

# 情報漏洩インシデント調査に基づく 漏洩原因のデータマイニング

明治大学総合数理学部

池上和輝 菊池浩明

# 研究背景

- 業務の電子化による情報漏洩の増加
- 漏洩インシデントの特徴が不明確で対応が困難
- 日本ネットワークセキュリティ協会(JNSA)による、漏洩インシデントの収集や被害額算出式の提案(JOモデル)
  - JOモデル：漏洩した要素や企業の規模、事後対応などから被害額算出



# 問題点

1. 漏洩原因と漏洩要素の関係が示されておらず、インシデントの特徴が不明確
2. JNSAデータでは、事後対応が普通(1)、悪い(2)の2値で主観的である

企業名	件数	原因	漏洩要素			事後対応度
			氏名	住所	メールアドレス	
マイクロソフト	559	誤操作			○	悪い
日本年金機構	1014653	不正アクセス	○	○		普通
ベネッセ	48580000	内部犯罪・不正	○	○		普通

# 研究目的・方法

- 独自のデータセットの作成（解決方法）

- 朝日新聞記事検索システム「聞蔵II」使用

- 原因と漏洩要素の関係を明らかにする（問題1）

- Rパッケージ“arules”を使用し連関規則を抽出

- 事後対応を客観的に分析する（問題2）

- 事後対応の分析
  - 事件発生から報道までの日数を収集



The screenshot shows the 'Monzou II' search interface. At the top, there are navigation links for '朝日新聞' (Asahi Shimbun), '聞蔵IIビジュアル' (Monzou II Visual), and 'English'. Below this, there are search filters for '朝日新聞 1945～' and '朝日新聞検索履歴 [5/17～17/17]'. The main search results area shows a table of articles. The table has columns for 'No.', '発行日' (Issue Date), '期次' (Issue), '題名' (Title), 'ページ' (Page), '文字数' (Character Count), '写真関係' (Photo Related), and '切り抜き' (Clipping). The first row shows an article from 2015年12月31日 (December 31, 2015) with the title '中顔カメラに防犯任せて 飯沼村 / 福島県' (Entrusted with security cameras, Ibanuma Village / Fukushima Prefecture). The second row shows an article from 2015年12月30日 (December 30, 2015) with the title '動いた歴史、刻まれた記憶 2015 プレーバック 1月～9月 = 訂正・お詫びあり' (History that moved, memories that were etched 2015 Playback January to September = Correction/Apology available). The third row shows an article from 2015年12月30日 (December 30, 2015) with the title '富森ガイド / 東京都' (Tomoson Guide / Tokyo).

No.	発行日	期次	題名	ページ	文字数	写真関係	切り抜き
03001	2015年12月31日	朝刊	福島県・1地方 中顔カメラに防犯任せて 飯沼村 / 福島県	015	00333文字	あり	<input type="checkbox"/>
03002	2015年12月30日	朝刊	復興集人 動いた歴史、刻まれた記憶 2015 プレーバック 1月～9月 = 訂正・お詫びあり	012	06459文字	あり	<input type="checkbox"/>
03003	2015年12月30日	朝刊	東京都・1地方 富森ガイド / 東京都	021	04452文字	あり	<input type="checkbox"/>

# 研究目的・方法

- 独自のデータセットの作成

- 朝日新聞記事検索システム「聞蔵II」使用

- 原因と漏洩要素の関係を明らかにする（問題1）

- Rパッケージ“arules”を使用し連関規則を抽出

- 事後対応を客観的に分析する（問題2）

- 事後対応の分析
  - 事件発生から報道までの日数を分析

The screenshot shows the '聞蔵IIビジュアル' search results page. The search criteria are '朝日新聞 1945~' and '朝日新聞掲載誌 15/1~17/12'. The results show a total of 1297 items, with the first 20 items displayed. The table below is a simplified version of the visible data.

