

# AIDAinformazioni

Rivista semestrale di Scienze dell'Informazione

NUMERO 1-2 / 2021

ANNO 39

GENNAIO-GIUGNO 2021

## Editoriale

FABRIZIA FLAVIA SERNA

## Contributi

STEFANO ALLEGREZZA

*La valutazione di interoperabilità dei formati elettronici nelle nuove Linee guida sulla formazione, gestione e conservazione dei documenti informatici*

SOFIA BARONCINI, MARILENA DAQUINO, FRANCESCA TOMASI

*Modelling art interpretation and meaning. A data model for describing iconology and iconography*

LUCA BERTOLONI

*Per una riflessione intorno al valore documentario della forma-canzone nell'attuale mediascape. L'Italia che canta in pandemia*

CLAUDIA CANTALE

*Le recensioni di TripAdvisor come strumento di analisi. Il caso delle Biblioteche Riunite "Civica e Ursino Recupero"*

ANNA IRENE CESARANO

*Il pensiero apocalittico tra teorie e riflessioni*

FABIO CUSIMANO

*Small Books, Small Screens. La digitalizzazione di antichi manoscritti di piccolo formato, dallo scaffale ai dispositivi mobili*

SALVATORE SPINA, GIUSEPPE SORBELLO

*Dagli archivi storici alle mappe genomiche. Il caso di Riposto*

MARC TANTI, HUGO MOQUET

*Identification sur Twitter des acteurs communiquant autour de la COVID-19 au début de l'épidémie. Analyse des messages et des rumeurs diffusés*

## Note e Rubriche

CLAUDIO GNOLI

*Organizzazione della conoscenza. Come mi vuoi, materiale o ideale?*

AIDAinformazioni N. 1-2 - ANNO 39 - GENNAIO-GIUGNO 2021

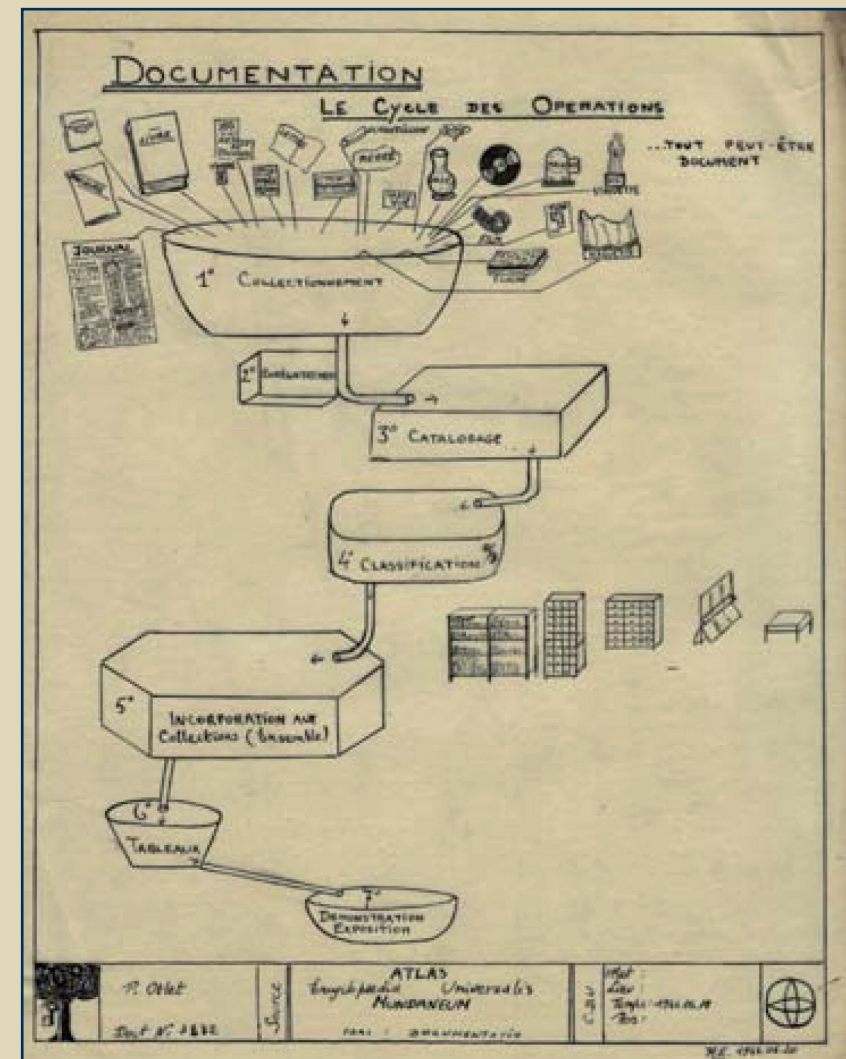
# AIDAinformazioni

RIVISTA SEMESTRALE DI SCIENZE DELL'INFORMAZIONE

NUMERO 1-2 / 2021

ANNO 39

GENNAIO-GIUGNO 2021



ARACNE



mundaneum

In copertina  
Disegno di Paul Otlet, Collections Mundaneum, centre d'Archives, Mons (Belgique).

ISSN 1121-0095

euro 20,00

ISBN 979-12-5994-484-9



9 791259 944849

ARACNE

# AIDAinformazioni

RIVISTA SEMESTRALE DI SCIENZE DELL'INFORMAZIONE

Fondata nel 1983 da Paolo BISOGNO

N. 1-2 — Anno 39 — gennaio-giugno 2021

Proprietario della rivista  
UNIVERSITÀ DELLA CALABRIA

Direttore Scientifico  
Roberto GUARASCI  
Università della Calabria

Direttore Responsabile  
Fabrizia Flavia SERNIA

## *Comitato scientifico*

Anna Rovella, UNIVERSITÀ DELLA CALABRIA  
Maria Guercio, SAPIENZA UNIVERSITÀ DI ROMA  
Giovanni Adamo, CONSIGLIO NAZIONALE DELLE RICERCHE  
Claudio Gnoli, UNIVERSITÀ DEGLI STUDI DI PAVIA  
Ferruccio Diozzi, CENTRO ITALIANO RICERCHE AEROSPAZIALI  
Gino Roncaglia, UNIVERSITÀ DELLA TUSCIA  
Laurence Favier, UNIVERSITÉ CHARLES-DE-GAULLE LILLE 3  
Madjid Ihadjadene, UNIVERSITÉ VINCENNES-SAINT-DÉNIS PARIS 8  
Maria Mirabelli, UNIVERSITÀ DELLA CALABRIA  
Agustín Vivas Moreno, UNIVERSIDAD DE EXTREMADURA  
Douglas Tudhope, UNIVERSITY OF SOUTH WALES  
Christian Galinski, INTERNATIONAL INFORMATION CENTRE FOR TERMINOLOGY  
Béatrice Daille, UNIVERSITÉ DE NANTES  
Alexander Murzaku, COLLEGE OF SAINT ELIZABETH, USA

## *Comitato di redazione*

Antonietta Folino, UNIVERSITÀ DELLA CALABRIA  
Erika Pasceri, UNIVERSITÀ DELLA CALABRIA  
Maria Taverniti, CONSIGLIO NAZIONALE DELLE RICERCHE  
Maria Teresa Chiaravallotti, CONSIGLIO NAZIONALE DELLE RICERCHE  
Assunta Caruso, UNIVERSITÀ DELLA CALABRIA

## *Segreteria di Redazione*

Valeria Rovella, UNIVERSITÀ DELLA CALABRIA

ARACNE

# AIDAinformazioni

RIVISTA SEMESTRALE

«AIDAinformazioni» è una rivista scientifica che pubblica articoli inerenti le Scienze dell'Informazione, la Documentazione, la Gestione Documentale e l'Organizzazione della Conoscenza. È stata fondata nel 1983 quale rivista ufficiale dell'Associazione Italiana di Documentazione Avanzata e nel febbraio 2014 è stata acquisita dal Laboratorio di Documentazione dell'Università della Calabria.

La rivista si propone di promuovere studi interdisciplinari oltre che la cooperazione e il dialogo tra profili professionali aventi competenze diverse, ma interdipendenti. I contributi possono riguardare *topics* quali Documentazione, Scienze dell'informazione e della comunicazione, Scienze del testo e del documento, Organizzazione e Gestione della conoscenza, Terminologia, Statistica testuale e Linguistica computazionale e possono illustrare studi sperimentali in domini specialistici, casi di studio, aspetti e risultati metodologici conseguiti in attività di ricerca applicata, presentazioni dello stato dell'arte, ecc.

«AIDAinformazioni» è riconosciuta dall'ANVUR come rivista di Classe A per l'Area 11 – Settore 11/A4 e censita per le Aree 10 – Scienze dell'antichità, filologico-letterarie e storico-artistiche; 11 – Scienze storiche, filosofiche, pedagogiche e psicologiche; 12 – Scienze giuridiche; 14 – Scienze politiche e sociali, così come dall'AERES (Agence d'évaluation de la recherche et de l'enseignement supérieur) che la annovera tra le riviste scientifiche dell'ambito delle Scienze dell'Informazione e della Comunicazione. La rivista è, inoltre, indicizzata in: ACNP – Catalogo Italiano dei Periodici; BASE – Bielefeld Academic Search Engine; ERIH PLUS – European Reference Index for the Humanities and Social Sciences – EZB – Elektronische Zeitschriftenbibliothek – Universitätsbibliothek Regensburg; Gateway Bayern; KVK – Karlsruhe Virtual Catalog; Letteratura Professionale Italiana – Associazione Italiana Biblioteche; The Library Catalog of Georgetown University; SBN – Italian union catalogue; Summon™ – by SerialsSolutions; Ulrich's; UniCat – Union Catalogue of Belgian Libraries; Union Catalog of Canada; LIBRIS – Union Catalogue of Swedish Libraries; Worldcat.

I contributi sono valutati seguendo il sistema del *double blind peer review*: gli articoli ricevuti dal comitato scientifico sono inviati in forma anonima a due *referee*, selezionati sulla base della loro comprovata esperienza nei *topics* specifici del contributo in valutazione.

## Condizioni di acquisto

(spese di spedizione in Italia incluse)

Abbonamento annuale ..... euro 42,00

Fascicolo ..... euro 26,00

## Per ordini

telefono 0039 06 87646960

e-mail: [info@adiuwaresrl.it](mailto:info@adiuwaresrl.it)

## Modalità di pagamento

Bonifico bancario intestato a:

Adiuvare S.r.l.

