the goal of transparent and citizen-centric solutions envisioned by the eCourts Project. While recognising the unprecedented potential of eCourts data, Mandyam in 2017 noted that this data is usable only when it is updated, uniform and navigable.<sup>15</sup> Issues such as misspellings, incorrect or incomplete entries, unexplained abbreviations, and lack of standardisation in court data impede the accurate calculation of case pendency.<sup>16</sup> In 2020, Damle and Anand conducted an in-depth study highlighting problems of standardisation, missing information and erroneous entries stemming from manually entered data and inconsistencies across districts.<sup>17</sup> In 2021, Internet Freedom Foundation and CivicDataLab wrote to the Hon’ble Chief Justice of India and Chairperson, e-Committee, Supreme Court of India, drawing attention to the persistent problems of non-standardised and incomplete entries in the eCourts platform.<sup>18</sup>

In 2019, the eCommittee took note of the problem of standardisation in fields on the eCourts platform in Phase II of the eCourts Project and characterised it as an issue of coordination between High Courts.<sup>19</sup> It has attempted to address some of these issues through the latest software deployed in 2018 under the eCourts Project, Case Information System 3.0 (CIS 3.0).<sup>20</sup> However, the exercise of cleaning and organising eCourts metadata in the analysis of the CALPRA evinces gaps that continue a decade after the launch of the eCourt platform. The next section highlights the areas of reform that can enhance the quality of the data and promote greater transparency.

9 ‘[India Justice Report: Ranking States on Police, Judiciary, Prisons and Legal Aid](#)’ (2022) 148.

10 For instance, CDL and IFF used eCourts data to develop a [tracker](#) to follow the application of Section 66A of the Information Technology Act after the Supreme Court declared it unconstitutional, and DAKSH tracks limited trends in specific case types ([writs](#), [bail applications](#) and [tax matters](#)) only in High Courts in select states.

11 See limitations of Vidhi (n 2) 13-17.

12 Damle et al. (n 1) 18.

13 Pooja Murty, Ritwika Sharma and Leah Verghese, ‘Do the Meek Inherit the Earth? A Study of Land Acquisition Litigation in Karnataka and Maharashtra’ (July 2022) 10.

14 HAQ-CDL Report (n 2) 27.

15 Kishore Mandyam, ‘[Reaping the Benefits of the e-Courts System](#)’ in Harish Narasappa and Shruti Vidyasagar (eds), *State of the Indian Judiciary* (DAKSH 2017).

16 Ibid.

17 Devendra Damle and Tushar Anand, ‘Problems with the e-Courts data’, No 314, National Institute of Public Finance and Policy (29 July 2020).

18 For the context in which this letter was written, please see Krishnesh Bapat, ‘We write to the Supreme Court e-Committee with CDL to improve the eCourts platform’ (Internet Freedom Foundation, 12 May 2021).

19 eCommittee, Supreme Court of India, eCourts Project Phase II: Objectives Accomplishment Report as per Policy Action Plan Document (2019) 59: “eCommittee has supplied database of more than 2000 Central Acts and Rules updated with National Codes. Previously Acts were entered as free text entry in local masters therefore there was no uniformity. In one State it is mentioned as IPC, in second State it was Indian Penal Code in third State Penal Code in India. The chronological order of Acts in the master was also not uniform or coherent between two establishments within the same District. This was creating problem[s] in searching entire national database on the strength of Acts and Sections. It was necessary to know how many cases under Section 376 or 302 of the Indian Penal Code or like sections or Acts are pending in the country. Its not possible unless Act and Section uniformity is established through a common mechanism which is acceptable and adoptable by all.”

20 eCommittee, Supreme Court of India, Case Management through CIS 3.0 (2018) 9-11.

