

# 課題提起フォーラム

「ナショナルな地域文化資源：  
地方紙の活用に向けて — 地方紙原紙の  
デジタル化状況調査から見えてきたこと」

2017年7月13日

東京大学大学院情報学環

DNP学術電子コンテンツ研究寄付講座

協力：日本新聞協会

# 地方紙デジタル化・活用プロジェクトの概要

## ■趣旨

デジタルデータになっていない古い地方新聞について、  
学術利用はもとより、我が国の貴重な文化資源の保存と  
活用を図る観点から、課題の整理と活用方法を探る。

## ■全体プログラム策定委員会の設置（2016年12月～）

- ・メンバー（敬称略）：植村八潮（専修大），柴野京子（上智大），  
丹羽美之（東大），松岡資明（学習院大），  
柳与志夫・時実象一・東由美子・  
平野桃子（以上 DNP講座）

# 地方紙デジタル化・活用プロジェクトの概要

## ■ 検討課題

- ①全国の主要な地方紙（県紙レベルを想定）のデジタル化状況の把握
- ②古い原紙デジタル化の問題点及びデジタル化記事活用の現状・課題の分析
- ③全国主要地方紙デジタル化の工程と運営管理の見通し
- ④知識基盤としての新聞記事活用のあり方とビジネスモデルの提示
- ⑤全体プログラムの問題点の把握、実施可能性の検証等を目的とする  
パイロット事業計画