No.	発行日	期次	題名	ページ	文字数	写真関係	切り抜き
03001	2015年12月31日	朝刊	福島県・1地方 中顔オメラに防犯任せて 飯沼村 / 福島県	015	00333文字	あり	<input type="checkbox"/>
	2015年12月31日	朝刊	東北編入 中顔オメラに防犯任せて 飯沼村 / 福島県	012	06459文字	あり	<input type="checkbox"/>
03002			動いた歴史、知られた記憶 2015 プレーバック 1月～9月 = 訂正・お詫びあり				<input type="checkbox"/>
	2015年12月31日	朝刊	東北編入・1地方 動いた歴史、知られた記憶 2015 プレーバック	021	04452文字	あり	<input type="checkbox"/>
03005			飯沼ガイド / 飯沼村				<input type="checkbox"/>

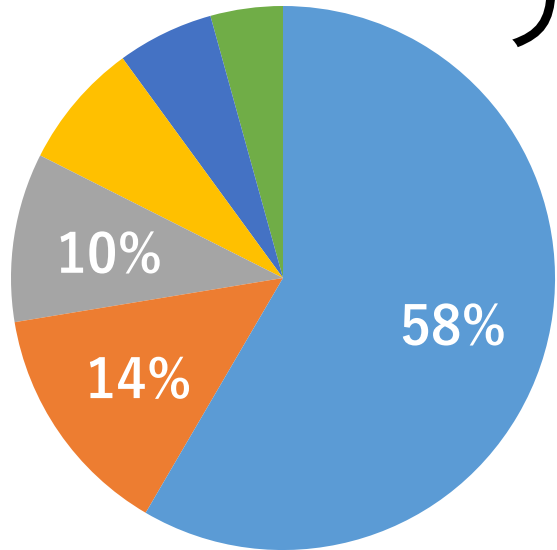
# データセットの作成

2015年データセットの比較(件)

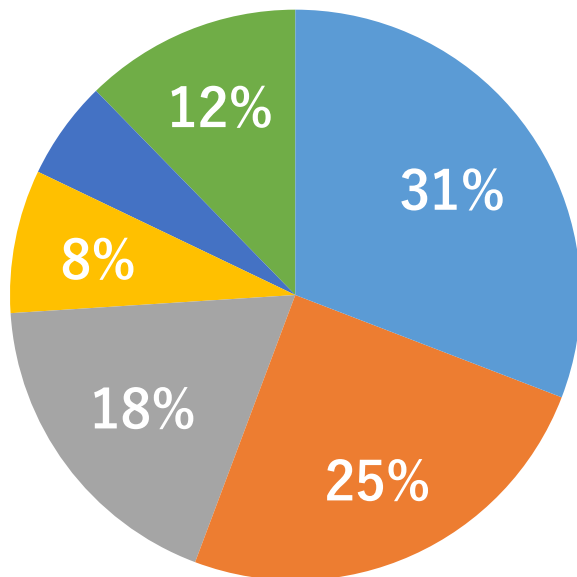
JNSA	本調査	共通
788	279	145

項目	JNSA	本調査	共通事例
公開日	2015/10/6	2015/9/7	2015/6/1
発生日		2015/9/2	2015/5/8
企業名	ゴールドボンド	福岡県	日本年金機構
業種	サービス業	公務	公務
持ち会社/親会社/関連企業	広島県		
管理委託先	委託先あり		
漏洩件数	77	35	1160000
漏洩原因	誤操作	紛失	不正アクセス
漏洩経路	電子メール	紙	インターネット
漏洩要素	氏名/電話番号	氏名/住所	氏名/ID
社内規則違反		0	1
被害の可能性		0	1
報道までの日数		5日	24日
社会的責任度など	有		有