## II. Areas of Reform in Maintenance of eCourts Data

The following limitations arising from the gaps in the quality of eCourts data emerged in the course of analysis of the metadata related to the CALPRA cases.<sup>21</sup>

### 1. Errors in entry affecting the reliability of the dataset

The first step in analysing the implementation of the the CALPRA from eCourts data entailed the identification of the cases in which the CALPRA provisions were mentioned. One method was to extract cases with the CALPRA mentioned under the “Act” field. This method hinged on the assumption that the data entry was accurate. However, some cases did not have the CALPRA mentioned under the Act field, although they were indeed the CALPRA cases. This necessitated the use of algorithms with specific search terms using permutations of the CALPRA to identify these cases. Despite these efforts, it is likely that the algorithm was unable to capture all the CALPRA cases, as all variations of references to the CALPRA could not have possibly been known or the CALPRA may not have been mentioned in the metadata.

Image 3: The CALPRA missing in metadata, but mentioned in judgment

Case Type	SESSIONS CASE	
Filing Number	111668/2016	Filing Date
Registration Number	100604/2016	Registration Date:
CNR Number	MHCC020106472016 (Note the CNR number for future reference)	

  

Case Status	
First Hearing Date	26th August 2016
Decision Date	05th February 2020
Case Status	Case disposed
Nature of Disposal	Contested--ACQUITTED
Court Number and Judge	25-COURT 25 ADDL SESSIONS JUDGE

  

Petitioner and Advocate	
[Redacted]	
Respondent and Advocate	
[Redacted]	

  

Acts	
Under Act(s)	Under Section(s)
INDIAN PENAL CODE	370(1),34
The Protection of Children from Sexual Offence Act, 2012	3,14
JUVENILE JUSTICE (CARE AND PROTECTION OF CHILDREN) ACT	23,26

  

**JUDGMENT**  
(Dictated and pronounced in open court)

1. The accused stands chargesheeted for the offences punishable under Sections 370 (1) of Indian Penal Code, section 3 and 14 of Child Labour (Prohibition and Regulation) Act, 1986 and Section 23 and 26 of the Juvenile Justice (Care and Protection of Children) Act, 2000.

Select paragraph of corresponding judgment

Screenshot taken on 6 August 2024 at 07:40

Conversely, some cases with the CALPRA mentioned in the “Act” field were not, in fact, the CALPRA cases and resulted in a false positive. Out of 13,481 cases that mentioned the CALPRA, 2,616 cases had to be removed owing to doubts about them being unrelated to child labour or because the judgment did not contain any references to the CALPRA.<sup>22</sup> While data cleaning helped eliminate some false positives from the data set, the possibility of errors exists.

<sup>21</sup> For a more detailed explanation of how these shortcomings affected our analysis, see Enfold (n 7).

<sup>22</sup> Various elimination methods were used to remove false positives. For instance, 24 cases with the CALPRA in the “Act” field had the entry “MACP” in the field “case type”. Motor Accident Claim Petition cases are usually associated with motor vehicle accident claims making it unlikely to be a the CALPRA case. Out of the 24 cases, 11 were disposed of with judgments available. On checking the judgments, all were found to be motor vehicle accident claims without child labour. On this basis, these 24 cases were eliminated.

Image 4: Case eliminated as metadata reference to the CALPRA was not backed by judgment

Case Type	COMPLAINT CASE ( C R )	
Filing Number	508/2016	Filing Date
Registration Number	508/2016	Registration Date:
CNR Number	ASDU030000282016 (Note the CNR number for future reference)	

  

<b>Case Status</b>	
First Hearing Date	18th April 2016
Decision Date	20th April 2019
Case Status	Case disposed
Nature of Disposal	Contested--Allowed
Court Number and Judge	7-Civil Judge Junior Division No. 1

  

<b>Petitioner and Advocate</b>	
[REDACTED]	
<b>Respondent and Advocate</b>	
[REDACTED]	

  

<b>Acts</b>	
<b>Under Act(s)</b>	<b>Under Section(s)</b>
Child Labour (Prohibition and Regulation) Act	13(1)ChildMarriageProhibitionv

  

Screenshot taken on 6 August 2024 at 07:48

IN THE COURT OF THE JUDICIAL MAGISTRATE FIRST CLASS, DHUBRI

CR Case No. 508/2016

**JUDGMENT**

1. This is a case under Prohibition of Child Marriage Act, 2006 seeking injunction under section 13(6) of the said Act. According to the complainant, his son [REDACTED] was born on [REDACTED] and at the time of alleged occurrence he was a minor and a student of class 12 and on 04.3.16 he was tried to be forcefully married to one [REDACTED] by all the accused persons/ respondents, but the said attempt was foiled by the complainant and his relatives but he fears that his son is in danger as the accused/respondents, are making preparation for getting his son married to [REDACTED] daughter of respondent/accused No 5 and therefore the present complaint is filed.