## ■ 全国主要地方紙のデジタル化状況調査実施（日本新聞協会との共同調査）

- ・ 調査対象：日本新聞協会加盟の地方新聞社（全73社）
- ・ 調査期間：2017年2月～4月
- ・ 回答社数：47社（回答率64.4%）

※別添：質問票

# 集計結果

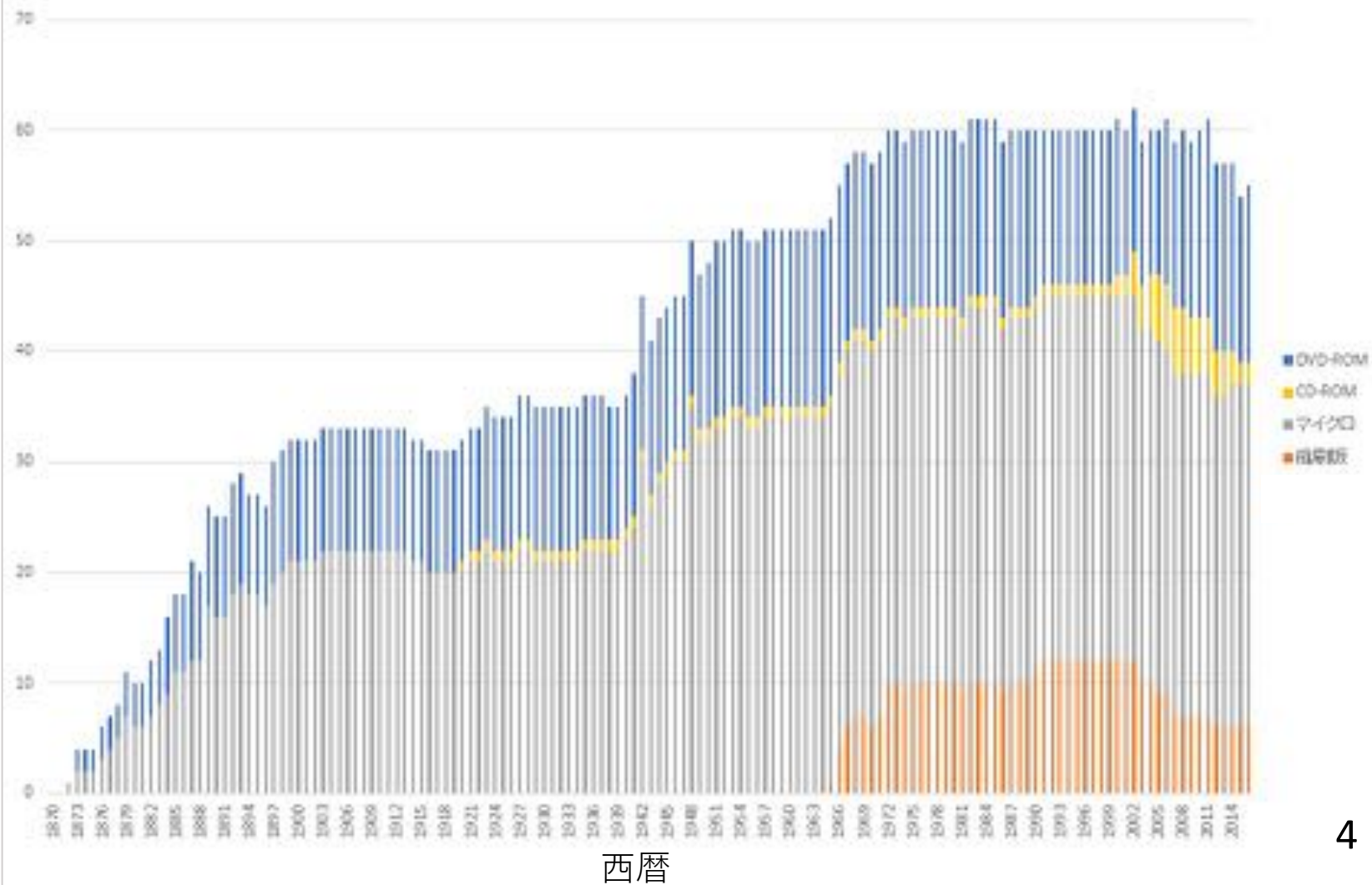
## ■目次

- 1 NDLサーチによる発行状況（各社からの訂正回答反映済み）
- 2 各社保存状況
- 3 NDL保存状況（NDLサーチによる）
- 4 公開（オンライン閲覧）
  - ① 日経テレコン
  - ② nifty
  - ③ G-Search
  - ④ その他
- 5 未公開（オンライン閲覧）
  - ① 紙面イメージ（記事イメージを含む）
  - ② 文字
  - ③ 写真