# データセットの統計と比較



本調査



JNSA

	JNSA	本調査
紛失・置き忘れ	243	163
誤操作	196	39
管理ミス	144	28
不正アクセス	64	21
盗難	44	16
その他	97	12

(参考) JNSA情報セキュリティインシデントに関する調査報告書別紙

# 研究目的・方法

- 独自のデータセットの作成

- 朝日新聞記事検索システム「聞蔵II」使用

- 原因と漏洩要素の関係を明らかにする（問題1）

- Rパッケージ“arules”を使用し連関規則を抽出

- 事後対応を客観的に分析する（問題2）

- 事後対応の分析
  - 事件発生から報道までの日数を分析

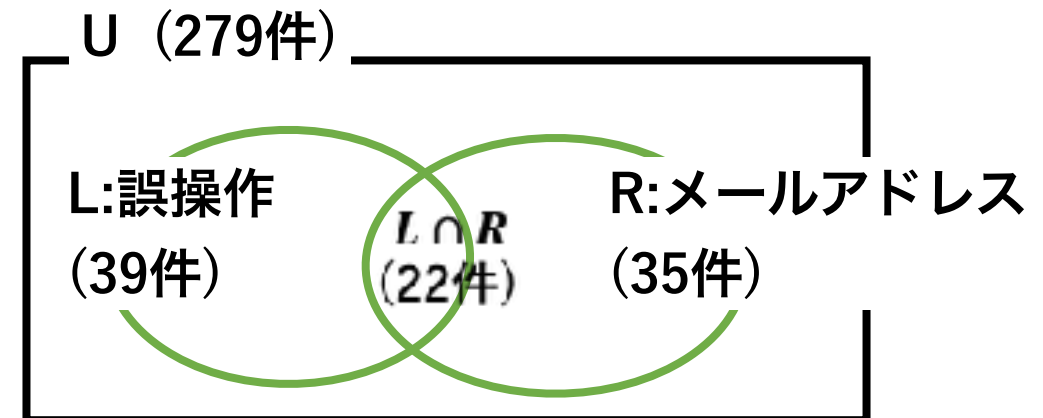
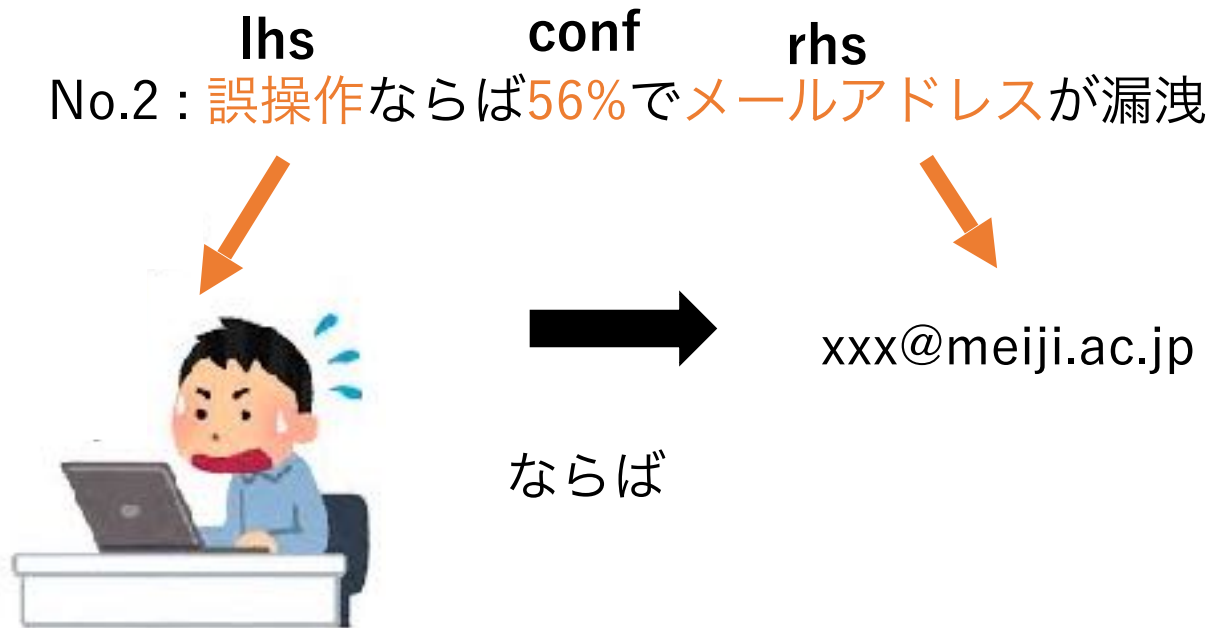
The screenshot shows the 'Monzou II' search interface. At the top, there are navigation links for 'English', 'お問い合わせ', '利用規定', and 'ログイン'. Below that, there are tabs for '朝日新聞 1945～', '朝日新聞経緯版 1877～1999', '記事集', '人物', '歴史写真', 'アフィリエイト', and '英文ニュース'. The main search area includes a search bar and buttons for '検索', 'クリア', and 'リセット'. Below the search bar, there is a message: '※グリーンで表示された記事は著作権などの関係で本文を表示できません。' and search statistics: '総件数: 1297件 通し番号: 1～20'. There are also buttons for '全選択', '全解除', and '+x件追加'. The search results are displayed in a table with columns: No., 発行日, 朝刊/夕刊, 題名, ページ, 文字数, 写真あり, and 切り抜き. The table contains three rows of results.