IBAN: IT 57 B 07092 38900 000001004504

BIC / Codice Swift: CCRTIT2TCRT

Banca di Credito Cooperativo dei Castelli  
Romani

Corso Giuseppe Garibaldi, 9, 00072 Ariccia RM

Causale: Abbonamento «AIDAinformazioni»

## Editore

Adiuvare S.r.l.

via Colle Fiorito, 2 (p. 2, int. 6)

00045 Genzano di Roma

P.IVA 15662501004

<http://www.aracneeditrice.eu>

[info@adiuwaresrl.it](mailto:info@adiuwaresrl.it)

ISBN 979-12-5994-484-9

edizione a stampa ISSN 1121-0095

edizione elettronica ISSN 1594-2201

# Indice

## Editoriale

- 7 Nella pandemia, le canzoni un balsamo dell'anima.  
Hanno aiutato a reagire  
*Fabrizia Flavia Sernia*

## Contributi

- 13 La valutazione di interoperabilità dei formati elettronici nelle nuove  
Linee guida sulla formazione, gestione e conservazione dei documenti  
informatici  
*Stefano Allegrezza*
- 39 Modelling art interpretation and meaning. A data model for describing  
iconology and iconography  
*Sofia Baroncini, Marilena Daquino, Francesca Tomasi*
- 63 Per una riflessione intorno al valore documentario della forma-canzone  
nell'attuale *mediascape*. L'Italia che canta in pandemia  
*Luca Bertoloni*
- 89 Le recensioni di TripAdvisor come strumento di analisi. Il caso delle Bi-  
blioteche Riunite "Civica e Ursino Recupero"  
*Claudia Cantale*
- 113 Il pensiero apocalittico tra teorie e riflessioni  
*Anna Irene Cesarano*
- 145 *Small Books, Small Screens*. La digitalizzazione di antichi manoscritti di pic-  
colo formato, dallo scaffale ai dispositivi mobili  
*Fabio Cusimano*

- 173 Dagli archivi storici alle mappe genomiche. Il caso di Riposto  
*Salvatore Spina, Giuseppe Sorbello*
- 195 Identification sur Twitter des acteurs communiquant autour de la COVID-19  
au début de l'épidémie. Analyse des messages et des rumeurs diffusés  
*Marc Tanti, Hugo Moquet*

### **Note e Rubriche**

- 221 Organizzazione della conoscenza. Come mi vuoi, materiale o ideale?  
*Claudio Gnoli*

# La valutazione di interoperabilità dei formati elettronici nelle nuove Linee guida sulla formazione, gestione e conservazione dei documenti informatici

STEFANO ALLEGREZZA\*

**ABSTRACT:** This paper aims at addressing the issues of file formats assessment for the preservation of digital documents, which is critically important because the chance of preserving documents created in these formats and minimizing the migration operations that will become necessary over time depends greatly on it. In particular, the paper shows the methodology proposed by the Italian Agency for Digital Government who recently published the Guidelines on the creation, management and preservation of digital documents. For the first time worldwide, this methodology is based on a qualitative and quantitative method based on the evaluation of some properties of file formats and assigning them a score; the sum of these scores constitutes the so-called interoperability index which makes it possible to establish whether the format is prone to obsolescence or not. The method is explained with examples that show its potential. Some suggestions for improvement and further developments are then discussed.

**Keywords:** File formats, Assessment, Evaluation, Selection, Digital preservation.

## 1. Introduzione

Il tema della scelta dei formati elettronici<sup>1</sup> per la produzione dei documenti informatici riveste grande importanza<sup>2</sup>, dal momento che è anche da una loro

\* Dipartimento di Beni culturali, Università degli Studi di Bologna, Italia.  
stefano.allegrezza@unibo.it.

1. Un formato elettronico è «l'insieme di codici e regole che, permettono, a partire da una sequenza binaria, di riprodurre mediante un software, a video o a stampa, il relativo oggetto informativo con lo stesso contenuto e nella stessa forma che gli ha dato l'autore», cfr. S. PIGLIAPOCO, *La memoria digitale delle amministrazioni pubbliche*, Maggioli, Rimini 2005, p. 64. Nel glossario di INTERPARES il formato è definito come «l'organizzazione dei dati all'interno dei file, di solito progettata con lo scopo di facilitare la memorizzazione, il recupero, il trattamento, la presentazione o la trasmissione dei dati per mezzo del software», cfr. <[http://www.interpares.org/ip3/ip3\\_terminology\\_db.cfm?team=3&status=glossary](http://www.interpares.org/ip3/ip3_terminology_db.cfm?team=3&status=glossary)> (ultima consultazione: 20/04/2021).

2. Come osserva MARIA GUERCIO «non vi è dubbio che i formati costituiscano un nodo importante del processo di conservazione a lungo termine: una scelta corretta di formati non proprietari, aperti, standard *de jure* e orientati ai dati facilita l'individuazione di soluzioni adeguate e riduce i costi», cfr.

accurata selezione che dipendono, da una parte, la possibilità di conservare nel tempo i documenti prodotti in tali formati, dall'altra, la possibilità di ridurre al minimo le operazioni di riversamento sostitutivo (*migration*) che si renderanno necessarie nel tempo. Su questi temi ferve una approfondita riflessione a livello internazionale da almeno due decenni, ma al di là di linee guida e raccomandazioni che orientano la scelta verso questo o quel formato sulla base di valutazioni di tipo qualitativo<sup>3</sup>, prima d'ora non era mai stata proposta una metodologia di selezione basata su valutazioni di tipo quantitativo. Per questo motivo, la metodologia proposta dall'Agenzia per l'Italia Digitale (AgID) con la pubblicazione delle recenti *Linee guida sulla formazione, gestione e conservazione dei documenti informatici*, (nel seguito: *Linee guida*), ed in particolare con quella dell'*Allegato 2 — Formati di file e riversamento* (nel seguito: *Allegato 2*), risulta di particolare interesse<sup>4</sup>. Le *Linee guida*, infatti, hanno introdotto delle interessanti novità sul tema dei formati elettronici — come la valutazione di interoperabilità

M. GUERCIO, *Archivistica informatica. I documenti in ambiente digitale*, Carocci Editore, Roma 2019.

3. Su questi temi si vedano: LIBRARY OF CONGRESS, *Recommended Formats Statement*, <<https://www.loc.gov/preservation/resources/rfs>> (ultima consultazione: 31/05/2021); LIBRARY OF CONGRESS, *Sustainability of Digital Formats. Planning for Library of Congress Collections*, <<https://www.loc.gov/preservation/digital/formats>> (ultima consultazione: 20/04/2021); DIGITAL PRESERVATION COALITION, *File Formats Assessments*, <[https://wiki.dpconline.org/index.php?title=File\\_Formats\\_Assessments](https://wiki.dpconline.org/index.php?title=File_Formats_Assessments)> (ultima consultazione: 20/04/2021); THE NATIONAL ARCHIVES OF THE UNITED KINGDOM, *Selecting File Formats for Long-Term Preservation (Digital Preservation Guidance Note no. 1)*, <<https://www.nationalarchives.gov.uk/documents/selecting-file-formats.pdf>> (ultima consultazione: 20/04/2021); J. ROG, C. VAN WIJK, *Evaluating File Formats for Long-term Preservation*, <<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.181.506&rep=rep1&type=pdf>> (ultima consultazione: 20/04/2021); G. DRAGO, *Recommended file formats for long-term archiving and for web dissemination in Phaidra*, <<https://phaidra.cab.unipd.it/static/EN-file-formats.pdf>> (ultima consultazione: 20/04/2021). Si segnalano per il loro particolare interesse, anche per motivi storici, i risultati del progetto DELOS, patrocinato dalla Commissione Europea; il documento finale del 2004 classificava i formati, discuteva la 'longevità digitale' e suggeriva criteri di valutazione dei formati per l'uso da parte degli archivi, cfr. DELOS, *File Formats Typology and Registries for Digital Preservation* (2004), <<http://web.archive.org/web/20110721194942/>> (ultima consultazione: 20/04/2021), <[http://www.dpc.delos.info/private/output/DELOS\\_WP6\\_d631\\_finalv2\(5\)\\_urbino.pdf](http://www.dpc.delos.info/private/output/DELOS_WP6_d631_finalv2(5)_urbino.pdf)> (ultima consultazione: 20/04/2021); e quelli del progetto InterPARES 2: *General Study 11 Final Report: Selecting Digital File Formats for Long-Term Preservation* (2007), <[http://www.interpares.org/ip2/ip2\\_case\\_studies.cfm?study=35](http://www.interpares.org/ip2/ip2_case_studies.cfm?study=35)> (ultima consultazione: 20/04/2021).

4. Le *Linee Guida sulla formazione, gestione e conservazione dei documenti informatici* sono state pubblicate sul sito dell'AgID il 9 settembre 2020, <[https://trasparenza.agid.gov.it/archivio19\\_regolamenti\\_0\\_5385.html](https://trasparenza.agid.gov.it/archivio19_regolamenti_0_5385.html)> (ultima consultazione: 20/04/2021). Sono articolate in un documento principale e in sei allegati che ne costituiscono parte integrante. Gli allegati sono i seguenti: Allegato 1 — *Glossario dei termini e degli acronimi*; Allegato 2 — *Formati di file e riversamento*; Allegato 3 — *Certificazione di processo*; Allegato 4 — *Standard e specifiche tecniche*; Allegato 5 — *Metadati*; Allegato 6 — *Comunicazione tra AOO di Documenti Amministrativi Protocollati*.

di cui si parlerà nel seguito — allo scopo di assicurare, da una parte, una gestione più efficiente dei documenti informatici, dall'altra, una conservazione a lungo termine più sostenibile.

## 2. La scelta dei formati elettronici nel “vecchio” Allegato 2

Prima di affrontare l'argomento riteniamo opportuno ricordare che già le precedenti Regole tecniche sul protocollo informatico, quelle sul sistema di conservazione e quelle sul documento informatico<sup>5</sup> trattavano le questioni legate ai formati elettronici. Infatti, uno dei cinque allegati, l'Allegato 2 — *Formati*, era specificatamente dedicato a questo argomento; in particolare, il tema della scelta dei formati veniva trattato nel Capitolo 3, laddove si specificava che «ai fini della formazione, gestione e conservazione, è necessario scegliere formati che possano garantire la leggibilità e la reperibilità del documento informatico nel suo ciclo di vita. La scelta tra i formati dipende dalle caratteristiche proprie del formato e dei programmi che lo gestiscono»<sup>6</sup>. Il Paragrafo 3.1 individuava le caratteristiche generali di cui bisognava tener conto nella scelta: l'apertura, la sicurezza, la portabilità, la funzionalità, il supporto allo sviluppo e la diffusione. Nel successivo Capitolo 4 si introducevano ulteriori elementi di valutazione distinguendo tra i *formati per la formazione e gestione dei documenti* (par. 4.1) e i *formati per la conservazione* (par. 4.2). Nel primo caso venivano indicati come elementi aggiuntivi da valutare «l'efficienza in termini di occupazione di spazio fisico e la possibilità di gestire il maggior numero possibile di metadati, compresi i riferimenti a modifiche o aggiunte intervenute sul documento»<sup>7</sup>. Si precisava, poi, che «le pubbliche amministrazioni indicano nel manuale di gestione i formati

5. Cfr. Decreto del Presidente del Consiglio dei Ministri (DPCM) 3 dicembre 2013, *Regole tecniche in materia di sistema di conservazione ai sensi degli articoli 20, commi 3 e 5-bis, 23-ter, comma 4, 43, commi 1 e 3, 44, 44-bis e 71, comma 1, del Codice dell'amministrazione digitale di cui al decreto legislativo n. 82 del 2005* (Supplemento ordinario (s.o.) alla Gazzetta ufficiale, Serie Generale n. 59 del 11/4/2014); DPCM 3 dicembre 2013, *Regole tecniche per il protocollo informatico ai sensi degli articoli 40-bis, 41, 47, 57-bis e 71, del Codice dell'amministrazione digitale di cui al decreto legislativo n. 82 del 2005* (s.o. alla Gazzetta ufficiale, Serie Generale n. 59 del 11/4/2014); DPCM 13 novembre 2014, *Regole tecniche in materia di formazione, trasmissione, copia, duplicazione, riproduzione e validazione temporale dei documenti informatici nonché di formazione e conservazione dei documenti informatici delle pubbliche amministrazioni ai sensi degli articoli 20, 22, 23-bis, 23-ter, 40, comma 1, 41, e 71, comma 1, del Codice dell'amministrazione digitale di cui al decreto legislativo n. 82 del 2005*. (Gazzetta ufficiale, Serie Generale n. 8 del 12/1/2015). Gli allegati erano i seguenti: Allegato 1 — *Glossario/Definizioni*; Allegato 2 — *Formati di file e riversamento*; Allegato 3 — *Standard e specifiche tecniche*; Allegato 4 — *Specifiche tecniche del pacchetto di archiviazione*; Allegato 5 — *Metadati*.

