

G-4

**EDINET API を用いたデジタルフォレンジック学習教材向け資料取得プログラムの評価**  
**Assessment of program to acquire documents for digital forensics learning materials using EDINET API**

○松浦芽生<sup>1</sup>, 五味悠一郎<sup>2</sup>

\*Mei Matsuura<sup>1</sup>, Yuichiro Gomi<sup>2</sup>

Abstract: The purpose of this research was to clarify that the use of digital-forensics learning materials can help beginning students easily learn digital-forensics technology and increase their interest in information security. In the previous research, dummy data were generated for the acquisition of digital-forensic techniques. However, it took a lot of time to generate dummy data. Therefore, we developed a generic document acquisition program using EDINET API to shorten the dummy data generation time. In this paper, we verify and evaluate that the developed program can be easily used by anyone to acquire materials in a short time.

1. はじめに

経済産業省による調査では日本の人口減少に伴い、情報セキュリティ人材の不足数は今後も拡大していくと、試算されている。こうした状況に対応するためには、情報セキュリティへの興味を向上させる適切な教材開発が必要であると考えられる。<sup>[1]</sup>

教材の範囲は「デジタル化の進展や訴訟の増大により必要性が高まっている」「IoT化の進展によりデジタルで扱う情報が膨大化し、重要性が増していく」と考えられるデジタルフォレンジックに絞り込むこととする。<sup>[2]</sup>

デジタルフォレンジックとは、種々のインシデントが発生した際に、コンピュータなどの情報処理機器上に残された証拠を確保し、将来起こり得る裁判に備えるための技術や手順のことである<sup>[3]</sup>。デジタルフォレンジック技術の習得には、多くのデータ（以下、ダミーデータとする。）が含まれたディスクイメージファイルを必要とするが、このファイルを講師が短時間で用意することは難しい。本稿では EDINET API を用いた資料取得プログラムによって、2022年度の生成方法よりも短時間で誰でも容易にダミーデータを生成できることを明らかにすることを目標とする。

2. 先行研究

2019年に服部は、初学者が短期間でデジタルフォレンジックを学ぶことができるシナリオ型教材を開発し、開発した教材を用いた講義と演習を実施した。<sup>[4]</sup>

2022年には、短時間で誰でも容易にデータ生成が行えることを目標として、個人情報自動生成プログラムの開発を行った。ダミーデータに含まれる情報として、個人情報を自動生成し、汎用的資料を手作業で作成し

た。結果はダミーデータの生成に多くの時間を要し、「ダミーデータの生成時間を短縮するためには、手作業で作成した汎用的資料を自動的に生成する必要がある」という課題が明らかになった。ダミーデータの生成ツールは Google Colaboratory を使用したが、被験者から「使い慣れていないツールを使うため、難しい」という意見が示された。

3. ダミーデータ生成実験について

本稿では本研究で開発したプログラムによって、2022年度の生成方法よりも短時間で誰でも容易にダミーデータを生成できることを明らかにするため、ダミーデータ生成実験を実施した。<sup>[5]</sup>

3.1. ダミーデータについて

本稿で実施したダミーデータ生成実験で、被験者が生成したダミーデータ一覧を Table 1 に示す。「海外顧客情報.txt」の生成には、2022年度に作成したプログラムを2023年度も使用した。2022年度は「海外顧客情報.txt」のみ自動生成であったが、2023年度は汎用的資料の取得プログラムを開発したことにより、すべてのファイルが自動生成となる。

Table 1. Dummy Data List

	2022年度ダミーデータ	2023年度ダミーデータ
	海外顧客情報.txt	
汎 用 的 資 料	確定申告書フォーマット.pdf	S100061V有価証券報告書.pdf
	基礎控除申告書.pdf	S10006YR有価証券報告書.pdf
	源泉徴収簿.pdf	S10007A9有価証券報告書.pdf
	扶養控除申告書.pdf	S10007JA有価証券報告書.pdf
	保険料控除申告書.pdf	S10007N7有価証券報告書.pdf

1 : 日大理工・学部・情報 2 : 日大理工・教員・情報

### 3.2. ダミーデータ生成実験方法について

ダミーデータの生成者は講師を想定しているため、被験者は情報技術に関して基礎的な知識を保有する情報系の学生6名とした。Table 1 に示したダミーデータのうち、海外顧客情報.txt と 2023 年度汎用的資料の生成ツールは Google Colaboratory を用いた。実験で使用するプログラムと手順書は Google ドライブで共有し、手順書に従いダミーデータの生成と生成所要時間の計測を行ってもらった。2023 年度実験手順書は 2022 年度実験手順書に使用ツールである Google Colaboratory の使用方法を追加し、スクリーンショットを増やすなどの改変を行った。

### 4. 評価方法

生成実験終了後に Google フォームからアンケートへ回答してもらい、アンケート結果から評価を行った。各アンケート項目には 1 点から 5 点の間隔尺度を設定した。アンケート結果より、2023 年度の生成方法のほうが 2022 年度の生成方法よりも短時間でダミーデータ生成が行えるかを検証した。