No.	発行日	朝刊/夕刊	題名	ページ	文字数	写真あり	切り抜き
01001	2015年:2月31日	朝刊	福島県会 - 1地方	0:5	30333文字	あり	
01002	2015年:2月33日	朝刊	東北信人	0:2	36459文字	あり	
01003	2015年:2月31日	朝刊	東京経済 - 1地方	021	32452文字	あり	



# 漏洩原因と漏洩要素の連関規則

No	lhs	rhs	support	confidence	lift	件数	例
1	紛失・置忘れ	氏名	0.557	0.957	1.117	156	タカラトミー
2	誤操作	メールアドレス	0.079	0.564	4.036	22	愛媛県
3	管理ミス	氏名	0.089	0.929	1.080	25	長崎大学病院
4	管理ミス	住所	0.057	0.571	1.317	16	静岡ガス
5	不正アクセス	クレジット情報	0.014	0.190	13.334	4	日本年金機構
6	不正アクセス	ID/パスワード	0.014	0.190	7.619	4	新日本プロレス



$$\text{Supp} = \frac{|L \cap R|}{|U|}$$

$$\text{Conf} = \frac{|L \cap R|}{|L|} \quad 9$$

# 研究目的・方法

## ・独自のデータセットの作成

- ・朝日新聞記事検索システム「聞蔵II」使用

## ・原因と漏洩要素の関係を明らかにする（問題1）

- ・Rパッケージ”arules”を使用し連関規則を抽出