6. Cfr. Allegato 2 — *Formati*, Capitolo 3 — *Criteri di scelta dei formati*, pp. 61–62.

7. Cfr. Allegato 2 — *Formati*, Paragrafo 4.1 — *Formati e prodotti per la formazione e gestione*, p. 63.



adottati per le diverse tipologie di documenti informatici motivandone le scelte effettuate; specificano altresì i casi eccezionali in cui non è possibile adottare i formati in elenco motivandone le ragioni<sup>8</sup>. Nel secondo caso, veniva specificato che era opportuno «privilegiare i formati che siano standard internazionali (*de jure* e *de facto*) o, quando necessario, formati proprietari le cui specifiche tecniche siano pubbliche, dandone opportuna evidenza nel manuale di conservazione dei documenti informatici»<sup>9</sup>. Inoltre, si faceva presente che «la scelta dei formati idonei alla conservazione oltre al soddisfacimento delle caratteristiche suddette deve essere strumentale a che il documento assuma le caratteristiche di immutabilità e di staticità previste dalle regole tecniche»<sup>10</sup>. Infine, si ricordava che «i formati per la conservazione adottati per le diverse tipologie di documenti informatici devono essere indicati nel manuale di conservazione motivandone le scelte effettuate; sono altresì specificati i casi eccezionali in cui non è possibile adottare i formati in elenco motivandone le ragioni»<sup>11</sup>.

Il successivo Capitolo 5 — *I formati indicati per la conservazione*, elencava una serie di formati ritenuti affidabili per la conservazione a lungo termine: il PDF e il PDF/A per i formati di descrizione di pagina; il TIFF e il JPG per i formati immagine; la famiglia di formati Office Open XML (OOXML) — come il DOCX, l’XLSX e il PPTX — e la famiglia di formati Open Document Format (ODF) — come l’ODT, l’ODS e l’ODP — per i documenti prodotti con gli strumenti di *office automation*; l’XML; il TXT; inoltre vi era un accenno al formato per i messaggi di posta elettronica (RFC 2822). Si trattava, quindi, di un elenco estremamente limitato — poco più di una dozzina di formati — che forniva indicazioni solo riguardo un numero ristretto di categorie di documenti informatici e non forniva alcuna indicazione in merito ad alcune categorie di formati — come quelle per i documenti sonori o audiovisivi — che pure stanno acquisendo un’importanza crescente negli archivi digitali oggi in formazione.

### 2.1. Il “nuovo” Allegato 2

Il nuovo Allegato 2 rappresenta un enorme passo avanti rispetto al precedente, sia per quanto riguarda le indicazioni metodologiche (finalmente si stabilisce un criterio misurabile, benché suscettibile di un certo grado di soggettività, per la scelta dei formati elettronici) che per quanto riguarda l’eshaustività con cui è trattato l’argomento (e questo risulta evidente anche dalla corposità dell’Allegato,

8. *Ibidem*.

9. Cfr. Allegato 2 — *Formati*, Paragrafo 4.2 — *Formati per la conservazione*, p. 63.

10. *Ibidem*.

11. *Ibidem*.

composto da ben 154 pagine — molte di più rispetto al precedente Allegato di sole 11 pagine). Già nel titolo si può osservare una differenza sostanziale. Mentre, infatti, il “vecchio” allegato era intitolato semplicemente *Formati*, il nuovo è intitolato *Formati di file e riversamento*, facendo intuire l'intenzione di andare ben oltre la semplice elencazione di formati elettronici ‘conservabili’ e spingendosi fino all'individuazione delle operazioni da compiere nel caso in cui si evidenzia che un formato elettronico stia diventando obsoleto.

L'allegato 2 è formato da tre capitoli (figura 1). Il primo, *Introduzione*, contiene le definizioni fondamentali ed una tassonomia dei formati elettronici. Il secondo, *Tipi di file*, costituisce la parte principale dell'Allegato ed è quello in cui si prendono in considerazione i formati elettronici (124 in totale) suddivisi nelle varie categorie<sup>12</sup>. Il terzo, *Raccomandazioni sui formati di file*, specifica le modalità per condurre la valutazione di interoperabilità, per il calcolo dell'indice di interoperabilità e per il riversamento di formato.

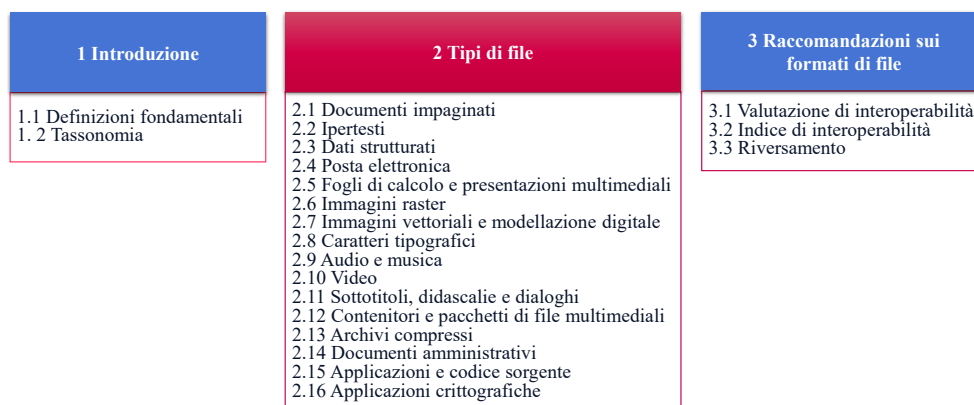


Figura 1. La struttura dell'Allegato 2.

L'Allegato distingue i formati in *generici* (o *generalisti*) e *specifici* (o *specialisti*). Alla prima categoria appartengono i formati che sono di uso comune (come DOCX, XLSX, PPTX, ODT, ODS, ODP, PDF, etc.) «rispetto ai quali tutte le P.P.AA. e le organizzazioni sul territorio nazionale: a) devono essere in grado di leggere file prodotti in questi formati, e b) seguono le indicazioni riportate nella tabella riassuntiva del formato per quanto concerne la produzione di nuovi documenti in questi

12. L'Allegato 2 suddivide i formati elettronici in 14 categorie: documenti impaginati; ipertestivi; dati strutturati; posta elettronica; fogli di calcolo e presentazioni multimediali; immagini raster; immagini vettoriali e modellazione digitale; caratteri tipografici; audio e musica; video; sottotitoli, didascalie e dialoghi; contenitori e pacchetti di file multimediali; archivi compressi; documenti amministrativi; applicazioni e codice sorgente; applicazioni crittografiche.

formati»<sup>13</sup>. Alla seconda categoria appartengono i formati, ulteriormente diversificati in base al settore merceologico o alla natura del documento informatico rappresentabile, verso i quali

sono identificati ulteriori obblighi e raccomandazioni solo verso le PP.AA. che, in quanto utenti professionali, trattano documenti di quella tipologia nell'ambito ristretto, o comunque delimitato, di quella categoria specifica. Tali categorie possono essere di riferimento anche per soggetti privati che trattano professionalmente i medesimi documenti informatici.<sup>14</sup>

Le Linee guida stabiliscono che la scelta dei formati elettronici da parte dei soggetti di cui all'art. 2 comma 2 e comma 3 del *Codice dell'Amministrazione Digitale* (CAD)<sup>15</sup> debba avvenire tra quelli previsti dall'Allegato 2. Nello scegliere tra di essi i formati elettronici da utilizzare per i propri documenti informatici, essi «possono effettuare una valutazione di interoperabilità che tenga conto dei seguenti fattori: formati aperti, non proprietari, standard *de iure*, estendibili, parlanti, completamente robusti, indipendenti dal dispositivo»<sup>16</sup>.

13. Cfr. Allegato 2 — *Formati di file e riversamento*, Paragrafo 1.2.3 — *Formati generici e specifici*, pp. 14–16. Vedi anche Linee guida, Paragrafo 3.6 — *Formati di file*, p. 29: «Le pubbliche amministrazioni garantiscono sempre la gestione dei formati classificati nell'Allegato 2 [...] come "generici", secondo la distinzione introdotta nell'Allegato 2 tra formati di file generici e specifici».

14. *Ibidem*.

15. I soggetti di cui all'art. 2, comma 2 del *Codice dell'Amministrazione digitale* (Decreto legislativo 82/2005) sono: a) le pubbliche amministrazioni di cui all'articolo 1, comma 2, del decreto legislativo 30 marzo 2001, n. 165, nel rispetto del riparto di competenza di cui all'articolo 117 della Costituzione, ivi comprese le autorità di sistema portuale, nonché alle autorità amministrative indipendenti di garanzia, vigilanza e regolazione; b) i gestori di servizi pubblici, ivi comprese le società quotate, in relazione ai servizi di pubblico interesse; c) le società a controllo pubblico, come definite nel decreto legislativo 19 agosto 2016, n. 175, escluse le società quotate di cui all'articolo 2, comma 1, lettera p), del medesimo decreto che non rientrino nella categoria di cui alla lettera b). L'art. 2, comma 3 del *Codice dell'Amministrazione digitale* specifica che «le disposizioni del presente Codice e le relative Linee guida concernenti il documento informatico, le firme elettroniche e i servizi fiduciari di cui al Capo II, la riproduzione e conservazione dei documenti di cui agli articoli 43 e 44, il domicilio digitale e le comunicazioni elettroniche di cui all'articolo 3 — bis e al Capo IV, l'identità digitale di cui agli articoli 3 — bis e 64 si applicano anche ai privati, ove non diversamente previsto».