# 集計結果 1

1 NDLサーチによる発行状況（各社からの訂正回答反映済み）

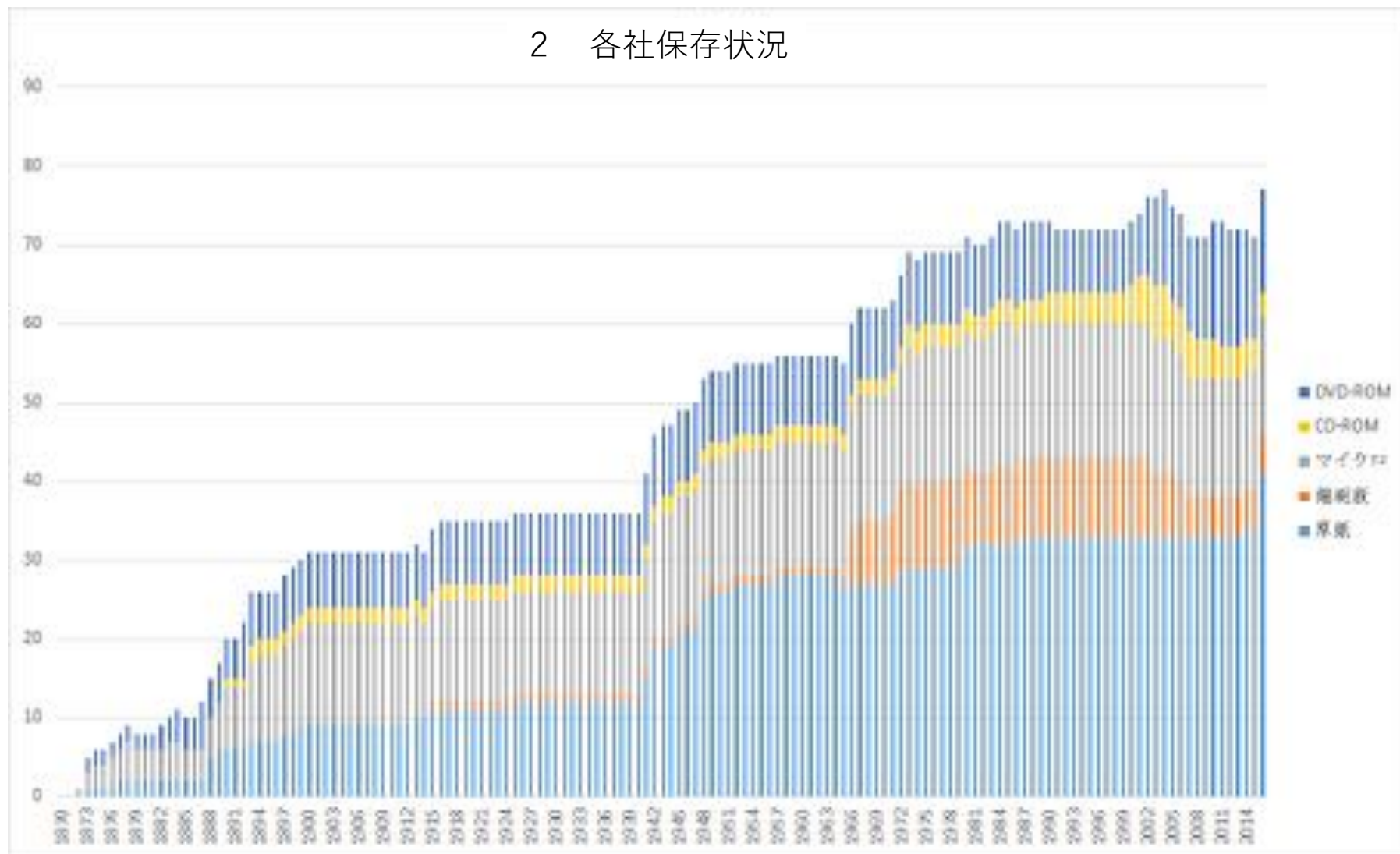
社数



# 集計結果 2

## 2 各社保存状況

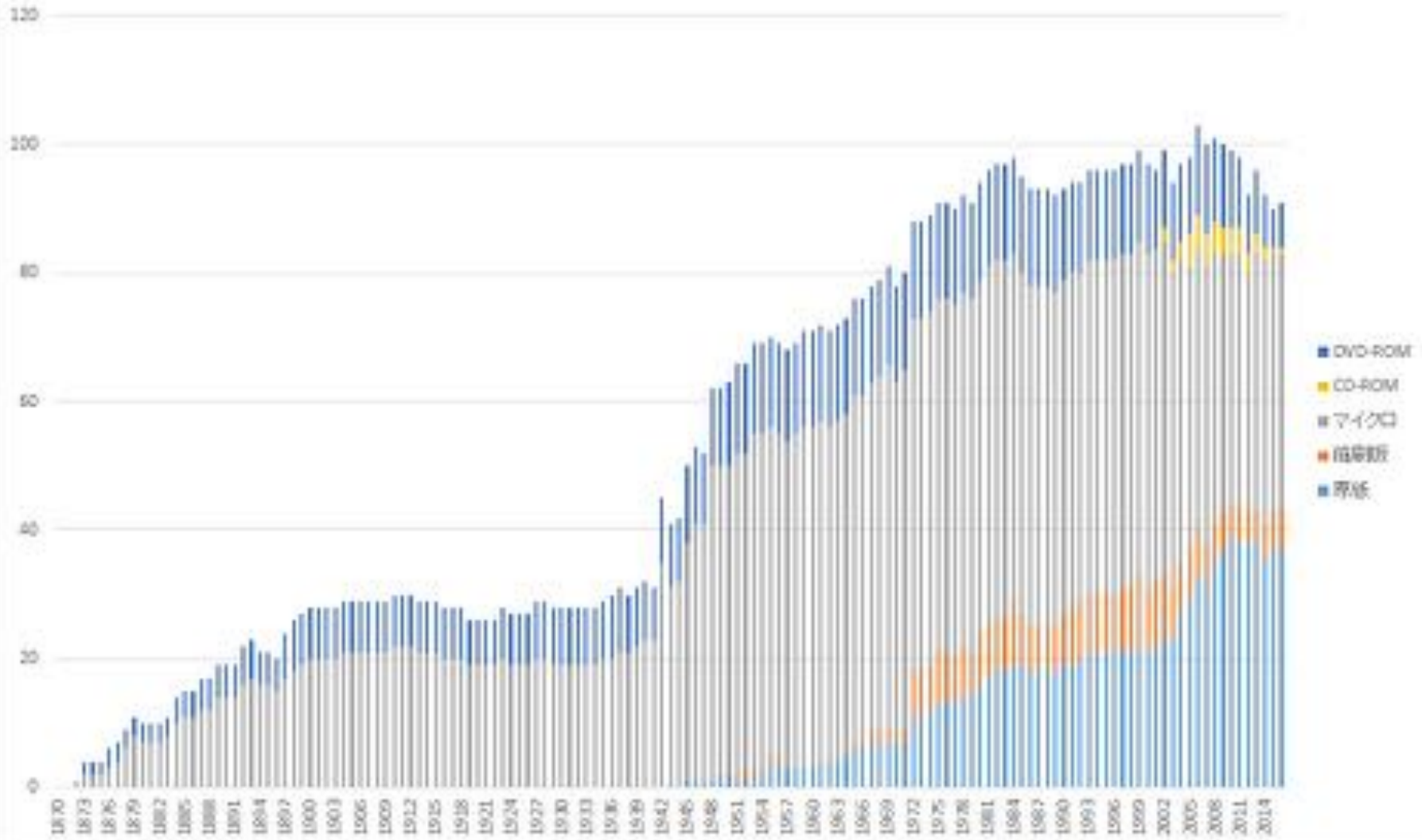
社数



西暦

# 集計結果 3

## 3 NDL保存状況 (NDLサーチによる)



社数

西暦

# 集計結果 4 - ①

4 - ① 公開（オンライン閲覧）-日経テレコン

社数

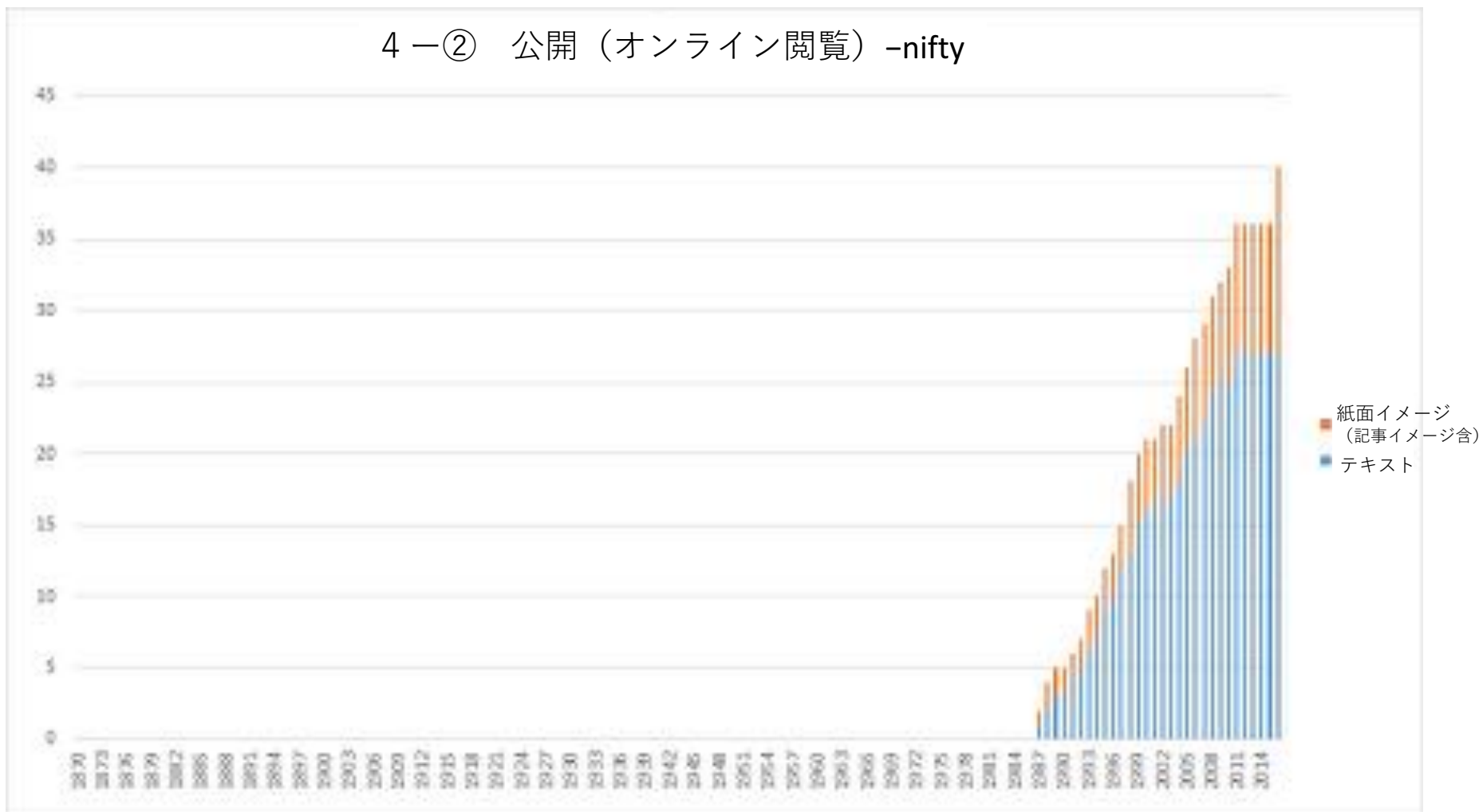


西暦



# 集計結果 4 - ②

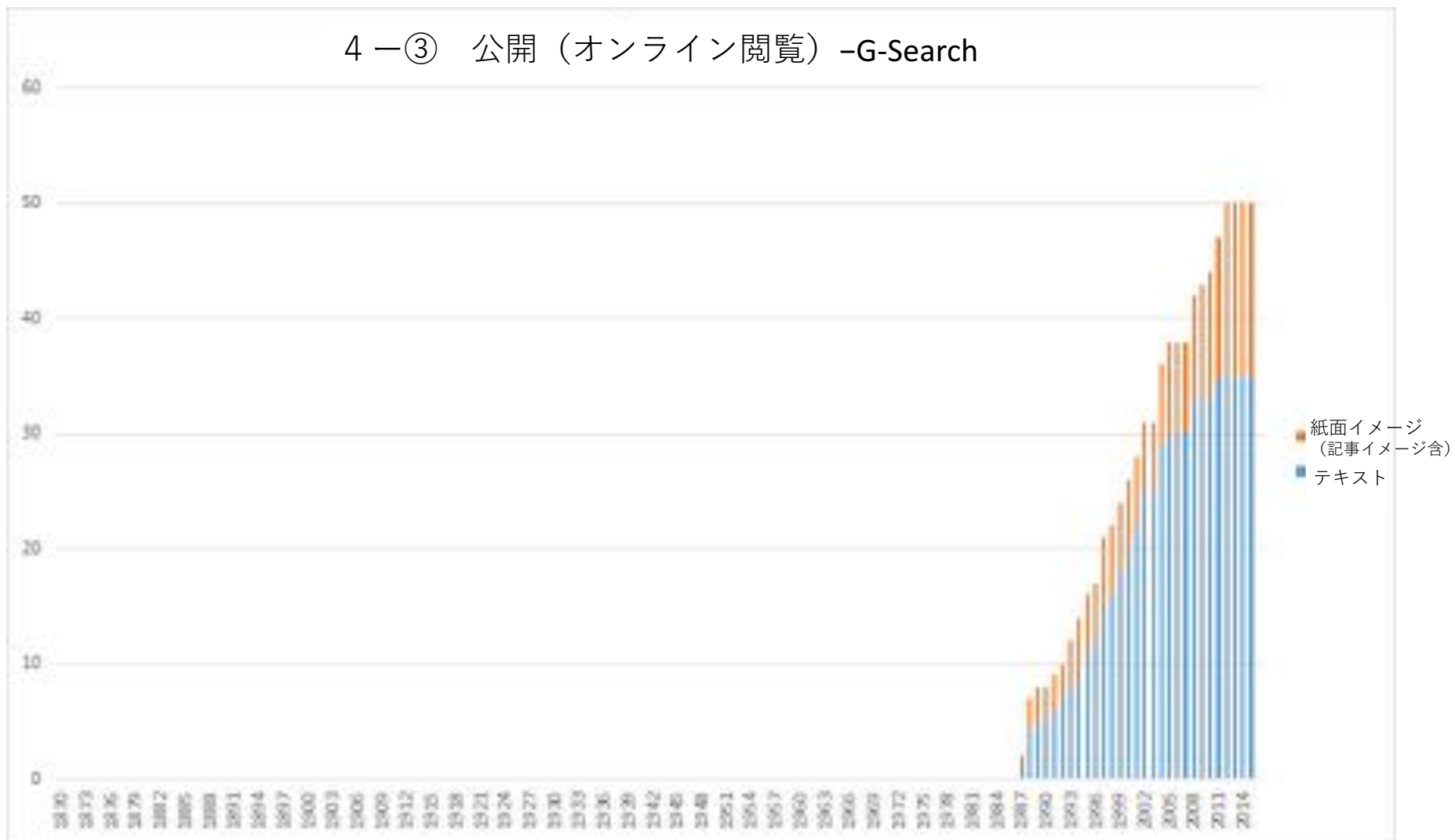
4 - ② 公開（オンライン閲覧）-nifty