## ・事後対応を客観的に分析する（問題2）

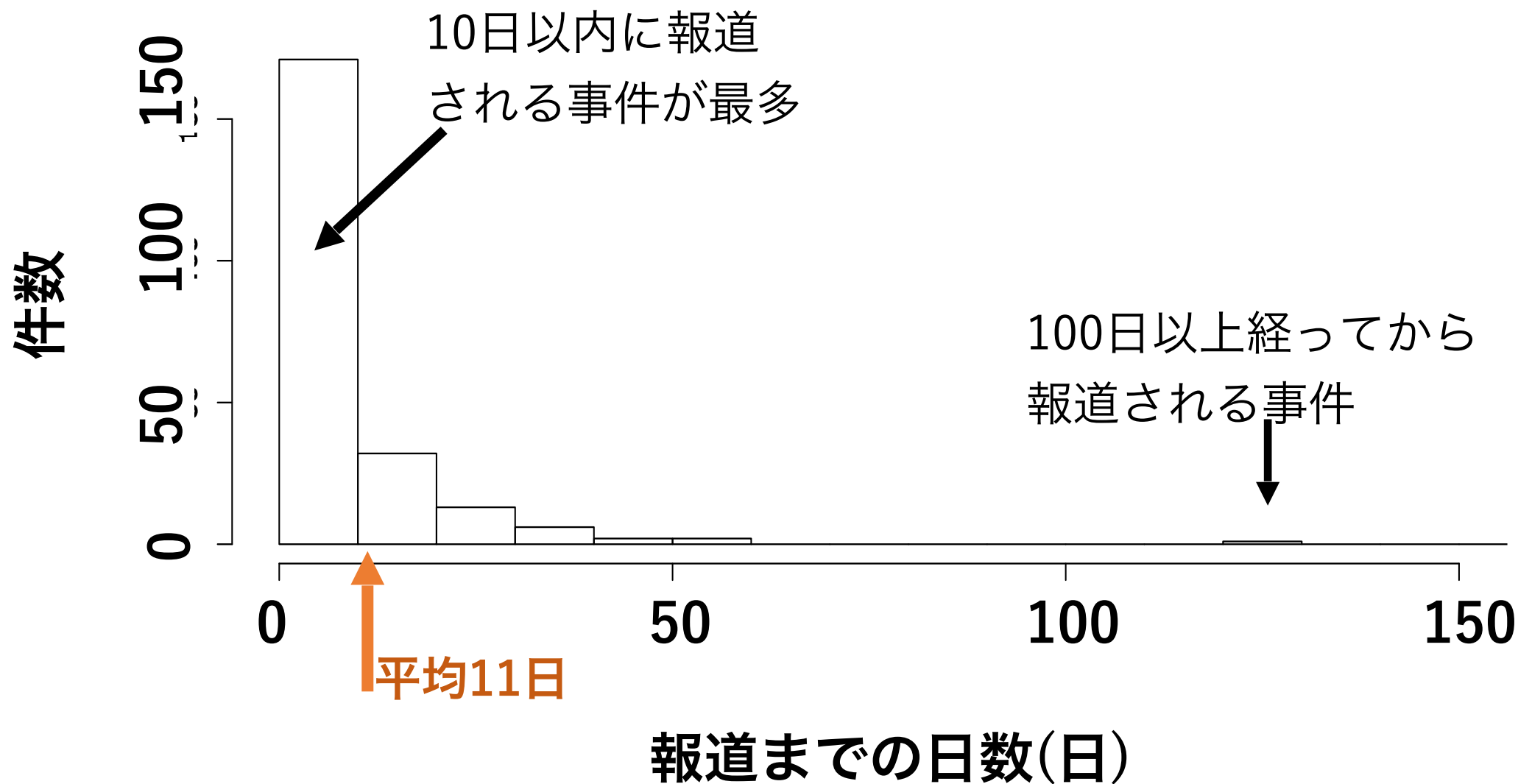
- ・事後対応の分析
  - ・事件発生から報道までの日数を分析



The screenshot shows the search results page of the Asahi Shimbun's 'Monzou II' system. The search criteria are set to '朝日新聞 1945～' and '週刊朝日・A社'. The results are sorted by '最新' (latest) and show a total of 1297 articles. The table below lists the first few results:

No.	発行日	版	面名	ページ	文字数	写真数	切り抜き
03001	2015年12月31日	朝刊	福島県・1地方	015	30333文字	あり	<input type="checkbox"/>
	中顔メラに防犯任せて 飯沼村 / 福島県						
	2015年12月30日	朝刊	奥付集人	012	36459文字	あり	<input type="checkbox"/>
03002	働いた歴史、刻まれた記憶 2015 プレーバック 1月～9月＝訂正・お詫びあり						
	2015年12月30日	朝刊	東京経済・1地方	021	32451文字	あり	<input type="checkbox"/>
03003	新卒ガイド / 東京版						