16. Cfr. *Linee guida*, Paragrafo 3.6 — *Formati di file*, p. 29. Lo stesso paragrafo precisa ulteriormente che «qualora l'ordinamento giuridico preveda, per particolari categorie di documenti elettronici, degli obblighi relativamente all'uso di formati di file specifici ovvero di vincoli aggiuntivi su formati generici (quali, ad esempio, l'uso di particolari dialetti o specializzazioni per formati generici), le pubbliche amministrazioni, assolvendo tali obblighi, accettano i suddetti documenti elettronici solo se prodotti nei formati o con i vincoli aggiuntivi obbligatori».

### 3. La valutazione di interoperabilità

Le Linee guida consentono di utilizzare formati diversi da quelli elencati nell'Allegato 2 ma subordinano questa possibilità all'obbligo di effettuare una *valutazione di interoperabilità*<sup>17</sup>. Questa consiste in un processo — articolato in diverse azioni<sup>18</sup>, alcune obbligatorie ed altre facoltative — che prendono in considerazione numerosi aspetti dei formati elettronici allo scopo di costruire un quadro il più possibile esaustivo relativamente al loro utilizzo all'interno di una organizzazione. Le azioni obbligatorie sono le seguenti:

- innanzitutto, va effettuata una ricognizione di tutte le procedure amministrative e/o processi di business, allo scopo di individuare tutte le tipologie di formati elettronici trattati (o trattabili);
- va poi redatto un *censimento* dei formati elettronici utilizzati (con particolare attenzione a quelli non elencati nel Capitolo 2 dell'Allegato 2);
- per ciascun formato elettronico, vanno elencati tutti i dettagli tecnici del medesimo: ad esempio, il nome del formato e degli eventuali dialetti, dei profili, dei *codec* e degli schemi operativi; l'appartenenza ad una delle due categorie dei formati *generici* e *specifici*; le versioni utilizzate nei documenti già esistenti, ovvero di quelli producibili dagli attuali applicativi; le altre caratteristiche tecniche non vincolate dalle specifiche di cui ai punti precedenti (ad esempio, le lingue adottate nei testi, il numero di canali audio, gli spazi colore, la risoluzione per le immagini, il *frame rate* per i documenti audiovisivi, il *bitrate* massimo, gli algoritmi di cifratura, la presenza di password, etc.);
- vanno poi elencati i processi di riversamento di formato attualmente in corso nell'organizzazione, con particolare riferimento ai software applicativi impiegati e alle procedure tecniche (automatiche, semiautomatiche o completamente manuali) adottate per configurare tali riversamenti, con lo scopo prioritario di renderli riproducibili;
- va valutata l'eventuale necessità di operare sui medesimi documenti informatici nell'arco di una finestra temporale futura;
- va valutata la sussistenza di leggi o altri tipi di obblighi in merito alla con-

17. Le Linee guida non forniscono una definizione di *interoperabilità*, tuttavia il *Codice dell'amministrazione digitale* all'art. 1 la definisce come la «caratteristica di un sistema informativo, le cui interfacce sono pubbliche e aperte, di interagire in maniera automatica con altri sistemi informativi per lo scambio di informazioni e l'erogazione di servizi». Questa definizione deve essere associata a quella di *cooperazione applicativa* definita nel Codice come «la parte del Sistema Pubblico di Connettività finalizzata all'interazione tra i sistemi informativi dei soggetti partecipanti, per garantire l'integrazione dei metadati, delle informazioni, dei processi e procedimenti amministrativi».

18. Elencate nell'Allegato 2, Paragrafo 3.1 — *Valutazione di interoperabilità*, pp. 147–148.

- servazione dei documenti informatici nel formato originale in cui sono stati prodotti od acquisiti;
- vanno valutati i formati elettronici di categorie specifiche, nonché l'opportunità di riutilizzo dei documenti informatici di ciascuna categoria da parte di pubbliche amministrazioni e organizzazioni che operano al di fuori dello specifico settore per il quale il formato e i suoi scenari d'utilizzo sono stati prefigurati;
- deve essere valutata la dipendenza dei formati elettronici da licenze d'uso, marchi e brevetti o altre forme di proprietà intellettuale, sistemi e architetture proprietarie;

Le azioni facoltative sono le seguenti:

- vanno elencate le motivazioni attuali che hanno portato alla scelta di ciascun formato elettronico per il trattamento dei documenti informatici, distinguendo, se del caso, tra i formati adottati per i documenti "in entrata" (ad esempio, quelli trasmessi dai cittadini, dalle imprese o da altre organizzazioni); quelli ad uso esclusivamente interno; quelli "in uscita" (ad esempio, quelli invitati ad altre organizzazioni o pubblicati on-line); infine, quelli archiviati ovvero mandati in conservazione;
- va valutata l'esistenza di standard o di iniziative di standardizzazione a livello internazionale, europeo e nazionale, relativamente alle tipologie di documenti informatici trattati;
- vanno valutati gli scenari ove successive modificazioni o revisioni dei documenti vengano prodotte in formati diversi da quello originale;
- va valutata la dipendenza dei formati elettronici da sistemi e architetture associati a costi di manutenzione ordinaria o straordinaria, senza la quale la capacità di elaborare i documenti codificati nei suddetti formati potrebbe essere a rischio o fortemente ridotta;
- infine, l'obsolescenza dei formati di file e delle tecnologie di archiviazione va collocata all'interno di una più ampia strategia di trasformazione digitale dell'organizzazione.

Al termine della valutazione deve essere redatto un dettagliato *rapporto*, che diventa parte integrante del Manuale di gestione dei documenti<sup>19</sup> insieme con l'elenco dei formati utilizzati.

19. Le Linee guida stabiliscono che nel manuale di gestione documentale sono riportati «relativamente ai formati dei documenti: a) l'individuazione dei formati utilizzati per la formazione del documento informatico, come introdotti nel paragrafo 3.6, tra quelli indicati nell'Allegato 2 "Formati di file e riversamento"; b) la descrizione di eventuali ulteriori formati utilizzati per la formazione di documenti in relazione a specifici contesti operativi che non sono individuati

Dal momento che la valutazione di interoperabilità è un processo piuttosto complesso e richiede competenze molto specialistiche nel settore dei formati elettronici, spesso non rinvenibili presso le pubbliche amministrazioni (PP.AA.) o gli altri soggetti a cui si applicano le Linee guida<sup>20</sup>, ci si potrebbe domandare quali siano i soggetti obbligati ad effettuarla e quali, invece, ne siano esentati. Dalla lettura delle Linee guida si deduce che sono obbligati i soggetti che utilizzano formati elettronici diversi da quelli elencati nel Capitolo 2 dell'Allegato 2 o quelli che utilizzano formati elettronici conformi all'Allegato 2 ma in difformità dalle raccomandazioni ivi presenti<sup>21</sup>. Ciò significa che, di fatto, sono obbligati tutti i soggetti, e questo almeno per due ordini di motivi:

- in primo luogo perché l'elenco dei formati presenti nell'Allegato 2 non è esaustivo e sono numerosi i formati elettronici in uso presso le organizzazioni che non vi sono elencati (come, ad esempio, il TXT che è utilizzato praticamente da tutte le organizzazioni, il formato BMP e il formato WEBP nella categoria delle immagini *raster*, il formato AAC nella categoria dei documenti sonori, il formato WEBM nella categoria dei documenti audiovisivi, il formato PST nella categoria dei formati per la posta elettronica, etc.);
- in secondo luogo perché vi sono numerosi formati che vengono utilizzati in difformità dalle indicazioni, come, ad esempio, il diffusissimo DOCX. Infatti, l'Allegato 2 specifica che i documenti in tale formato devono essere prodotti nel profilo *strict* mentre nella quasi totalità dei casi vengono prodotti nel profilo *transitional*, che è quello utilizzato di *default* da Microsoft Word, e che non è considerato interoperabile.

Ad ogni modo, anche se esistesse un'organizzazione che non rientra in nessuno dei due casi appena visti, la valutazione di interoperabilità «in quanto parte della gestione informatica dei documenti, viene effettuata periodicamente e, comunque, ogni anno, allo scopo di individuare tempestivamente cambiamenti nelle condizioni»<sup>22</sup> (ad esempio, formati che sono diventati obsoleti e quindi ne-

nell'Allegato 2 "Formati di file e riversamento"; c) le procedure per la valutazione periodica di interoperabilità dei formati e per le procedure di riversamento previste come indicato al paragrafo 3.7 e nell'Allegato 2 "Formati di file e riversamento", cfr. Linee guida, Paragrafo 3.5 — *Manuale di gestione documentale*, p. 27.

20. Si ricorda che le Linee Guida e i connessi Allegati sono applicabili ai soggetti indicati nell'art. 2 commi 2 e 3 del *Codice dell'amministrazione digitale*, cfr. nota 15.

21. Cfr. Allegato 2, Paragrafo 3.1 — *Valutazione di interoperabilità*, p. 147: «La valutazione di interoperabilità è redatta con cadenza annuale dalle PP.AA. che trattino documenti informatici in formati diversi da quelli di cui al presente Allegato, ovvero conformi a questi formati ma utilizzati disapplicando gli obblighi e raccomandazioni ivi contenuti».

22. Cfr. Linee guida, Paragrafo 3.6 — *Formati di file*, p. 29.

cessitano di essere sottoposti al processo di riversamento sostitutivo) e permettere quindi di mettere subito in atto azioni correttive.

L'Allegato 2 al Paragrafo 3.1 specifica ulteriormente che la valutazione di interoperabilità «può essere redatta da qualunque ente pubblico o privato tratti documenti informatici». Ciò significa che può essere effettuata anche dai privati che, pur non essendone obbligati, possono trarre giovamento da questo strumento utile per «prevenire il rischio della loro obsolescenza tecnologica (rischio che incrementa con il divergere delle caratteristiche del formato da quelle sopraelencate); mitigare il rischio del *vendor lock-in*, che sussiste soprattutto per i formati proprietari, non codificati in standard *de iure*, o per quelli dipendenti dal dispositivo; facilitare il più possibile un loro futuro riversamento in altro formato, prediligendo caratteristiche quali l'estendibilità, la codifica testuale, la compatibilità in avanti e, in misura minore, la robustezza»<sup>23</sup>.

#### 4. Il modello semplificato di valutazione di interoperabilità

Per venire incontro alle difficoltà che le pubbliche amministrazioni e gli altri soggetti obbligati potrebbero incontrare nel condurre una valutazione di interoperabilità, l'Agenzia per l'Italia Digitale ha proposto un modello semplificato basato sulla determinazione dell'*indice di interoperabilità*. Si tratta di un metodo quali-quantitativo che consiste nel valutare un formato elettronico prendendo in considerazione una serie di nove caratteristiche (figura 2)<sup>24</sup> a ciascuna delle quali viene attribuito un valore numerico ('punteggio'); la somma dei valori associati a tutte le caratteristiche prese in considerazione fornisce come risultato l'*indice di interoperabilità* che può variare tra un minimo pari a 0 (zero) ed un massimo pari a 20<sup>25</sup>. Il formato più interoperabile è quello che raggiunge un indice di interoperabilità pari a 20, quello meno interoperabile è quello che consegue un indice pari a 0. Un valore pari a 12 viene considerato come soglia di sufficienza: sono "accettabili" tutti quei formati che raggiungono un indice di interoperabilità pari o superiore a 12; valori inferiori indicano problematiche oggettive che vanno affrontate il prima possibile ricorrendo, ad esempio, a processi di riversamento o ad altre metodologie.