Select paragraph of corresponding judgment

References to repealed legislations such as the Child (Pledging of Labour) Act, 1933 (CPOLA)<sup>23</sup> were also found, though with references to Sections 3, 3A and 14A, which do not appear in CPOLA, but are provisions under the CALPRA. This points to a strong possibility that these cases were under the CALPRA, but were mistakenly tagged under CPOLA.

These errors render the comprehensive identification of cases under the CALPRA a challenge, and make data cleaning, even with the aid of technological tools, a time consuming and laborious process. They also limit the conclusions that can be drawn from metadata because the principal dataset (a) may not contain all the cases under the selected legislation and (b) may contain cases that do not involve the selected legislation<sup>24</sup>. Other studies have similarly found gaps in the eCourts data a barrier to curating a comprehensive dataset for analysis of a legislation.<sup>25</sup>

## 2. Lack of uniformity in the meaning of the fields

Though the main fields such as case type, Act, Section, court name, judge name, purpose of hearing, etc. are the same across the country and fixed in the eCourts pages, the details for these fields are not selected from a predefined list of values but are entered as free flowing text. This results in multiple variations across and within States.<sup>26</sup> For instance, the field “case type” in some courts indicated the type of proceeding, with varying levels of granularity, whereas in other districts, the same field indicated the legislations applicable<sup>27</sup>.

Below are some of the varying types of information under the “case type” field in Bihar where all the cases have the CALPRA as the applicable legislation as per the field “Act”:

<sup>23</sup> The Repealing and Amending Act 2016.

<sup>24</sup> This problem has been described as a limitation to Damle's analysis of the operation of the Negotiable Instruments Act, 1881 using eCourts data in Damle et al. (n 1) 18 and DAKSH's analysis of litigation under the Land Acquisition Act in Maharashtra and Karnataka in Pooja Murty, Ritwika Sharma and Leah Verghese, 'Do the Meek Inherit the Earth? A Study of Land Acquisition Litigation in Karnataka and Maharashtra' (July 2022) 10; See also Damle and Anand (n 17), 11: "Cases could be tagged under a procedural law, substantive law, or both. Non-uniformity in tagging cases means that there is a high likelihood of a search for cases under an Act would contain an incomplete set of cases for a given court complex. This inconsistency in tagging of cases results in incomplete information about the actual litigation under a given statute."

<sup>25</sup> HAQ-CDL Report (n 2) 18-24, 29-31.

<sup>26</sup> DAKSH, 'Deciphering Judicial Data: DAKSH's Database' (August 2020) 8-19.

<sup>27</sup> Damle and Anand (n 17).

Image 5: Intra-state variations in entries under “case type”

CJM Div Dalsinghsarai			
Case Details			
Case Type	CRI. CASE		
Filing Number	626/2022	Filing Date	21-06-2022
Registration Number	246/2022	Registration Date:	07-07-2022
CNR Number	BRSM110006262022 (Note the CNR number for future reference)		<a href="#">View QR Code / Cause Title</a>

Screenshot taken on 17 May 2024 at 09:45

Case type indicates that it is a criminal case.

CJM Division Araria			
Case Details			
Case Type	GR POLICE CASES		
Filing Number	4278/2022	Filing Date	19-09-2022
Registration Number	2124/2022	Registration Date:	19-09-2022
CNR Number	BRAR020042782022 (Note the CNR number for future reference)		<a href="#">View QR Code / Cause Title</a>

Screenshot taken on 15 June 2024 at 19:00

CJM Div. Patna City			
Case Details			
Case Type	CR. CASE COMPLAINT (O)		
Filing Number	41/2010	Filing Date	01-04-2010
Registration Number	41/2010	Registration Date:	02-10-2018
CNR Number	BRPA110008372010 (Note the CNR number for future reference)		<a href="#">View QR Code</a>

Screenshot taken on 15 June 2024 at 18:55

Case type, in these cases, indicates the manner the criminal case was filed.

