

Web アーカイブの現状と課題

06L4011 杉浦 友哉

1. Web アーカイブとは

Web アーカイブとは Web 上のホームページを収集して保存・公開するもので、変化が激しくすぐに改廃される Web ページを記録して将来に残そうというものである。

2. Web アーカイブの技術と手法、法的課題

- (1) 収集方法: クローラを使う方法と手作業
- (2) 収集範囲: 選択的収集と網羅的収集
- (3) 表層 Web と深層 Web
- (4) 公開上の問題: ファイル形式の変化に対応した閲覧性の確保
- (5) 著作権の問題: 著作権法とプロバイダ責任制限法と個人情報の保護

3. Web アーカイブ活動の現状

- (1) IIPC: 2003 年欧米の国立図書館を中心とした 12 機関で発足した組織
- (2) Internet Archive: バルク式収集方法で Web アーカイブを構築
- (3) 各国のプロジェクト: MINERVA、PANDORA、netarchive.dk など
- (4) 国立国会図書館 WARP: 国立国会図書館が行っている Web アーカイブ事業

3. 調査

Internet Archive により各企業・組織のホームページの変遷を分析した。

(1) 本田技研工業株式会社

	ニュース		製品とサービス					企業活動		会社情報			
	トピックス		製品とサービス					企業活動	会社情報				
1997/4/15	○		○					○	○				
	更新情報	広報発表	クルマ	バイク	汎用製品	お客様窓口		モーター スポーツ		投資家 情報	HONDA について	テクノ ロジー	
2005/1/1	○	○	○	○	○	○		○ ○	○	○	○	○	○
							リコー ール	CSR					
2009/12/12	○	○	○	○	○	○	○	○	○	○	○	○	○

ニュース、製品とサービス、企業活動、会社情報というスタイルはほとんど変化ない。

(2) 愛知県警察本部

	ニュース			県警について						交通		犯罪対策	
	新着 情報			県警察案内		法律 改正					交通 事故		
1999/2/19	○			○		○					○		
				愛知県警	警察 署						交通 安全	犯罪対 策	ハイ テク
2000/8/18	○			○	○	○				○		○	○
		県警	捜査にご 協力を			法律 施行	統計	110 番	プレゼント コーナー			安全な 暮らし	
2009/12/12	○	○	○	○	○	○	○	○	○	○		○	○

2000 年以降「交通安全」、「犯罪対策」のページが充実した。

(3) 西日本旅客鉄道株式会社

	お詫び	お知らせ		乗客向け情報		鉄道の話題		企業情報					
		Hot News			トラベルブ ラザ	鉄道ひ ろば						グルー プ情報	
1998/12/12		○			○	○							○
	お詫び	プレスリ リース	お知 らせ	運行 情報	お出かけネ ット		鉄道ギ ャラリ	IR 情報	法人 取引	企業 案内	支店 情報	グルー プ会社	
2005/4/28	○	○	○	○	○		○	○	○	○	○	○	
			お知らせ	ご報 告			鉄道フ ァン						関連事 業
2009/12/12	○	○	○		○	○		○		○			○

2005 年の福知山線の脱線事故以降、トップページの中央が「おわびの言葉」となっている。

5. 考察

Internet Archive を用いることにより、各社のホームページのコンセプトの変化、画面の見せ方の変化等が解析できる。このように Web アーカイブはインターネットの歴史を解析する上で有効なツールと思われる。