23. Cfr. Allegato 2, Capitolo 3 — *Raccomandazioni sui formati di file*, p. 146.

24. Cfr. Allegato 2, Paragrafo 3.2 — *Indice di interoperabilità*, p. 149.

25. Per la verità, facendo la somma dei valori massimi attribuibili alle varie caratteristiche da valutare, si ottiene un valore pari a 21, ma si tratta evidentemente di un refuso.



Figura 2. Le caratteristiche da valutare per calcolare l'indice di interoperabilità.

Per determinare l'indice di interoperabilità devono essere valutate le seguenti caratteristiche dei formati<sup>26</sup>:

- a) la *standardizzazione* (da 0 a 3): ai formati che hanno ottenuto il riconoscimento come standard *de jure* da un organismo di standardizzazione (come l'International Organization for Standardization (ISO), l'Ente nazionale italiano di unificazione (UNI), il World Wide Web Consortium (W3C), il Comitato europeo di normazione (CEN), l'International Telecommunication Union (ITU), la Society of Motion Picture and Television Engineers (SMPTE), etc.) viene attribuito il punteggio massimo, pari a 3; quelli che non hanno avuto tale riconoscimento ma sono diventati comunque, grazie alla loro larghissima diffusione, standard *de facto* conseguono un punteggio pari a 2. Infine, un formato può anche non essere né standard *de jure* né *de facto* (punteggio pari a 0), ma tale caratteristica tendenzialmente lo esclude dall'elenco dei formati accettabili;
- b) l'*apertura* (da 0 a 3): i formati aperti, ovvero quelli per i quali sono state pubblicate e rese disponibili — eventualmente anche a pagamento — le specifiche del formato<sup>27</sup>, conseguono il massimo punteggio, pari a 3; al

26. Per una approfondita trattazione delle questioni relative ai requisiti dei formati elettronici compatibili con un processo di conservazione digitale si rimanda a: S. ALLEGREZZA, *Requisiti e standard dei formati elettronici per la produzione di documenti informatici*, in «ARCHIVI & COMPUTER», a. XIX, 2-3 / 2009, 2009, pp. 42-82; S. ALLEGREZZA, S. PIGLIAPOCO, *Produzione e conservazione del documento digitale. Requisiti e standard per i formati elettronici*, EUM, Macerata, 2008.

27. La *specificata* di un formato elettronico è quel documento che descrive dettagliatamente sia le modalità tecniche che occorre seguire per la *lettura* di un documento codificato secondo



- contrario, i formati chiusi, ovvero quelli per i quali non sono state rese disponibili le specifiche, conseguono il punteggio minimo, pari a 0.
- c) la *non proprietà* (da 0 a 4): i formati non proprietari, ovvero quei formati su cui non vi sono diritti di proprietà intellettuale e le cui specifiche non sono gestite da un'organizzazione privata ma da una comunità di sviluppatori che cooperano per la gestione condivisa delle stesse (ad esempio, la comunità di LibreOffice) o da un organismo di standardizzazione, conseguono il punteggio massimo, pari a 4; i formati proprietari, ovvero quelli che sono stati creati da una organizzazione privata (ad esempio, una software house), che ne detiene i diritti di proprietà intellettuale e ne gestisce le specifiche, conseguono un punteggio variabile, che va da 0 a 3; in particolare: i formati proprietari ma di utilizzo libero conseguono un punteggio pari a 3; quelli proprietari che consentono la lettura di documenti già codificati in tale formato ma non la produzione di nuovi conseguono un punteggio pari a 2; infine, quelli che non consentono neanche la lettura dei documenti codificati secondo tale formato conseguono un punteggio pari a 0<sup>28</sup>;
- d) l'*estendibilità* (da 0 a 2): i formati estendibili, ovvero concepiti fin dall'inizio per ammettere successive versioni che ne aumentino progressivamente le funzionalità, conseguono il punteggio massimo, pari a 2; viceversa quelli non estendibili conseguono un punteggio pari 0<sup>29</sup>.

quel formato (ovvero, come deve essere interpretata la sequenza di bit (bitstream) che lo costituisce), sia quelle necessarie per la *scrittura* (ovvero, come deve essere prodotta la sequenza di bit che lo costituisce). A queste, che sono le informazioni minime, possono aggiungersi altre informazioni, ad esempio, quelle relative alla modalità di gestione. Si tratta, di solito, di documenti piuttosto voluminosi: ad esempio, la specifica del formato PDF nella versione 1.7 è costituita da un documento estremamente tecnico composto da oltre 1000 pagine; la specifica della famiglia di formati OOXML è contenuta in un documento di oltre 5000 pagine. Esempi di formati proprietari sono quelli che erano gestiti dal pacchetto di Office automation Microsoft Office fino alla versione 2003 (DOC, XLS, PPT, etc.), i quali sono stati, per diversi anni, anche chiusi perché le loro specifiche non erano mai state rese note. Invece, la famiglia di formati OOXML (DOCX, XLSX, PPTX, etc.) è non proprietaria ed aperta; il formato PDF, pur essendo stato per molti anni proprietario, è sempre stato aperto in quanto le sue specifiche sono sempre state liberamente accessibili. Altri esempi di formati aperti sono il TXT, l'RTF, l'ODF, l'XML, l'HTML e l'XHTML, il JPEG e il PNG per le immagini raster; l'SVG per le immagini vettoriali; il FLAC per l'audio con compressione *lossless*, il Musepack e l'Ogg Vorbis per l'audio con compressione *lossy*; l'Ogg Theora e l'xvid per il video con compressione *lossy*.

28. Potrebbero essere possibili anche altri tipi di vincoli (pagamento di *royalty*, sottoscrizione di contratti di riservatezza o contratti vincolanti relativamente a particolari utilizzi quali lo sfruttamento commerciale, ecc.), che andranno valutati caso per caso.

29. La *ratio* di questa distinzione risiede probabilmente nel fatto che «i formati non estendibili [...] possono comunque essere soggetti a revisioni, che però potrebbero, per tali formati, richiedere una reingegnerizzazione o un adattamento più difficoltoso rispetto a formati estendibili, probabilmente anche a scapito delle compatibilità di cui al punto precedente». Cfr. Allegato 2, paragrafo 1.2.2 — *Classificazione dei formati*.

- e) il *livello di metadati* (da 0 a 3): il livello del modello per i metadati segue l'analoga classificazione adottata nelle *Linee guida nazionali per la valorizzazione del patrimonio informativo pubblico*<sup>30</sup>, emanate dall'Agenzia per l'Italia Digitale, ove al livello 1 vengono attribuiti 0 punti, al livello 2 viene attribuito 1 punto, e così via fino al livello 4 cui sono attribuiti 3 punti. Per i formati descritti nel Capitolo 2 dell'Allegato 2, tale valore viene indicato nelle tabelle riassuntive di ciascun formato.
- f) la *robustezza* (da 0 a 2): i formati non robusti, ovvero quelli che, nel caso di corruzione della sequenza di bit che costituisce un documento informatico, non consentono il recupero di nessuna parte del documento originario, conseguono il punteggio minimo, pari a 0; i formati completamente robusti, che prevedono meccanismi per verificare l'eventuale perdita di integrità e permettono il recupero dell'intero documento (o, nel caso di formati contenitore<sup>31</sup>, delle parti integre), conseguono il punteggio massimo, pari a 2; i formati parzialmente robusti conseguono un punteggio pari a 1.

30. Le *Linee guida nazionali per la valorizzazione del patrimonio informativo pubblico* sono pubblicate all'indirizzo <<https://docs.italia.it/italia/daf/lg-patrimonio-pubblico/it/stabile/index.html>> (ultima consultazione: 20/04/2021). La classificazione qualitativa dei metadati prende in considerazione due fattori principali: il legame tra dato–metadato e il livello di dettaglio. Sulla base di tali fattori viene data la seguente classificazione: 1) *livello 1*: legame dato–metadato: nessun legame in quanto i dati non sono accompagnati da un'opportuna metadatozione; livello di dettaglio: nessuno in quanto i metadati non sono presenti; 2) *livello 2*: legame dato–metadato: il legame è debole perché i dati sono accompagnati da metadati esterni, (e.g., inclusi nella pagina di *download* del *dataset* o in file separati); livello di dettaglio: i metadati forniscono informazioni relativamente a un *dataset*, quindi sono informazioni condivise dall'insieme di dati interni a quel *dataset*; 3) *livello 3*: legame dato–metadato: il legame è forte perché i dati incorporano i metadati che li descrivono; livello di dettaglio: i metadati forniscono informazioni relative a un *dataset*, quindi sono informazioni condivise dall'insieme di dati interni a quel *dataset*; 4) *livello 4*: legame dato–metadato: il legame è forte perché i dati incorporano i metadati che li descrivono; livello di dettaglio: i metadati forniscono informazioni relative al singolo dato, quindi col massimo grado di dettaglio possibile.

31. Un formato contenitore (*container format*, detto anche *wrapper format* o *package format*) è un «formato di file progettato per consentire l'inclusione (“imbustamento” o *wrapping*), in uno stesso file, di una o più evidenze informatiche soggette a differenti tipi di codifica e al quale possono essere associati specifici metadati». Cfr. Allegato 1 — *Glossario dei termini e degli acronimi* alle Linee guida. In sostanza, si tratta di un formato elettronico che può contenere diversi tipi di oggetti digitali, ciascuno dei quali distinguibile dagli altri. Esistono varie tipologie di formati contenitore: quelli utilizzati per i documenti della famiglia OOXML e ODF (formati di pacchetto); quelli utilizzati per i documenti sonori ed audiovisivi (formati contenitore multimediali); quelli utilizzati per la creazione di documenti compressi come ZIP e RAR (formati di compressione); quelli utilizzati per i documenti sottoscritti con firma digitale nei tre formati CADES, PAdES e XAdES (formati per buste crittografiche), etc. Ad esempio, il formato DOCX è un formato di pacchetto: il contenitore, formato da un pacchetto di file distribuiti in più cartelle, è compresso con l'algoritmo ZIP e presentato come un unico file.

- g) l'*indipendenza da dispositivi* (da 0 a 4): i formati indipendenti dal dispositivo, ovvero rappresentabili in maniera attendibile e sempre allo stesso modo indipendentemente dalla piattaforma hardware e dal software impiegato, conseguono il punteggio massimo, pari a 4; quelli dipendenti dal dispositivo conseguono il punteggio minimo, pari a 0.