# 集計結果 4ー③

4ー③ 公開（オンライン閲覧）-G-Search

社数

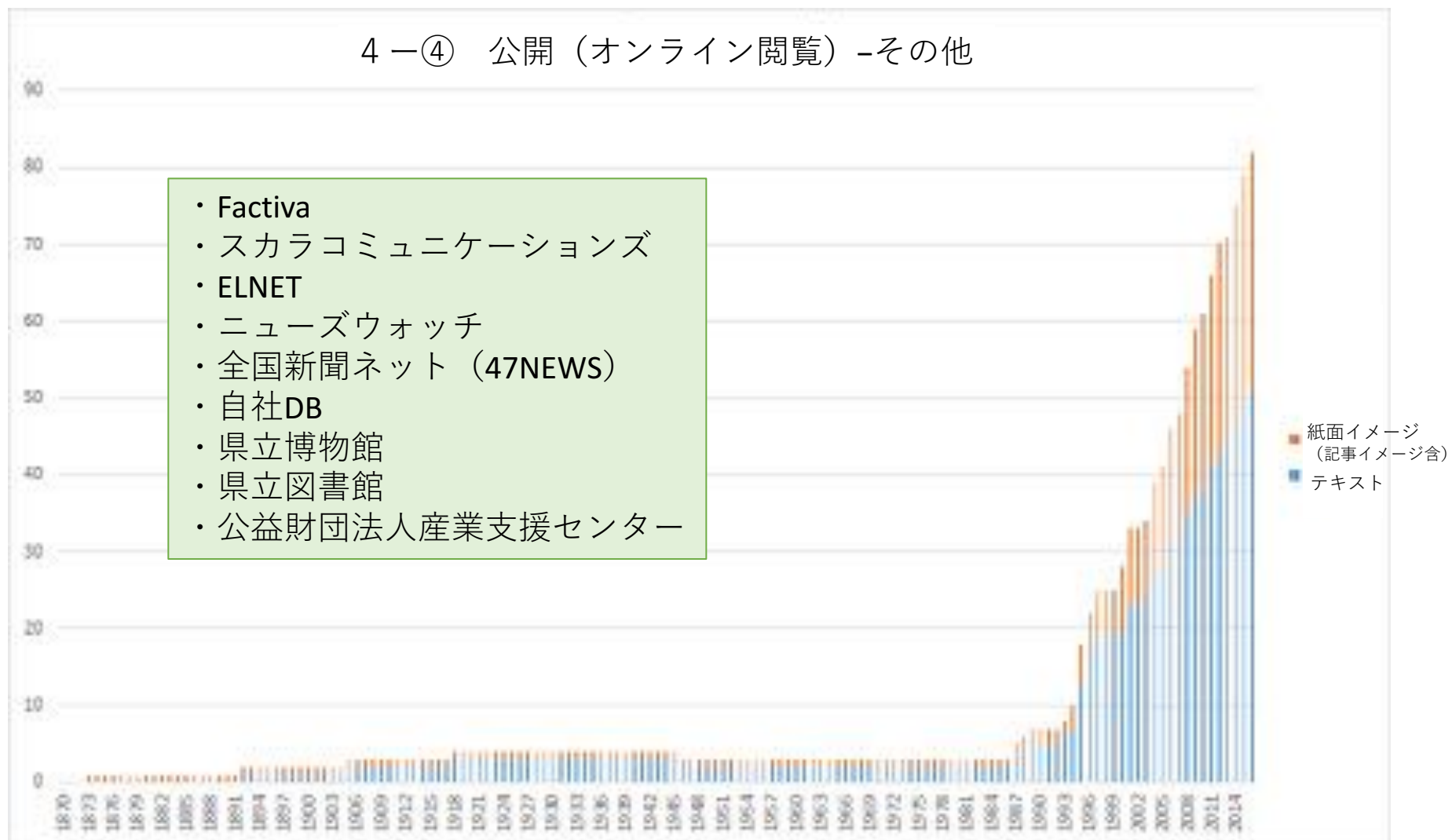


西暦

# 集計結果 4 — ④

4 — ④ 公開 (オンライン閲覧) - その他

社数

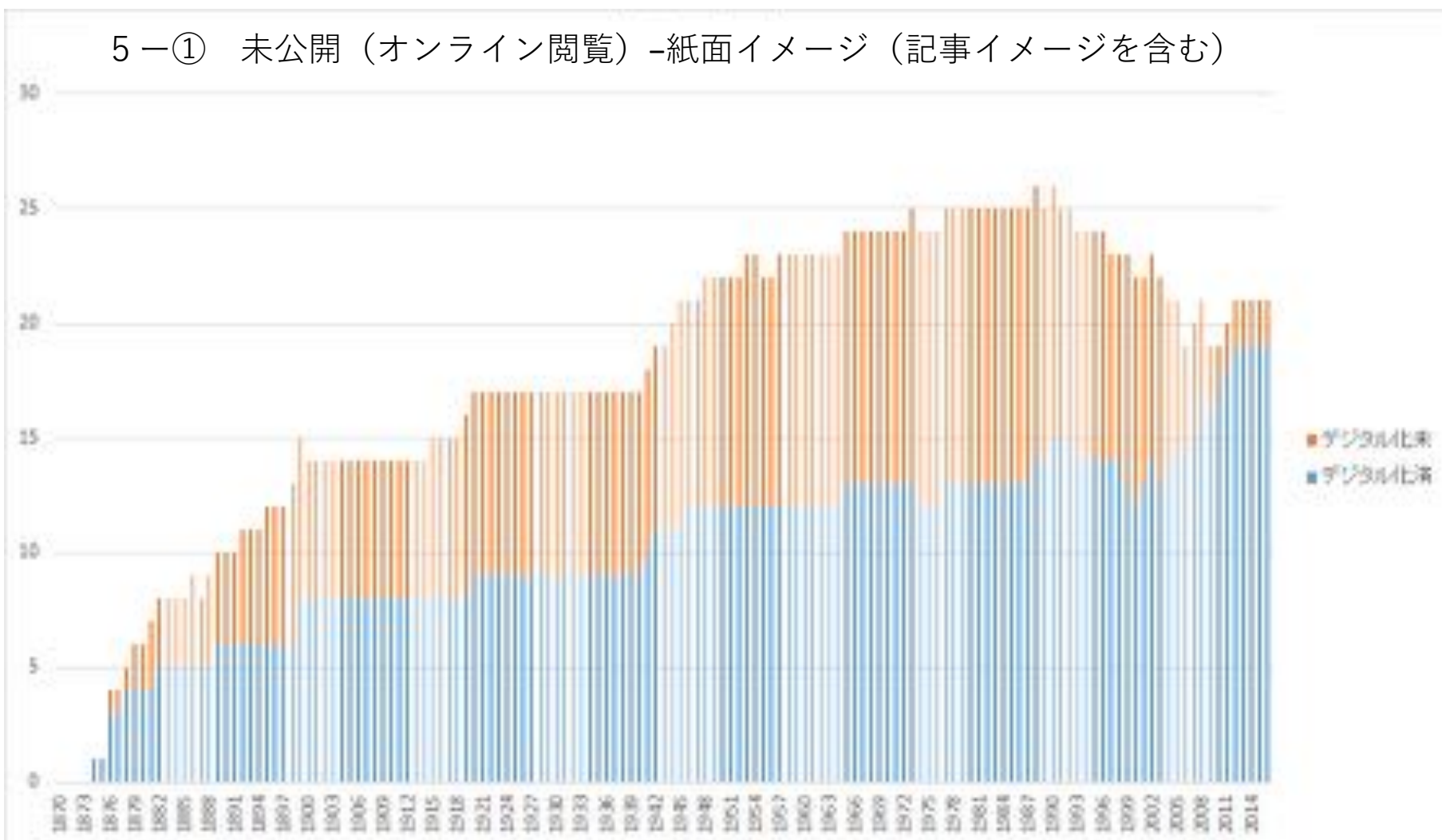


西暦

# 集計結果 5ー①

5ー① 未公開（オンライン閲覧）-紙面イメージ（記事イメージを含む）

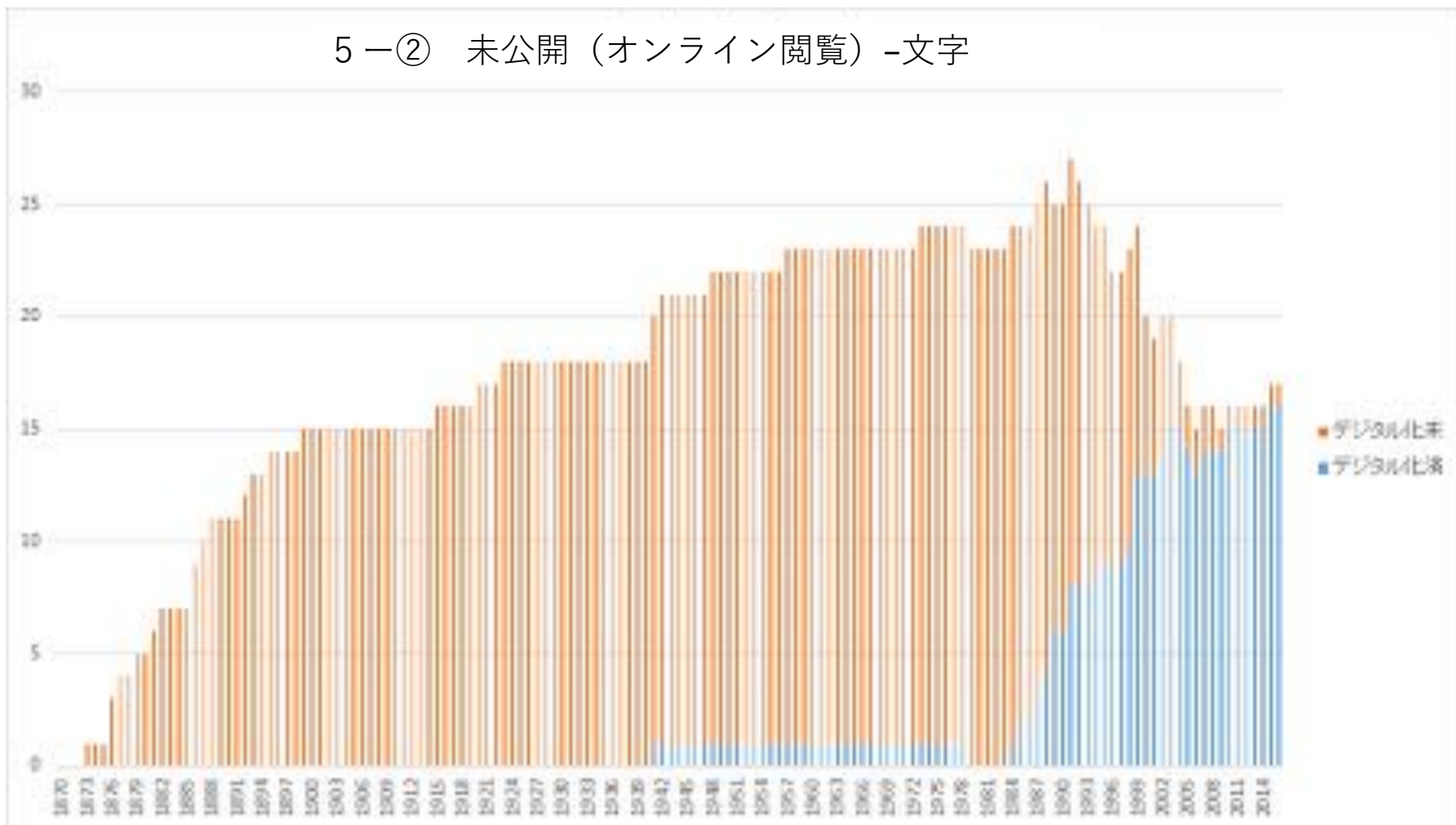
社数



西暦

# 集計結果 5－②

5－② 未公開（オンライン閲覧）-文字



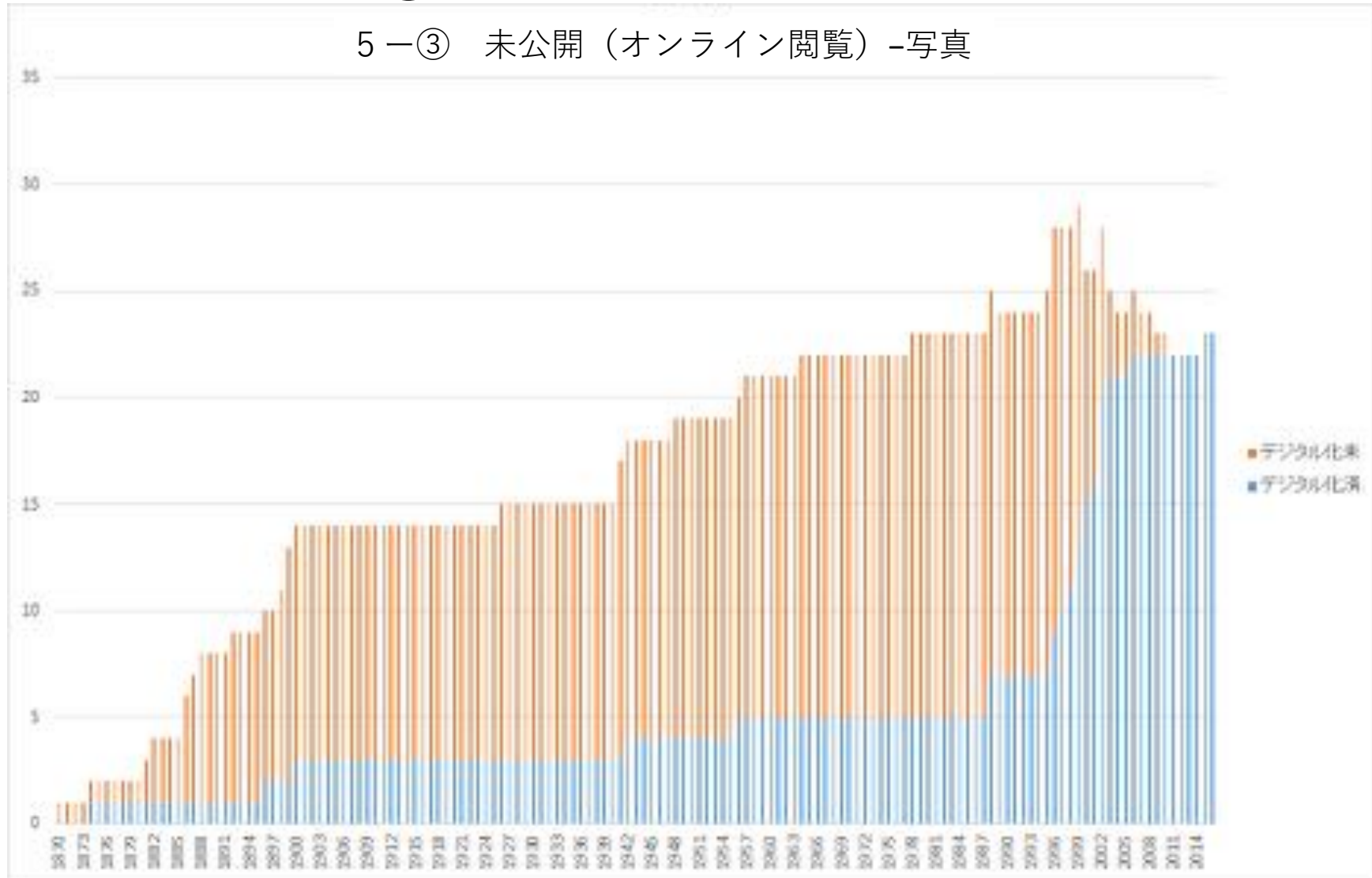
社数

西暦

# 集計結果 5ー③

5ー③ 未公開（オンライン閲覧）-写真

社数