CJM Div. Saran at Chapra

Case Details

Case Type	WEIGHT AND MEASUREMENT ACT		
Filing Number	210018/2015	Filing Date	11-02-2015
Registration Number	210018/2015	Registration Date:	11-02-2015
CNR Number	BRSR020114782015 (Note the CNR number for future reference)		<a href="#">View QR Code / Cause Title</a>

Case Status

First Hearing Date	28th May 2017
Next Hearing Date	29th October 2024
Case Stage	Appearance
Court Number and Judge	43-ACJM-9

Petitioner and Advocate

1) State of Bihar

Respondent and Advocate

1) [REDACTED]

Acts

Under Act(s)	Under Section(s)
Child Labour Act	14(1)

Screenshot taken on 17 May 2024 at 10:26

In this case, a legislation is placed instead of the type of proceeding. Notably, this legislation does not appear in the field of applicable acts.

Such variations make it **difficult to gain insight into the exact type of case proceeding** and the number of cases thereunder within a State, and across States. Further, it complicates the determination of the dataset consisting of cases under a particular legislation. For instance, in the figure above, one of the cases appears to be under the improbable combination of both the CALPRA and the Weights and Measures Act. While attempts were made to ascertain the facts, in pending cases where there is no judgment, this is impossible. The limitation that the entry of either the CALPRA or the Weights and Measures Act may be erroneous cannot thus be overcome.

A similar lack of clarity exists in other fields. Entries in the fields of “Under Act(s)” and “Under Section(s)” are jointly read to glean the charges in criminal cases. It is unclear whether these charges appearing in the metadata reflect the charges in the chargesheet or the charges framed by the court.

### ***Unclear and incomplete information***

Certain entries had local connotations, were indecipherable abbreviations or were incomplete, making the entry unclear.<sup>28</sup> While court practices and references may vary from State to State according to the applicable court rules, such variations may require additional entry types within the same field, **in the absence of a key specific to each jurisdiction, such entries cannot be deciphered**. For instance, there were certain entries in the “case type” field that were not comprehensible such as “C-7”, “C.t.”, “C.g.” and “C2(cia)”. Similarly, in the field “purpose of hearings”, there were entries such as “Batta”, “Put Up”, “NOT HEARD AT 2.45 P.M.” from which it was not possible to discern the nature of hearing. **The purpose of hearings can provide vital insights on how judicial time is spent and how many effective hearings take place in a case but because of entries like this analysis of this nature was limited.**

In some cases, there are no entries in the fields at all, and this has been highlighted in other reports too. For instance, the HAQ-CDL Report on POCSO Act found that key aspects, including FIR year and districts were missing.<sup>29</sup> Damle and Anand noted that the statute name field often lacked entries or contained procedural statutes only, while the section number, essential dates, final orders and related matters fields were empty in several cases.<sup>30</sup>

28 HAQ: Centre for Child Rights and Forum Against Sexual Exploitation of Children, ‘Implementation of the POCSO Act Goals, Gaps and Challenges’ (2018) 50: “Even the court data was difficult to comprehend due to use of abbreviations that may not be commonly known or used. As the police and the courts gear up towards a digitised data management system, these details will have to be borne in mind to reduce discrepancies and ensure more accurate information is collected and disseminated.”

29 HAQ-CDL Report (n 2) 25-26.

30 Damle (n 24) 12-18.

### **3. Variations in the form of entry affect granular analysis**

Where there is consistency in the broad meaning of the field, there are multiple variations of the same entry with differences in spelling, abbreviations and references peculiar to specific jurisdictions. For instance, in the the CALPRA study, the name of the forum, “District and Additional Sessions Judge” had 70 variations<sup>31</sup>. While this is understandable, as such variations may enable litigants and lawyers to easily identify the forum, reorganisation of this data can promote access to granular data on the types of cases being dealt with and the working of Special Courts.