Vi sono poi altre due caratteristiche valutabili (a cui, tuttavia, l'Allegato 2 non attribuisce un punteggio specifico):

- h) la *compatibilità in avanti e all'indietro*: i formati compatibili all'indietro (retrocompatibili) sono quelli che possono essere utilizzati con una delle versioni precedenti dell'applicativo che lo ha prodotto; i formati compatibili in avanti sono quelli che è possibile utilizzare con una delle versioni del software successive a quella con cui è stato prodotto<sup>32</sup>.
- i) la *codifica testuale o binaria*: i formati testuali sono quelli che consentono di estrarre il contenuto informativo codificando ogni *byte* della sequenza binaria con il carattere corrispondente (ad esempio, nella codifica ASCII)<sup>33</sup>; viceversa i formati che non ammettono questo tipo di rappresentazione si dicono binari; in entrambi i casi viene attribuito un punteggio pari a 0.

Sulla base di questa valutazione viene assegnato un valore numerico per ciascuna delle nove caratteristiche di un formato (ad eccezione delle ultime due, alle quali, come si è già detto, l'Allegato 2 non attribuisce un punteggio specifico). Facendo la somma dei valori attribuiti a ciascuna caratteristica si ottiene un valore complessivo che rappresenta l'indice di interoperabilità.

Particolare attenzione deve essere prestata nel caso dei *formati contenitori* nel qual caso la valutazione deve riguardare non solo il formato del contenitore ma ogni singolo oggetto digitale contenuto al suo interno; inoltre deve considerare come addendo il valore più basso (cioè peggiore) per ciascuno di tali oggetti. Ad esempio per un formato contenitore multimediale deve essere valutato il formato delle singole essenze (audio, video) in esso contenute; per un formato

32. Si noti che, di solito, la compatibilità in avanti è garantita (ad esempio, con Microsoft Word 2019, che produce documenti nel formato DOCX, è possibile — ad oggi — leggere i documenti di testo creati nel precedente formato DOC); invece la compatibilità all'indietro è raramente assicurata.

33. Esempi di formati testuali sono il TXT, l'XML, l'HTML. Sono binari i formati MP3, WAV, MP4, solo per citare alcuni tra quelli più conosciuti. Si noti che anche i formati utilizzati per la produzione di documenti di testo possono essere binari, come nel caso del formato DOC, oppure testuali, come nel caso del formato RTF, oppure, ancora, sia testuali che binari, come nel caso del formato DOCX che, essendo un formato di pacchetto, contiene al suo interno sia oggetti digitali testuali (ad esempio, i file XML che contengono la componente testuale del documento) che binari (ad esempio, le immagini inserite nel documento).

di pacchetto andranno valutati i formati di tutti gli oggetti digitali compresi nel pacchetto e, qualora il pacchetto comprenda, a sua volta, dei formati contenuti, andranno anch'essi valutati con il medesimo criterio appena esposto<sup>34</sup>.

## 5. Alcuni esempi di calcolo dell'indice di interoperabilità

Per comprenderne meglio il meccanismo di determinazione dell'indice di interoperabilità, proviamo ad applicarlo al caso di alcuni formati elettronici di utilizzo comune. Cominciamo considerando il formato DOC, il formato di *default* prodotto con il programma Microsoft Word fino alla versione 2003. Andando ad effettuare la valutazione delle nove caratteristiche necessaria per il calcolo dell'indice di interoperabilità si possono ottenere i risultati riportati in tabella 1.

Tabella 1. Calcolo dell'indice di interoperabilità per il formato DOC.

Caratteristica	Intervallo	Valutazione	Valore
standardizzazione	da 0 a 3	Il formato DOC non è standard <i>de jure</i> ma certamente, vista la sua ancora notevole diffusione, può essere considerato uno standard <i>de facto</i>	2
apertura	da 0 a 3	Le specifiche del formato, inizialmente chiuse, sono ora aperte	3
non proprietà	da 0 a 4	Il formato è proprietario	0
estendibilità	da 0 a 2	Il formato è stato abbandonato dalla stessa Microsoft e sostituito con il formato DOCX	0
livello di metadati	da 0 a 3	Valore individuato nell'Allegato 2	3
robustezza	da 0 a 2	Trattandosi di un formato binario, la robustezza è limitata	1

34. L'Allegato 2 riporta, a titolo di esempio, il calcolo dell'indice di interoperabilità «di uno specifico pacchetto di file ove tutti i file utilizzino formati indipendenti dal dispositivo tranne un file multimediale, la cui busta e tutte le sue essenze sono anch'esse indipendenti dal dispositivo tranne una sola essenza audio (che ad esempio potrebbe richiedere la presenza di un dispositivo hardware proprietario per la sua decodifica)». In tal caso «il pacchetto andrà considerato comunque come "dipendente dal dispositivo" e quindi gli verrà applicato un modificatore pari a 0 (anziché a +3), proprio a causa della presenza dell'unico file il cui formato è dipendente dal dispositivo», cfr. Allegato 2, p. 149.

indipendenza da dispositivi	da 0 a 4	Sono noti i problemi di visualizzazione di documenti in formato doc in ambienti diversi da Windows	2
compatibilità in avanti e all'indietro	n.d.	Il formato è compatibile in avanti	-
testuale o binario	n.d.	Il formato è binario	-
<b>TOTALE</b>			<b>11</b>

Il punteggio ottenuto non raggiunge la soglia di sufficienza e quindi il formato non è interoperabile. Si noti che la valutazione presenta un certo grado di soggettività in considerazione del fatto che valutatori diversi potrebbero attribuire punteggi diversi per ciascuna delle caratteristiche prese in considerazione. Questo non deve essere visto necessariamente come un fattore negativo, dal momento che ci possono essere situazioni diverse che richiedono valutazioni diverse. A tal proposito l'Allegato 2 precisa che «l'indice di interoperabilità è meramente indicativo, in quanto non tiene conto di alcuna peculiarità che un'organizzazione possa avere in merito a specifici formati di file»<sup>35</sup>.

Consideriamo ora il formato DOCX<sup>36</sup>. A differenza del formato .DOC, che memorizza i dati del documento in un singolo file binario, il formato DOCX utilizza le *Open Packaging Conventions*<sup>37</sup> per creare un 'pacchetto' di file, compressi con l'algoritmo ZIP, all'interno del quale sono raccolte le varie componenti necessarie per rappresentare il documento. All'interno di un documento in formato DOCX vi sono dei file XML e tre cartelle, "docProps", "Word" e "\_rels", che contengono le proprietà del documento, il contenuto e le relazioni tra i file. Questa struttura è progettata per rendere il contenuto del documento più ac-

35. Cfr. Allegato 2, p. 149.

36. Per esplorare manualmente i contenuti di un file DOCX è sufficiente rinominare l'estensione ".DOCX" in ".ZIP" e quindi decomprimere il file risultante con qualsiasi utilità di decompressione ZIP.

37. Le *Open Packaging Conventions* (OPC) sono delle convenzioni, inizialmente create da Microsoft, che definiscono le modalità per la creazione di formati contenitore attraverso la combinazione di file XML e non XML, solitamente raggruppati insieme in un unico contenitore e compressi mediante l'algoritmo ZIP; in questo modo si ha il vantaggio di lasciare intatti ed indipendenti gli oggetti digitali incorporati nel documento e si ottengono documenti di dimensioni generalmente più piccole rispetto alla somma delle dimensioni degli oggetti digitali di cui sono costituiti. Queste convenzioni sono utilizzate da diversi formati elettronici di tipo contenitore, come l'XPS (Microsoft XML Paper Specification), l'OXPS (Open XML Paper Specification), i formati della famiglia OOXML (Office Open XML) come il DOCX, l'XLSX e il PPTX; il formato VSDX (Microsoft Visio 2013 Drawing Format) utilizzato da Microsoft Visio, il formato DWFx (Design Web Format) di AutoCAD, etc., cfr. *Open Packaging Conventions Fundamentals*, <<https://docs.microsoft.com/en-us/previous-versions/windows/desktop/opc/open-packaging-conventions-overview>> (ultima consultazione: 20/04/2021).

cessibile. Ad esempio, il testo del documento viene salvato utilizzando un file di testo normale e le immagini incorporate nel documento vengono memorizzate come singoli file immagine. Questi file possono anche includere informazioni di formattazione della pagina, dati dell'autore e note di revisione del documento. Si tratta, quindi, di un "formato di pacchetto" per il quale occorrerebbe andare ad esaminare le singole componenti che si trovano all'interno del pacchetto, che possono essere anche molto diverse, dal momento che in un documento in formato DOCX è possibile incorporare non solo immagini (in vari formati: JPG, PNG, GIF, ecc.) ma anche contenuti audio (anch'essi in vari formati: WAV, WMA, MP3, ecc.) e video (anche qui in vari formati: WMV, AVI, MP4, ecc.). Limitiamoci, per semplicità, al caso del formato DOCX con il solo contenuto testuale. In tal caso, evidentemente molto particolare e solo in minima parte rappresentativo dei documenti presenti in un qualsiasi archivio digitale, il calcolo dell'indice di interoperabilità può essere fatto come segue:

**Tabella 2.** Calcolo dell'indice di interoperabilità per il formato docx.

<b>Caratteristica</b>	<b>Intervallo</b>	<b>Valutazione</b>	<b>Valore</b>
standardizzazione	da 0 a 3	Il formato DOCX è standard de jure (ISO 29500) ed è diventato ormai anche standard de facto	3
apertura	da 0 a 3	Le specifiche del formato sono aperte	3
proprietà	da 0 a 4	Il formato è non proprietario	4
estendibilità	da 0 a 2	Il formato è estendibile	2
livello di metadati	da 0 a 3	Valore individuato nell'Allegato 2	3
robustezza	da 0 a 2	Trattandosi di un formato contenente solo la componente testuale (codificata in XML) la robustezza è elevata	2
indipendenza da dispositivi	da 0 a 4	I documenti in formato DOCX possono essere letti su diverse piattaforme hardware e software, anche se a volte viene evidenziata una non perfetta compatibilità	3
compatibilità in avanti e all'indietro	n.d.	Il formato è compatibile in avanti	-
testuale o binario	n.d.	Il formato è in parte binario ed in parte testuale (le componenti in XML sono testuali)	-
<b>TOTALE</b>			<b>20</b>

Il risultato che si ottiene è pari a 20, quindi il formato, in questa forma estremamente semplificata, risulta interoperabile. Per completezza sarebbe necessario ripetere il calcolo con riferimento a ciascuna delle varie combinazioni che possono presentarsi all'atto pratico, considerando, ad esempio, documenti in formato DOCX contenenti anche immagini, contenuti audio e contenuti video, oltre ovviamente al testo. È possibile che da una siffatta analisi derivino delle conclusioni inaspettate: ad esempio potrebbe verificarsi che il formato risulti interoperabile nel caso in cui vengano incorporate immagini codificate in determinati formati interoperabili (come il JPG) mentre potrebbe risultare non interoperabile nel caso in cui vengano incorporati contenuti video in formati non interoperabili.