西暦

# フリーアンサーまとめ

## ①権利・人権

- 人権に絡む記事の場合、非公開。
- 事件や事故など、人権に関わるものや弊社に著作権がない寄稿・投稿は見出しのみ公開。
- 著作権者からの了解が得られたもの、自社に著作権のある記事、写真を公開している。
- 記事ごとに人権のチェックを行い、配慮すべきものは掲載後5年で削除。

## ②外部公開

- 公開しているデータはテキスト全文と切り抜きイメージのPDFで、写真と紙面イメージは非公開。
- 切り抜きPDFは写真も公開している。
- 最終的に有料電子版ユーザー向けに公開の可能性。別途、料金設定してビジネス向けに公開する可能性もあり。
- 全て有料で提供。
- 公開基準等はまだ定めていない。
- 契約別で記事テキストの公開基準を設けている。外部機関へは紙面イメージを提供していない。
- 写真はサムネイル画像程度は公開。
- 広告は依頼があれば、マスキングをかける。
- 自社取材のテキストデータと写真データはほとんど自社ニュースサイト上に公開している。
- テキストがない主な記事も見出しを登録し紙面で読めるようにしている。紙面は全面公開。
- 明治昭和の紙面は、表現やプライバシーへの配慮で調査研究に限って提供しており、大学や図書館などに限定している。
- 社内利用のみで公開。マイクロフィルムは全面公開。

# フリーアンサーまとめ

## ③外部公開（非公開）

- 事件・事故、当社に著作権がない記事・図表は非公開
- 事件・事故に関する場合は2014年7月までは非公開、それ以降は匿名化して公開しています。
- 記事内容によっては非公開。
- 著作権のないもの（共同記事、寄稿など）は見出しのみ。
- 非公開の写真データは顔写真や戦前戦後の風景写真など1部のみ。