Entries such as “District and Additional Sessions Judge 6-cum-Special Judge SC and ST Act” and “District and Adl Sessions Judge X cum FTC Crime Against Women” indicate the special courts designated for specific types of cases. However, as this is not a consistent practice across jurisdictions, conclusive information about special courts dealing with child labour cases is unavailable. Further, multiple variations in spelling and abbreviations and inclusion of the name of the district or the judge, require manual checking and considerable resources to harmonise entries as it is challenging for an algorithm to recognise all possible colloquial variations of the designation of a judge. This inhibits the depth of analysis on forums dealing with certain special types of cases.

With respect to the study on the CALPRA, the extent to which Children’s Courts were trying cases of child and adolescent labour could not be effectively gauged. Under Section 25 of the Commissions for the Protection of Child Rights Act, 2005 a Sessions Court can be designated to be a Children’s Court for speedy trial of offences against children or of violation of child rights. According to the metadata, only eight cases were tried in a Children’s Court in one district in Bihar. It is possible that there are more Children’s Courts dealing with cases of child labour, but this may not have been captured during data entry. As there is no consistency or guideline mandating that the special status of courts and designated courts be specified in eCourts either in the forum name or as a separate field, a valuable opportunity to track efficacy of special courts set up to implement specific legislations is lost.

31. For a list of variations in entries in various fields, and how they were reclassified for the the CALPRA study, refer to “reclassified variables” dataset uploaded at <https://justicehub.in/dataset/analysis-the-CALPRA-2015-2023/resource/c1027618-b24c-4d53-8e48-0ba5621a4fca/>.

*Scale of re-organisation of entries due to multiple unique variations*

Metadata Fields analysed						
Case type	Legislations “Under Acts”	Judge designation	Nature of Disposal		Purpose of Hearing	
			Criminal Trials	Bails and Antic- ipatory Bails	Disposed Criminal Trials	Pending Criminal Trials
294 uniques Case types	161 unique entries	1249 unique designations	165 unique outcomes	41 unique outcomes	302 unique purpose of hearing	207 unique latest purpose of hearing
4 Categories	88 standardised legislations names	6 Categories	7 Categories	3 Categories	7 Categories	
Criminal Trials  Appeals and Revisions  Criminal Misc  Unclear/ unknown	For e.g. all variations of POCSO Act were standardised to ‘Protection of Children from Sexual Offences Act, 2012”	Judicial Magistrate of the Second class  Judicial Magistrate of the First Class  Chief Judicial Magistrate  Sessions Judge (including Additional Sessions Judge)  Assistant Sessions Judge  Civil Judge	Acquittal  Conviction  Committal  Conciliation/ Compromise  Pending  Transfer  Others (Abated, Absconded, Allowed, Discharged, Dismissal, Disposed, Miscellaneous, Sec.258 Cr.P.C., Stopped, Untraced)	Bail Granted  Bail Rejected  Others (Dismissal, Withdrawn, Disposed, Contested, Transfer, absolute)	Preliminary Trial Procedures  Appearances, Summons and Warrants  Statement of Accused and Evidence  Arguments  Others  Orders/Judgment  Not Clear/Not Available	

#### 4. Suboptimal presentation of available information

Certain information on eCourts is presented in a manner that makes it difficult to comprehend, analyse or extract for further analysis. For instance, there is no uniform method of depicting an exhaustive list of outcomes of a case. In 857 the CALPRA cases, the field “nature of disposal” included entries such as “judgement”, “disposed” and other entries that only indicate that the matter has been disposed of without description of the actual outcome such as conviction, acquittal, discharge or other outcome.