# 報道までの日数のヒストグラム

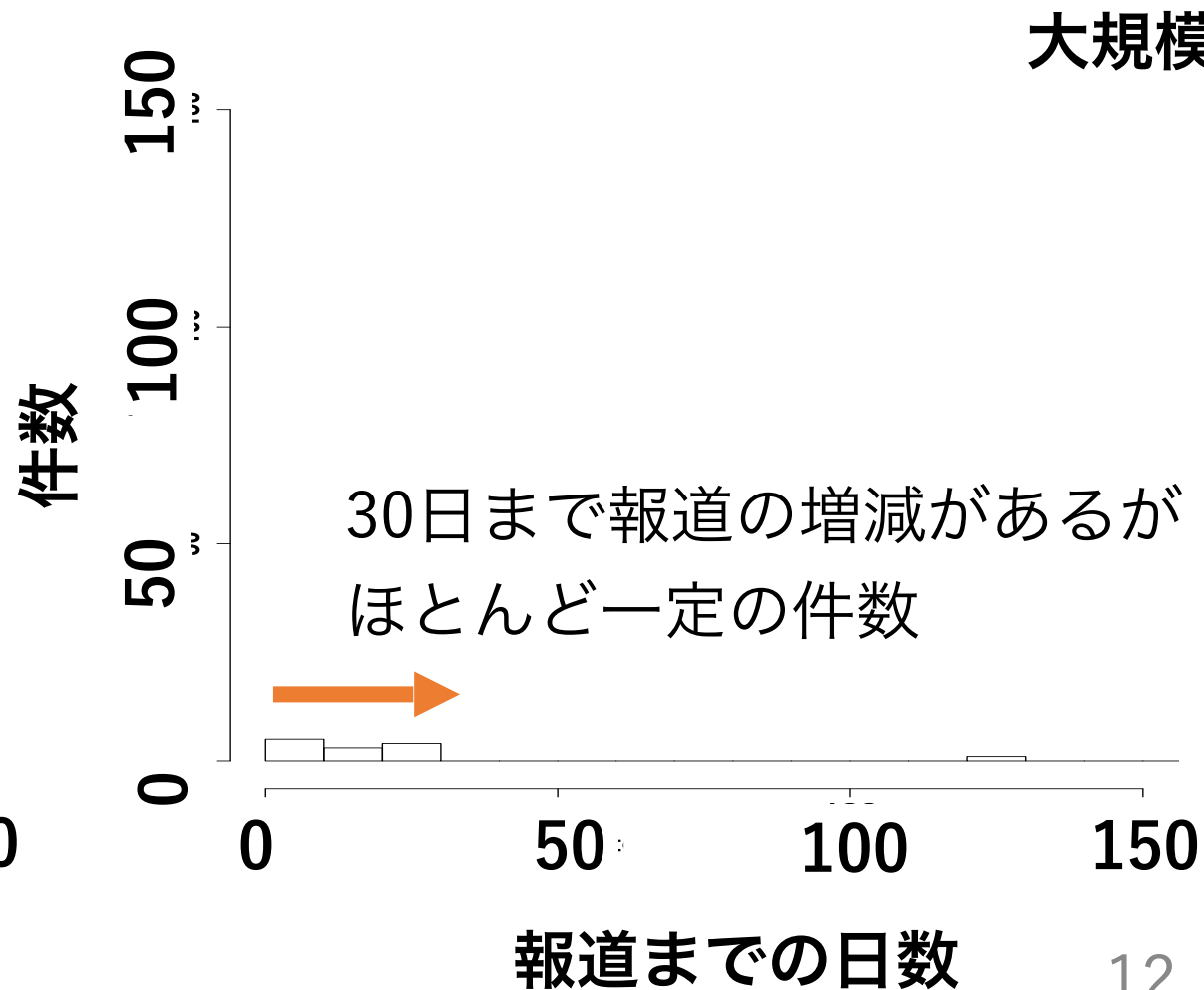