- 切抜きをOCR処理した未校閲の生のテキストデータは非公開です。

## ④体制

- 公開の体制が整っていない。

## ⑤データ提供

- ネット公開の記事のみ外部DBに提供
- 事件・事故ものなど人権にかかわるものや、著作権・肖像権など権利の問題がある記事は提供していません。写真は画質の良くないものを表示させています。



# フリーアンサーまとめ

## ⑥保存

- 原紙：製本にして保存（欠落あり）
- 記事イメージデータは戦災や震災で明治・大正・昭和初期の紙面がかなり欠落。
- 明治後半からの写真が残っています。
- 欠落が多い。古いものはもちろんだが、昭和40年以降のマイクロフィルムにも欠落コマが見つかる。

## ⑦デジタル化

- 写真データについては、年次には関係なく、重要な写真から順次デジタル化している
- 未処理写真が多数あり、詳細把握できていません。
- 写真は分類区分で入力されているため、未処理分が多数あります。
- デジタル化にあたり記事のジャンルで入力順を決め、重要度（続報有無、見出しサイズ）で選別しています。
- 外部公開に適さない事件・事故記事を後回しにしました。
- 紙焼き写真のデジタル化は終わりましたが公開していません。
- 過去写真のデータ化、過去記事のテキスト化作業中。
- 所蔵する紙面マイクロフィルムのデジタル化は完了している。2008年2月以降は個々の記事の「切り抜きイメージ」もある。
- 原紙は製本で残っているが、これをバラして再撮影するかどうか。よい方策がないか検討中。
- 東大が保有している欠落部分のデジタル化を行いたい。

# 課題のまとめ

1. 保存（消失した資料の補充）
2. 紙面・記事・写真のイメージデータ化
3. 記事のテキスト化
4. デジタルデータの公開・維持体制（著作権・肖像権・人権等への配慮）
5. 採算性