Image 6: Imprecise descriptions of case outcomes

**Case Details**

Case Type	PRC	
Filing Number	460/2019	Filing Date
Registration Number	186/2019	Registration Date:
CNR Number	<b>ASKK030004602019</b> (Note the CNR number for future reference)	

**Case Status**

First Hearing Date	09th April 2019
Decision Date	11th July 2019
Case Status	Case disposed
Nature of Disposal	<b>Contested--Judgement Delivered</b>
Court Number and Judge	5-Addl. Chief Judicial Magistrate

DISTRICT:-KOKRAJHAR  
IN THE COURT OF Addl. CHIEF JUDICIAL MAGISTRATE, KOKRAJHAR

GR Case No.-156/2018 & PRC-186/2019,  
U/S-3/14 of Child Labour (P & R) Act.

**6) O-R-D-E-R**

In the result, the accused is guilty u/s 14 of the Child Labour (P & R) Act.

Then I heard the accused as well as the prosecution on the sentence of punishment. Accused has stated that he brought the victim for financial support of the child and for the child friendly atmosphere for the favour of his wife. He has no previous criminal record. He would not repeat any such incident in near future. Thus, he pleaded to have mercy.

Considering the nature of the case, circumstance of the occurrence of the alleged offence as well as all other antecedents of the accused, I am of the view that the accused deserves to get a lenient punishment. But then, if the accused is not sentenced with minimum punishment, then also a wrong message will carry to the society.

Considering all these facts, I am of the view that the accused is not entitled any benefit under Section-3/4 of the Probation of Offenders' Act-1958 and section 360 of CrPC. Accordingly, with the aforesaid reasons, I convicted the accused and sentenced him with fine of Rs.12,000/- only (Rupees twelve thousand only), in default, rigorous imprisonment for the term of six months (06) for the offence of Section 14 of the Child Labour

Screenshot taken on 6 August 2024 at 08:15

Select paragraph of corresponding judgment

In the case above, *State of Assam v Swapon Kr. Sen*,<sup>32</sup> the nature of disposal merely mentions that judgment was delivered when there was conviction under the CALPRA.

**Case Details**

Case Type	G R CASE	
Filing Number	53/2014	Filing Date
Registration Number	53/2014	Registration Date:
CNR Number	<b>ASGL080000152014</b> (Note the CNR number for future reference)	

**Case Status**

First Hearing Date	07th August 2017
Decision Date	10th September 2018
Case Status	Case disposed
Nature of Disposal	<b>Contested--CONVICTED</b>
Court Number and Judge	2-Civil Judge (Jr. Div.) cum JMFC

IN THE COURT OF JUDICIAL MAGISTRATE FIRST CLASS,  
BOKAKHAT, GOLAGHAT

G.R.Case No.- 53/14

25. In view of the above discussion, it is concluded that the prosecution has been able to successfully establish the offence under section 323 I.P.C. against the accused person [REDACTED] reasonable doubt. Accordingly, the accused person of this case, Sri [REDACTED] is found guilty under section 323 I.P.C. and he is convicted under the said section of I.P.C. On the other hand, prosecution has failed to establish the guilt of the accused person under section 23 of the juvenile justice (Care and Protection of Children) Act, 2000 and section 14 of the Child Labour (Prohibition and Regulation) Act, 1986 and hence, the accused person Sri [REDACTED] is acquitted of the offences under section 23 of the juvenile justice (Care and Protection of Children) Act, 2000 and section 14 of the Child Labour (Prohibition and Regulation) Act, 1986.

Screenshot taken on 6 August 2024 at 08:24

Select paragraph of corresponding judgment

In the case above, *State of Assam v Sri Bicky Sharma*,<sup>33</sup> the nature of disposal indicated that there was a conviction. However, there was no conviction under the CALPRA and the accused was convicted of an offence under the IPC.

<sup>32</sup> G.R. Case No.156/2018 decided by the Additional Chief Judicial Magistrate, Kokrajhar on 11 July 2019.

<sup>33</sup> G.R.Case No.53/14 decided by Judicial Magistrate First Class, Bokakhat, Golaghat on 10 September 2018.