Analoghe considerazioni devono essere fatte per gli altri tipi di formati contenitori, come quelli multimediali: anche in tal caso occorre valutare tutte le possibili combinazioni con tutti i possibili formati di oggetti digitali che possono essere inseriti all'interno del contenitore. Per fare un esempio, si consideri il formato contenitore MKV (Matroska video), utilizzato diffusamente per la produzione e la conservazione di documenti audiovisivi. Al suo interno (figura 3) possono essere presenti essenze video in vari formati (MPEG-2, MPEG-4, WMV, Realvideo, Adobe Flash, H.264, H.265, VP8 VP9, Theora, etc.), essenze audio anch'esse in vari formati (PCM, MP2, MP3, AAC, Vorbis, WMA, RA, AC3, DTS, FLAC, etc.), nonché sottotitoli (nei formati USE, SRT, ASS/SSA, OGG WRIT, WebVTT, etc.) e metadati. Per valutare compiutamente il formato si dovrebbero prendere in esame tutte le combinazioni possibili (ricordando che, nel caso di presenza di più oggetti, occorre prendere come valore di riferimento quello peggiore) ed identificare sia quelle interoperabili che quelle non interoperabili. La complessità dell'operazione sta nell'elevato numero di oggetti digitali che possono essere presenti all'interno del contenitore, e, di conseguenza, nel numero assai elevato di combinazioni possibili, ciascuna delle quali deve essere oggetto di valutazione. Si potranno, quindi, avere dei casi in cui, in virtù di una certa combinazione, il formato risulta interoperabile e casi in cui, in virtù di combinazioni differenti, il formato risulta non interoperabile. Di tutto ciò si dovrà tener nella valutazione di interoperabilità.

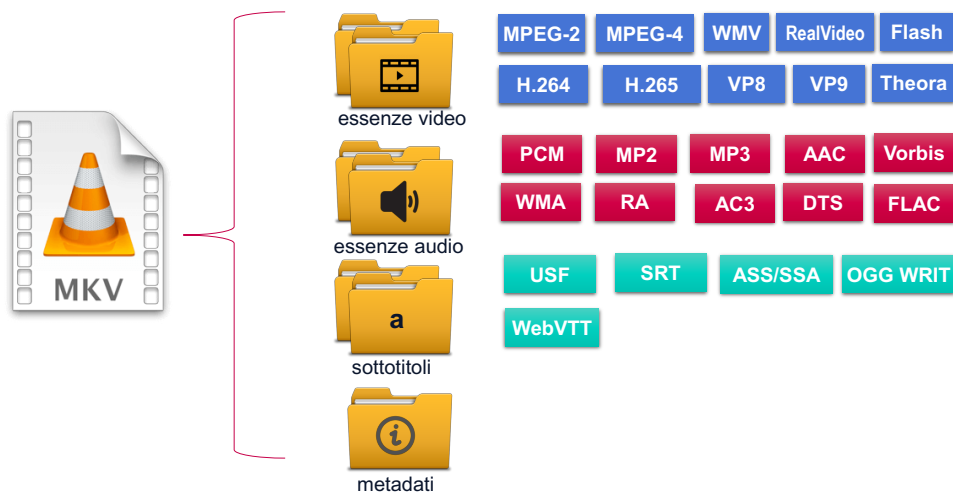


Figura 3. Gli oggetti digitali che possono essere presenti all'interno del formato mkv.

Questo esempio dimostra la complessità della questione, che non può ridursi alla valutazione di casi semplificati ma deve prevedere tutti le possibili situazioni che possono presentarsi.

Consideriamo ora il ben noto formato MP3, utilizzato per la produzione di documenti sonori. Il calcolo dell'indice di interoperabilità può essere condotto come segue:

Tabella 3. Calcolo dell'indice di interoperabilità per il formato MP3.

Caratteristica	Intervallo	Valutazione	Valore
standardizzazione	da 0 a 3	Il formato MP3 è standard de jure (ISO/IEC 11172-3 e ISO/IEC 13818-3) oltre che standard de facto	3
apertura	da 0 a 3	Le specifiche del formato sono aperte	3
proprietà	da 0 a 4	Il formato è proprietario ma di utilizzo libero	3
estendibilità	da 0 a 2	Il formato non è estendibile	0
livello di metadati	da 0 a 3	Valore individuato nell'Allegato 2	3
robustezza	da 0 a 2	Trattandosi di un formato compresso con compressione di tipo lossy non è robusto	0



indipendenza da dispositivi	da 0 a 4	I documenti in formato MP3 possono essere letti su diverse piattaforme hardware e software	4
compatibilità in avanti e all'indietro	n.d.	Il formato è compatibile in avanti	-
testuale o binario	n.d.	Il formato è binario	-
<b>TOTALE</b>			<b>16</b>

Quindi anche il formato MP3, ad oggi, può essere considerato interoperabile, anche se, trattandosi di un formato che si sta incamminando verso l'obsolescenza<sup>38</sup>, andrà 'tenuto sotto controllo' ripetendo la valutazione almeno ogni anno.

Qualora, a seguito della valutazione di interoperabilità, si rilevasse che un formato elettronico non è interoperabile (e quindi potenzialmente a rischio) occorre procedere al riversamento verso un formato più interoperabile. Le modalità di esecuzione di tale operazione vengono descritte nel paragrafo 3.3 — *Riversamento* dell'Allegato 2. I riversamenti vengono pianificati a seguito di una nuova valutazione di interoperabilità, considerando i costi e i benefici di una tale operazione. In ogni caso le pubbliche amministrazioni devono effettuare operazioni di riversamento verso «formati che ne migliorino l'interoperabilità, o comunque non la peggiorano (come, stabilito, ad esempio, mediante il calcolo dell'indice di interoperabilità)»<sup>39</sup>. In particolar modo *non* sono opportuni riversamenti da un formato aperto verso formati chiusi; da un formato non proprietario, ovvero proprietario a libero utilizzo, verso formati proprietari; da un formato non dipendente dal dispositivo verso formati dipendenti da dispositivo; da un formato parlante verso formati muti<sup>40</sup>.

## 6. Considerazioni conclusive e sviluppi futuri

Come si è già avuto modo di far notare, nel complesso il “nuovo” Allegato 2 rappresenta un enorme passo avanti rispetto al “vecchio” Allegato 2. L'impostazione è indubbiamente buona e la decisione di rendere la scelta dei formati

38. Nel 2017 il suo sviluppatore e proprietario, l'Istituto Fraunhofer, ha ufficialmente dichiarato che non avrebbe più mantenuto il formato MP3 in quanto non più in grado di competere con i formati attuali, più efficienti e con qualità superiori. Cfr. *MP3 ufficialmente morto. Troppo obsoleto per gli standard odierni*, <<https://focustech.it/2017/05/16/mp3-ufficialmente-morto-obsoleto-gli-standard-odierni-141821>> (ultima consultazione: 20/04/2021).

39. Cfr. Allegato 2 — *Riversamento*, p. 151.

40. Cfr. Allegato 2, Paragrafo 3.3 — *Riversamento*, p. 150.

elettronici il meno soggettiva possibile attraverso il calcolo dell'indice di interoperabilità, basato sulla valutazione delle caratteristiche dei formati e sulla determinazione di un valore numerico, appare corretta e pienamente valida sotto il profilo metodologico. Purtroppo, non è possibile eliminare del tutto la soggettività nel processo valutativo, e questo potrebbe rappresentare una criticità da risolvere o, in qualche modo, mitigare, come si dirà meglio più avanti.

Tornando al calcolo dell'indice di interoperabilità, riteniamo ci sia margine per qualche azione di miglioramento, al fine di renderlo uno strumento ancor più efficace.

In primo luogo va osservato che alcune delle caratteristiche oggetto di valutazione non sono facilmente determinabili, e questo per diversi ordini di motivi. In alcuni casi, la difficoltà dipende dal fatto che le fonti di informazioni sono insufficienti o non esistono affatto. Se per i formati più diffusi è relativamente facile acquisire le informazioni necessarie per la valutazione, per i formati meno diffusi ciò non è assolutamente vero e può essere necessario reperire e analizzare approfonditamente le specifiche dei formati, spesso costituite da documenti estremamente tecnici e composti da centinaia o migliaia di pagine. Prendiamo, ad esempio, la caratteristica dell'*estendibilità*: come si può sapere se un formato è stato progettato fin dall'inizio per essere estendibile? In alcuni casi ciò non costituisce un problema (si pensi al PDF/A che è stato creato fin dall'inizio come formato estendibile, e questa caratteristica è stata ampiamente comunicata), ma in altri, specialmente per i nuovi formati, non è dato conoscere a priori se il produttore ha in animo di sviluppare ulteriori versioni di quel formato. Spesso è possibile sapere se un formato è estendibile solo a distanza di tempo, dopo che sono state rilasciate le successive versioni.

Altre volte le fonti di informazioni ci sono ma il reperimento dello specifico dato necessario per valutare una particolare caratteristica di un formato risulta difficile e a volte impossibile. Ad esempio, come valutare la caratteristica del *livello di metadati*? Per i formati compresi tra i 124 descritti nel Capitolo 2 dell'Allegato 2, il valore numerico da attribuire è riportato direttamente nelle relative schede e, quindi, lo si può desumere direttamente da esse<sup>41</sup>. Ma come fare per i formati non presenti nell'elenco? Il più delle volte bisogna fare ricorso, ancora una volta, all'analisi della specifica del formato, ma non è detto che l'informazione di cui si ha bisogno sia immediatamente rintracciabile dal momento che, come si è già avuto modo di far notare, si tratta spesso di documenti assai tecnici e voluminosi; ad ogni modo, la ricerca richiede solitamente parecchio tempo e competenze tecniche non alla portata di tutti e non di rado si conclude con un nulla di fatto perché l'informazione desiderata non è presente.

41. A questo proposito si fa presente che nelle schede relative ai vari formati non vengono riportati valori se non quelli relativi al livello di metadati; sarebbe utile riportare la valutazione anche per le altre caratteristiche.

In alcuni casi, poi, le caratteristiche solo valutabili solo sperimentalmente. Prendiamo, ad esempio, la *robustezza*: questa caratteristica è misurabile solo ricorrendo ad apposite verifiche sperimentali, effettuate utilizzando specifici software che vanno artificialmente a ‘corrompere’ la sequenza di bit (simulando in questo modo i fenomeni di corruzione dovuti al trascorrere del tempo, a difetti nei supporti di archiviazione, ad operazioni di riversamento non andate a buon fine, ad errori nella trasmissione dei documenti attraverso le reti telematiche, etc.) e cercando poi di recuperare integralmente o parzialmente il contenuto<sup>42</sup>. Ma anche in questo caso, la possibilità di effettuare tali verifiche richiede competenze che, il più delle volte, vanno ben oltre quelle presenti tra il personale dei soggetti obbligati alla valutazione in interoperabilità.