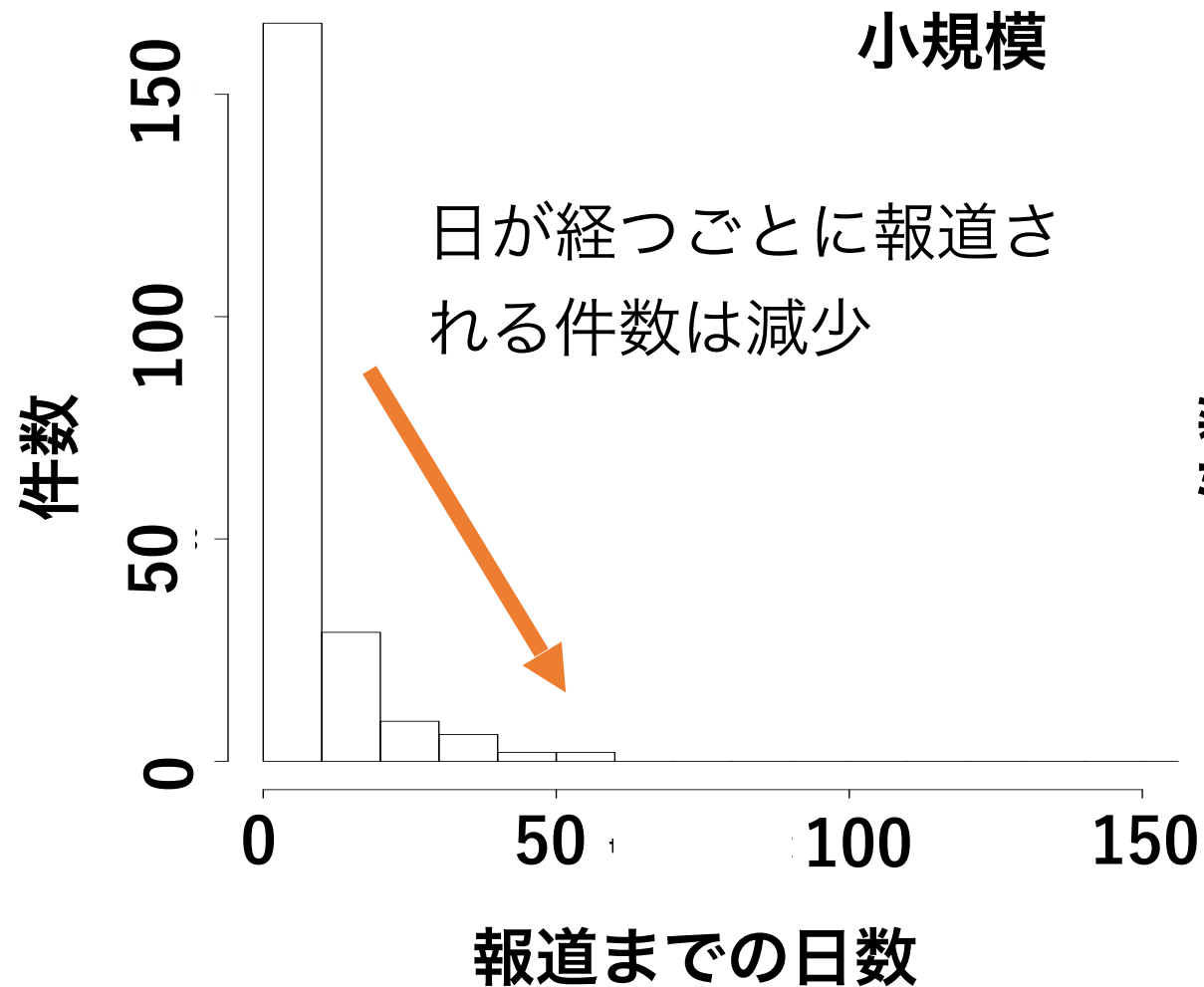