Further, in cases where the “nature of disposal” indicates a conviction, it is not clear which charges the accused is convicted of. For instance, there were cases wherein the outcome indicated that the accused was convicted, but the judgment revealed that the conviction was for an offence under the IPC, not the CALPRA. More granular information indicating the outcomes vis-à-vis each of the offences with which the accused may have been charged will expand the quality of the evidence on utilisation of legal provisions, and help identify trends in outcome as per the provisions applied. It can also help identify specific areas for capacity building to strengthen investigation and prosecution, as well as enhance support to victims. For instance, if the granular data reveals that convictions are higher for allied offences under IPC, instead of the CALPRA, it would necessitate a qualitative inquiry into the reasons for the failure of the charge under the CALPRA so that corrective measures can be taken.

As a result of incorrect and incomplete entries, coupled with the numerous variations in the entries, the opportunity for granular analysis of the charges is lost. Apart from considerable time and resources being spent on cleaning and categorising data because of these types of entries, some analysis was rendered impossible. For instance, the proportion of criminal trials under the CALPRA that involved child labour and adolescent labour, respectively, could not be gauged because of the lack of specificity in the entries on sections. Section 3 of the the CALPRA read with the penalty prescribed in Section 14(1) prohibits child labour and Section 3-A read with penalties under Section 14(1-A) prohibits adolescent labour in hazardous industries. Given that the field “sections” in the metadata is supposed to capture the exact provisions applied in a case, the proportion of cases involving child labour, adolescent labour and both, could have been obtained. However, the sub-sections under Section 14 were not mentioned in most cases, rendering it impossible to discern the exact offence. Where sub-sections were mentioned, multiple variations in the metadata with combinations of spaces, brackets and special characters complicated the analysis

Another instance is with respect to information relating to cases that get transferred between forums. When a case is transferred, the case is assigned a fresh Case Number Record (CNR) in the next forum. These are difficult to trace as the CNRs are often not linked to each other on the eCourts portal. In such cases, it becomes difficult to ascertain the entire duration and track the outcome in the case.

While this data is intended for public consumption, these instances show that the sub-optimal presentation results in an information gap or requires considerable time and effort to extract information

### III. Conclusion and Suggestions

Court data from the eCourts platform is a valuable source of data that can advance accountability, enhance the understanding of the implementation of laws and guide policy reform. However, the inconsistencies and errors in data entry, geographical variations and the suboptimal manner in which data is presented hinders comprehensive and effective analysis. Considerable time and resources are required in making the data fit for analysis and, in any event, findings are limited by the quality of data.

There are ongoing efforts amongst civil society to make the data amenable to nuanced analysis. For instance, DAKSH attempted to harmonise eCourts data by building a database based on select fields noting that “[w]hile the ideal goal would be to ensure uniformity in all the data available online [on eCourts], the objective of this ... is to focus on a few key variables that are preventing a more in-depth analysis of judicial data, i.e., case types, stages, and subject matter classification.”<sup>34</sup> Similarly, Justice Hub is making available a platform for researchers in the field of law and justice to contribute and download data, including eCourts data in cleaned formats. At the same time, the eCommittee of the Supreme Court has attempted to address some of the gaps in the data available. The latest case management software CIS 3.0 deployed under the eCourts Project meant to standardise fields and some entries<sup>35</sup>.

The requirements of researchers and data analysts differ from the eCourt platform’s original intent to provide information to litigants and lawyers regarding their cases. However, addressing these gaps and enhancing the data platform to accommodate diverse user needs could significantly contribute to transparency, accountability and evidence-based policy-making.

**It is urged that the eCommittee recognize data analysts, civil society organisations and researchers as crucial stakeholders and specific user categories of the eCourts data, ensuring that the eCourts platform evolves to serve not only its primary users but also the broader research community, ultimately strengthening the Indian justice and legal system.**

The discrepancies and gaps in eCourts data may be a symptom of larger systemic issues, challenges faced by implementing agencies including coordination between the National Informatics Centre that designs solutions for eCourts and local courts, quality of digital infrastructure, and constraints related to personnel and resources. While strategies and resources required in the long term to address these factors are beyond the scope of this report, the following immediate suggestions may be considered by the eCourts Committee to strengthen eCourts as a data source to advance evidence based monitoring of laws and interventions:



1. **Data Key:** The level of granularity in the entries presented and what they mean should be uniform to the extent possible across districts with uniform instruction for data entry. This needs to be supplemented with a State and district specific key of terms and what they signify so as to help users understand each entry and guide insights that can be drawn from the information available. The data-entry key may be developed in consultation with all States and be available for reference by all persons accessing the ecourts platform. An exhaustive list of possible entries with their full form and meaning for each field can be a part of this key This will promote uniformity in data-entry and thereby enhance the quality of the data.