In secondo luogo, sarebbe utile avviare un confronto sulle caratteristiche dei formati elettronici da prendere in considerazione, insieme con i relativi punteggi, al fine di migliorare la capacità dell’indice di interoperabilità di costituire una valutazione fedele del livello di interoperabilità dei formati. Ad esempio, alcune caratteristiche, come l’*indipendenza dal dispositivo*, appaiono al giorno d’oggi meno utili rispetto a quanto non lo fossero alcuni anni fa, dal momento che i formati elettronici ‘moderni’ sono in grande maggioranza indipendenti dal dispositivo; altre appaiono prive di una sufficiente rilevanza ai fini dell’interoperabilità, come ad esempio l’*estendibilità*, dal momento che la decisione di privilegiare i formati estendibili rispetto a quelli non estendibili potrebbe non essere condivisa da tutti. Anche la distinzione tra codifica *testuale* o *binaria* appare non così rilevante, anche in considerazione del fatto che i formati testuali sono ormai pochi, e confinati prevalentemente tra i formati per i documenti a prevalente contenuto testuale. Altre caratteristiche, invece, non sono state incluse nell’elenco di quelle sottoposte a valutazione; il caso paradigmatico è quello dell’*adozione*, considerata unanimemente una delle caratteristiche più importanti dal momento che l’ampia adozione di un formato elettronico costituisce uno dei deterrenti principali nei confronti della sua obsolescenza<sup>43</sup>. Anche l’attribuzione dei punteggi alle caratteristiche oggetto di valutazione potrebbe essere affinata, al fine

42. La robustezza è una caratteristica che è facilmente valutabile in senso *relativo* ma meno in senso *assoluto*. Infatti, è relativamente semplice stabilire se un formato sia più o meno robusto di un altro. Ad esempio, nel campo dei formati immagine i formati compressi sono meno robusti di quelli non compressi, e i formati con compressione di tipo *lossy* sono meno robusti di quelli con compressione di tipo *lossless*. Ciò significa che il formato TIFF è più robusto del formato JPEG o del PNG, tanto per fare un esempio che riguarda formati noti. Invece non è facile parlare di robustezza in senso assoluto.

43. Quando un formato è ampiamente adottato, il rischio che diventi obsoleto è molto basso. Al contrario, è possibile che formati che possiedono tutte le migliori caratteristiche per essere considerati interoperabili (sono standard *de jure*, non proprietari, aperti, etc.) vengano presto abbandonati e diventino obsoleti se non raggiungono un’ampia adozione.

di attribuire maggiore importanza a quelle caratteristiche che effettivamente incidono sulla possibilità di conservazione a lungo termine di un formato elettronico ed ottenere risultati più aderenti alla realtà. Ad esempio, potrebbe essere rivisto il “peso” attribuito alla caratteristica dell'*indipendenza dal dispositivo*, che può raggiungere un valore massimo pari a 4, e che conferisce a tale caratteristica un'importanza pari a quella di altre caratteristiche, come la *standardizzazione* o la *non proprietà*, che vengono generalmente considerate di livello superiore per quanto ad importanza.

Infine, vorremmo tornare sulla questione della soggettività delle valutazioni alla quale abbiamo accennato all'inizio e che riteniamo un aspetto particolarmente critico. Infatti, non è assolutamente da escludere la possibilità che la valutazione di uno stesso formato elettronico condotta da soggetti diversi porti a risultati anche notevolmente differenti. Come si è visto, questo potrebbe non essere necessariamente una situazione da evitare (ad esempio, ci potrebbero essere situazioni particolari in cui lo stesso formato debba essere considerato non interoperabile da un soggetto ed interoperabile dall'altro, magari perché quest'ultimo deve sottostare a regole interne o disposizioni che prevedono obbligatoriamente l'utilizzo di quel formato) ma, almeno in linea teorica, sarebbe auspicabile che la valutazione dello stesso formato fornisse gli stessi risultati indipendentemente dal soggetto che la compie. Questo richiede che le caratteristiche da prendere in esame siano misurabili con un basso livello di indeterminazione, ma questo risultato è particolarmente difficile da ottenere.

Una soluzione alternativa per ridurre al minimo la soggettività, a nostro parere preferenziale, potrebbe essere quella di costituire, a livello centrale, un tavolo di lavoro permanente formato da un team di esperti con competenze specifiche sulle questioni relative ai formati elettronici, che provveda ad effettuare le valutazioni di interoperabilità dei formati elettronici e le renda pubblicamente disponibili per tutti. Si tratterebbe, in buona sostanza, di costituire un *Registro nazionale dei formati elettronici*<sup>44</sup>, che potrebbe all'inizio comprendere le valutazioni dei formati di utilizzo più generale (considerando che molti dei formati elettronici sono comuni a più organizzazioni) e si arricchirebbe via via con le valutazioni dei formati che di volta in volta in volta verrebbero effettuate (anche a seguito delle richieste avanzate dai vari soggetti obbligati)<sup>45</sup>. Il registro conter-

44. Una previsione di questo genere è già presente *in nuce* nell'*Introduzione* all'Allegato 2: «il presente Allegato, per la natura stessa dell'argomento trattato, viene periodicamente aggiornato sulla base dell'evoluzione tecnologica e dell'obsolescenza dei formati e potrà essere pubblicato online sotto forma di Avvisi, ovvero di un registro dei formati sul sito istituzionale dell'Agenda per l'Italia Digitale». Cfr. Allegato 2, Capitolo 1 — *Introduzione*, p. 3.

45. Soluzioni di questo genere sono già state implementate sia a livello europeo che internazionale. Ad esempio, i National Archives of the United Kingdom hanno sviluppato il registro tecnico PRONOM (Public Record Office and Nôm), che contiene un database ricercabile di

rebbe le valutazioni di tutte le caratteristiche dei formati necessarie per il calcolo dell'indice di interoperabilità; verrebbe mantenuto continuamente aggiornato e potrebbe fornire anche le indicazioni sui formati che sono ancora attuali, su quelli che stanno andando incontro all'obsolescenza ma che ancora non lo sono diventati, su quelli che sono ormai diventati obsoleti, su quelli che sono ormai estinti<sup>46</sup>. Un registro del genere sarebbe estremamente utile e costituirebbe una fonte autorevole ed affidabile di informazioni, a cui tutti potrebbero attingere, e che andrebbe a vantaggio di tutti coloro che sono chiamati ad effettuare la valutazione di interoperabilità.

In conclusione, la valutazione di interoperabilità proposta dall'AgID, pur suscettibile di miglioramento, va senza dubbio nella direzione giusta ed ha l'indiscutibile pregio di aver avviato un percorso che, una volta perfezionato, porterà all'individuazione di una metodologia per la valutazione dei formati elettronici affidabile, robusta e proponibile come *best practice* anche a livello internazionale.

informazioni tecniche sui formati di file, gli strumenti software e gli ambienti tecnici necessari per accedervi. Nel passato, le Harvard University Libraries avevano sviluppato il registro GDFR (Global Digital Format Registry), un registro tecnico che è stato attivo dal 2005 al 2009. Dal 2010 al 2016 è stato attivo anche il registro UDFR (Unified Digital Format Registry), che aveva cercato di 'unificare' le funzioni di due registri esistenti, PRONOM e GDFR ed era il risultato di un progetto sviluppato dalla University of California Curation Center (UC3) presso la California Digital Library (CDL) e finanziato dal National Digital Infrastructure Preservation Program (NDIIPP) della Library of Congress. I registri attualmente attivi contengono numerose informazioni sui formati elettronici, ma nessuno di questi contiene tutte le informazioni necessarie per condurre la valutazione di interoperabilità nella sua forma semplificata attraverso il calcolo dell'indice di interoperabilità. Questa limitazione è una ulteriore spinta verso la costituzione di un Registro nazionale dei formati.

46. Sulla falsariga della sezione *The 'Bit List' of Digitally Endangered Species* presente nel sito della Digital Preservation Coalition (DPC), che contiene una lista delle 'specie digitali' in pericolo di estinzione, suddivise nelle categorie *Lower risk, Vulnerable, Endangered, Critically endangered, Practically extinct, Concern*. La lista è continuamente aggiornata dalla comunità che aderisce alla coalizione e consente di scoprire quali materiali digitali sono più a rischio, così come quelli che sono relativamente sicuri grazie alla conservazione digitale. Continuando a mantenere aggiornato questo elenco nei prossimi anni, il DPC si prefigge l'obiettivo di celebrare i risultati dei grandi sforzi in materia di conservazione digitale che si sono compiuti in questi anni e che hanno consentito che le categorie diventassero meno 'preoccupanti', pur sottolineando la necessità di un ulteriore impegno per salvaguardare quei materiali che sono ancora considerati 'in grave pericolo'. Cfr. <[www.dpconline.org/digipres/champion-digital-preservation/bit-list](http://www.dpconline.org/digipres/champion-digital-preservation/bit-list)> (ultima consultazione: 20/04/2021).

## Riferimenti bibliografici

- AGENZIA PER L'ITALIA DIGITALE, *Linee Guida sulla formazione, gestione e conservazione dei documenti informatici* <[https://trasparenza.agid.gov.it/archivio19\\_regolamenti\\_o\\_5385.html](https://trasparenza.agid.gov.it/archivio19_regolamenti_o_5385.html)>.
- ALLEGREZZA S., PIGLIAPOCO S., *Produzione e conservazione del documento digitale. Requisiti e standard per i formati elettronici*, EUM, Macerata 2008.
- \_\_\_\_\_, *Requisiti e standard dei formati elettronici per la produzione di documenti informatici*, in «ARCHIVI & COMPUTER», a. XIX, 2-3 / 2009, 2009.
- DELOS, *File Formats Typology and Registries for Digital Preservation* (2004), <[web.archive.org/web/20110721194942/http://www.dpc.delos.info/private/output/DELOS\\_WP6\\_d63I\\_finalv2\(5\)\\_urbino.pdf](http://www.dpc.delos.info/private/output/DELOS_WP6_d63I_finalv2(5)_urbino.pdf)>.
- DIGITAL PRESERVATION COALITION, *File Formats Assessments*, <[https://wiki.dpconline.org/index.php?title=File\\_Formats\\_Assessments](https://wiki.dpconline.org/index.php?title=File_Formats_Assessments)>.
- DRAGO G., *Recommended file formats for long-term archiving and for web dissemination in Phaidra*, <<https://phaidra.cab.unipd.it/static/EN-file-formats.pdf>>.
- GUERCIO M., *Archivistica informatica. I documenti in ambiente digitale*, Carocci Editore, Roma 2019.
- INTERPARES 2, *Selecting Digital File Formats for Long-Term Preservation. General Study 11. Final Report* (2007), <[http://www.interpares.org/ip2/ip2\\_case\\_studies.cfm?study=35](http://www.interpares.org/ip2/ip2_case_studies.cfm?study=35)>.
- LIBRARY OF CONGRESS, *Recommended Formats Statement*, <<https://www.loc.gov/preservation/resources/rfs>>. PIGLIAPOCO S., *La memoria digitale delle amministrazioni pubbliche*, Maggioli, Rimini 2005.
- ROG J., VAN WIJK C., *Evaluating File Formats for Long-term Preservation*, <[citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.181.506&rep=rep1&type=pdf](http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.181.506&rep=rep1&type=pdf)>.
- THE NATIONAL ARCHIVES, *Selecting File Formats for Long-Term Preservation*, <<https://www.nationalarchives.gov.uk/documents/selecting-file-formats.pdf>>.