### 5. 結果と考察

ダミーデータの平均生成所要時間を Table 2 に示す。

Table 2. Dummy data average generation time

2022年度ダミーデータ生成結果		2023年度ダミーデータ生成結果	
ファイル名	平均生成所要時間	ファイル名	平均生成所要時間
海外顧客情報.txt	14分12秒	海外顧客情報.txt	14分12秒
確定申告書フォーマット.pdf	4分32秒	S10006IV有価証券報告書.pdf	2分31秒
基礎控除申告書.pdf		S10006YR有価証券報告書.pdf	
源泉徴収簿.pdf		S10007A9有価証券報告書.pdf	
扶養控除申告書.pdf		S10007JA有価証券報告書.pdf	
保険料控除申告書.pdf		S10007N7有価証券報告書.pdf	
合計	18分44秒	合計	16分43秒

Table 2 より、2023 年度の汎用的資料は 2022 年度よりも 121 秒早い、2 分 31 秒で生成することができた。開発した資料取得プログラムにより、2022 年度よりも短時間でダミーデータの生成が行えた。ダミーデータ全体の生成所要時間をさらに短縮するには、海外顧客情報生成プログラムの改良が必要であると考えられる。

アンケート結果を Table 3 に示す。

Table 3. Questionnaire results

質問事項	最小尺度	最大尺度	平均値	標準偏差
2022年度ダミーデータ生成方法と2023年度ダミーデータ生成方法の難易度の比較	2022年度のほうが難しい	2023年度のほうが難しい	2.33	0.94
2022年度ダミーデータ生成方法と2023年度ダミーデータ生成方法の手順量の比較	2022年度のほうが手順量が多い	2023年度のほうが手順量が多い	1.33	0.47
海外顧客情報生成の難易度はどうだったか	難しい	易しい	3.00	1.00
2022年度ダミーデータ生成の難易度はどうだったか	難しい	易しい	3.00	1.53
2023年度ダミーデータ生成の難易度はどうだったか	難しい	易しい	4.83	0.37
2022年度実験手順書のわかりやすさはどうだったか	わかりにくい	わかりやすい	2.00	1.00
2023年度実験手順書のわかりやすさはどうだったか	わかりにくい	わかりやすい	4.33	1.11

Table 3 より、2022 年度ダミーデータ生成難易度は平均値 3.00、標準偏差 1.00 で、2023 年度ダミーデータの生成難易度は平均値 4.83、標準偏差 0.37 であるため、2022 年度よりも誰でも容易にダミーデータ生成が行え

たとえる。2022 年度手順書のわかりやすさは平均値 2.00、標準偏差 1.00 で、2023 年度手順書のわかりやすさは平均値 4.33、標準偏差 1.11 であった。2023 年度手順書は、2022 年度手順書よりもわかりやすかったといえる。

### 6. まとめと今後の課題

本稿では資料取得プログラムを用いて被験者実験を行い、一部手作業であった 2022 年度よりも短時間で誰でも容易にダミーデータ生成が行えることを検証した。

アンケート結果より、2023 年度に作成した手順書は 2022 年度に作成された手順書よりわかりやすいことが明らかになった。実験結果とアンケート結果から、開発したプログラムを用いることで、汎用的資料は 2022 年度の生成方法よりも短時間で誰でも容易に生成できると明らかになった。よって、本稿の目標は達成できた。

今回の実験結果から、ダミーデータの生成に時間を要していた原因は、海外顧客情報の生成にあったことが明らかになった。今後はダミーデータ生成所要時間を短縮するため、海外顧客情報生成プログラムの改良や生成手順の再考を行う。

### 7. 参考文献

- [1] 商務情報政策局 情報処理振興課. “IT 人材の最新動向と将来推計に関する調査結果”. 経済産業省. 2016-06-10. [https://www.meti.go.jp/shingikai/economy/daiyoji\\_sangyo\\_skill/pdf/001\\_s02\\_00.pdf](https://www.meti.go.jp/shingikai/economy/daiyoji_sangyo_skill/pdf/001_s02_00.pdf). (参照 2023-2-26)
- [2] 特定非営利活動法人デジタル・フォレンジック研究会. 「基礎から学ぶデジタル・フォレンジック 入門から実務での対応まで」. 安富潔, 上原哲太郎 編. 日科技連出版社. 236p. 2019.
- [3] 上原哲太郎, 桜庭信之, 白濱直哉, 野崎周作, 八槇博史, 山本清子. 「デジタル・フォレンジックの基礎と実践」. 佐々木良一 編. 東京電機大学出版局. 290p. 2017.
- [4] 服部慶. “デジタルフォレンジック技術を短期間で習得する教材の開発”. 令和元年度日本大学理工学部学術講演会予稿集. <https://www.cst.nihon-u.ac.jp/research/gakujutu/66/pdf/G-27.pdf>. (参照 2023-09-28)
- [5] 松浦芽生, 五味悠一郎. “EDINET API を用いたデジタルフォレンジック学習教材向け資料生成プログラムの開発”. 日本ソフトウェア科学会第 40 回大会講演論文集. 2023.