2. **Standardised Entries:** The entries for each field needs to be standardised such that there is only one spelling or abbreviation for each possible entry across all districts. Entries in free form need to be minimised and drop down menu options with standardised references to entries may be considered. The options can be customised to take into account jurisdiction-specific procedures and terminology and this may be supplemented with a key explaining each option. Additional information such as the court number

<sup>34</sup> DAKSH (n 26).

<sup>35</sup> eCommittee, Supreme Court of India, [Case Management through CIS 3.0](#) (2018) 9-11; DAKSH (n 26) 20-21.

where applicable, special status of the court and name of the judge can be entered as separate fields rather than sporadically mentioned or combined in the judge designation field.



3. **Comprehensive Reference to Applicable Legislations:** With respect to the Act and Section fields, all substantive applicable laws and provisions need to be mentioned.<sup>36</sup> This will allow for the reliable creation of datasets based on these fields and facilitate analysis of specific legislation. In addition, it will enable a clearer understanding of how different laws interact and are applied to address various issues—an analysis that is challenging with other existing data sources and is a significant advantage of eCourts data.



4. **Automated data entry:** Intelligent and automated data entry needs to be considered to reduce human error in data entry.<sup>37</sup> For instance, the option of applying a repealed legislation for a case filed after its repeal either should not exist or should be accompanied with a prompt or caution that the legislation opted for is now repealed.



5. **Disaggregated data on Disposal:** In the field of nature of disposal, the disposal needs to be recorded against each provision and legislation applicable. For instance, in criminal cases, the outcome against each offence can be specified distinctly.



6. **Linking of details of transfer:** Proceedings subsequent to transfer need to be linked on the same landing page so that the entire duration and outcome of the case can be tracked.



7. **Regular Consultations:** The eCourts Committee may consider conducting regular regional consultations to seek suggestions and solutions from users, researchers, data scientists and technology solution providers to improve data entry, collection and presentation.

<sup>36</sup> Not all legislations applicable are mentioned on the eCourts platform. See Damle and Anand (n 17) 8: “Karnataka and Tamil Nadu report no cases under SR-Act [Specific Relief Act, 1963]. This is a result of the variance between what each state considers a case. Contract disputes in these states are filed under provisions of the Code of Civil Procedure (CPC), without any reference to the substantive law under which the dispute falls.”

<sup>37</sup> See Damle and Anand (n 17) 20



### About Enfold

Enfold Proactive Health Trust is a non-profit based in Bengaluru founded in 2001, and is working towards building safer, rights affirmative, inclusive spaces and justice systems for children and communities.

For more information, visit [www.enfoldindia.org](http://www.enfoldindia.org)

Contact us: [info@enfoldindia.org](mailto:info@enfoldindia.org)



### About CivicDataLab

CivicDataLab works at the intersection of data, technology, design and social science to strengthen data for public good in India and other countries. The organisation harnesses the potential of open knowledge movements to institutionalise data practices, strengthen data capacities and enable citizens to participate in governance. It works with governments, civil societies, think-tanks, universities and other stakeholders to enhance their data and technology capacity to adopt data-driven decision-making at scale. Current areas of expertise include digital public goods & infrastructure (DPGs & DPI), climate action, public finance, urban development, open contracting and law & justice.

Contact us: [info@civicaldatalab.in](mailto:info@civicaldatalab.in), Website: [www.civicaldatalab.in](http://www.civicaldatalab.in) Twitter: <https://x.com/civicaldatalab>

LinkedIn: <https://in.linkedin.com/company/civicaldatalab>