# 規模別報道までの日数

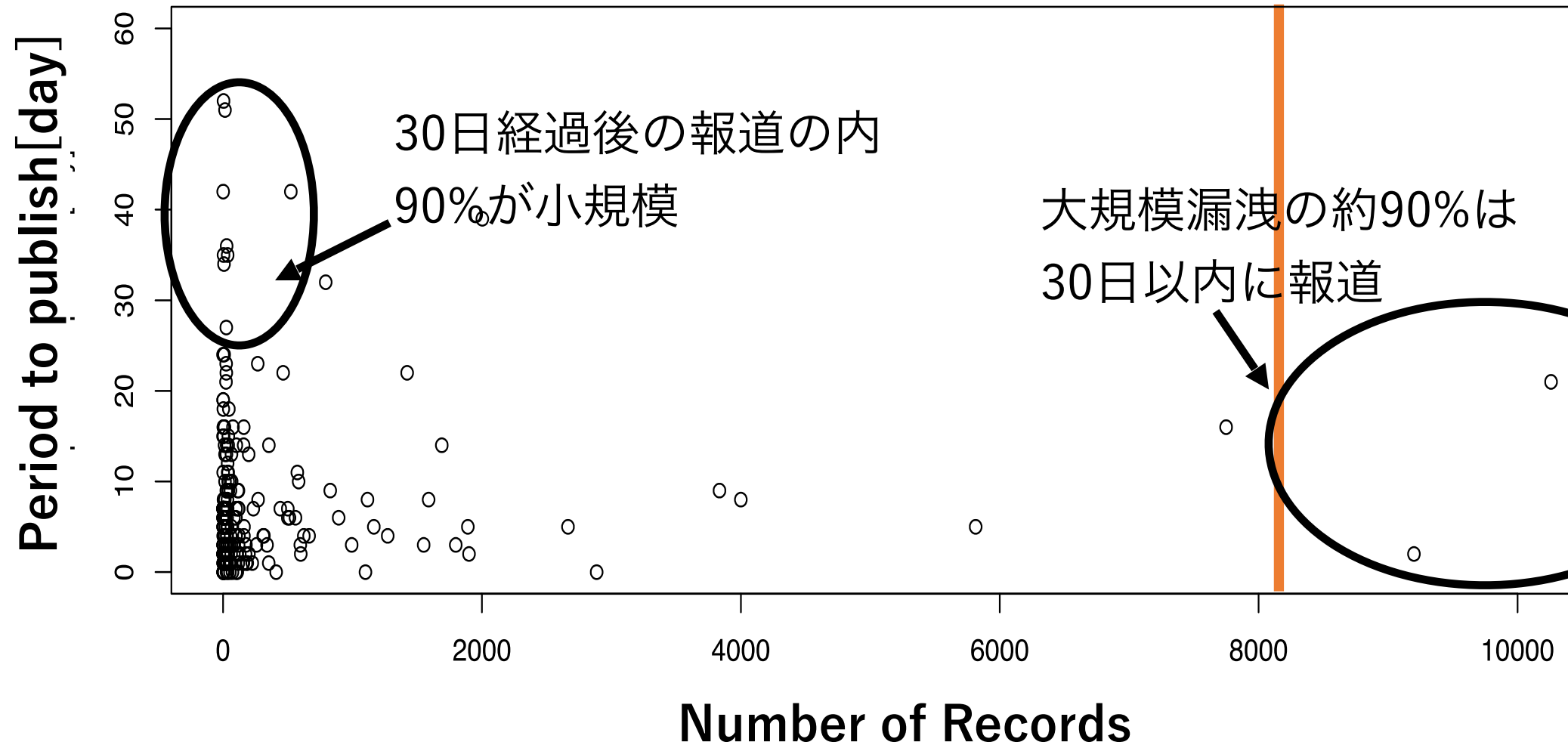
平均漏洩件数:8104件

小規模

大規模



# 報道までの日数と漏洩規模の散布図



# 結論

- 漏洩原因と漏洩要素の特徴を明らかにした（問題1）
  - 紛失/置忘れは約95%で氏名漏洩
  - 誤操作は約56%でメールアドレス漏洩
- 事後対応の分析(問題2)
  - 30日以上かかるインシデントの内、約90%は小規模な漏洩
  - 大規模な漏洩の約90%は30日以内に報道