Un'impresa comune degli Archivi svizzeri



Manuale di applicazione

Contenuto

1	Prefa	azione	2		
2	Requisiti di sistema				
3		i aperti / feedback			
4	Introduzione alla doppia convalida PDF/A				
5	Installazione di KOST-Val				
6	Conf	igurazione di KOST-Val	4		
	6.1	Componenti del file di configurazione "kostval.conf.xml"	4		
7	Risor	rse KOST-Val	8		
8	Avvio della convalida				
	8.1	GUI KOST-Val	g		
	8.2	Avviare la convalida manualmente	11		
9	Сору	right	12		
	9.1	Licenza API 3-Heights™ PDF/A Validator	13		
	9.2	Licenza pdfaPilot CLI	14		
	9.3	Licenza egov-validationclient-cli	14		
10	Appe	endice	15		
	10.1	Struttura del programma	15		
	10.2	Convalida del formato del principio funzionale	16		

KOST-Val_Manuale_di_applicazione_v2.2.0.0.docx CR 21.02.2024

1 Prefazione

KOST-Val è un'applicazione basata su Java per la convalida della struttura e del contenuto di file PDF/A, JP2, JPEG, TIFF, PNG, FLAC, WAVE, MP3, MKV, MP4, XML, SIARD e i cosiddetti Submission Information Package (SIP) per la consegna di informazioni digitali. Questa applicazione è rilasciata sotto licenza GPL3+ ed è resa disponibile al pubblico da CECO come open source. KOST-Val si basa su componenti non modificati di altri produttori, che sono direttamente integrati nel codice sorgente di KOST-Val. Gli utenti di KOST-Val sono tenuti a rispettare le condizioni di licenza di tutti questi componenti. Informazioni dettagliate si possono trovare nel capitolo 9.

I risultati (compresi i messaggi relativi a incongruenze o errori) vengono emessi per ogni fase e scritti in un file di registro di convalida.

Le singole fasi di convalida/controllo vengono eseguite una dopo l'altra. Se possibile, la convalida continua anche in caso di errori, per ridurre il numero di cicli di correzione.

2 Requisiti di sistema

- Microsoft Windows a 64 bit
- Almeno 512 MB di RAM
- Almeno 20 GB di spazio su disco rigido

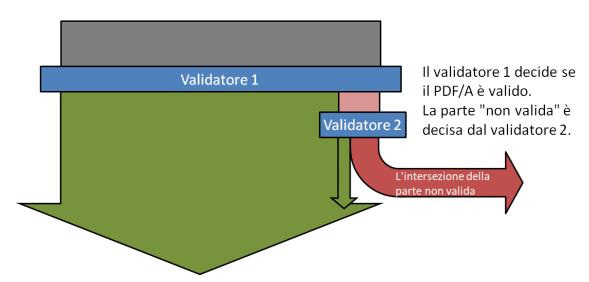
3 Punti aperti / feedback

I punti aperti, da bug ed errori a richieste di integrazioni e domande, si trovano sulla piattaforma di sviluppo GitHub alla voce Issues (https://github.com/KOST-CECO/KOST-Val/issues) e possono essere segnalati a kost-val@kost-ceco.ch.

Questo elenco può e deve essere ampliato da chiunque e viene modificato dal team di sviluppo.

4 Introduzione alla doppia convalida PDF/A

Per il PDF/A, KOST-Val offre l'opzione della doppia convalida. Un file PDF/A viene prima controllato da un primo validatore. Se il risultato non è valido, viene controllato da un secondo validatore. Il file PDF/A è considerato valido se almeno uno dei validatori lo identifica come tale, e non valido se entrambi i validatori lo identificano come tale.¹



La doppia convalida PDF/A può essere utilizzata solo se l'archivio consente di accettare file PDF/A potenzialmente non validi. In caso contrario, la doppia convalida PDF/A non deve essere utilizzata.

Per la doppia convalida, vengono utilizzati sia 3-Heights™ PDF/A Validator di PDF-Tools che pdfaPilot di callas. Se è acceso un solo validatore, viene eseguita automaticamente una sola convalida.

La base concettuale della doppia convalida è l'osservazione che anche i validatori PDF/A di alta qualità possono produrre risultati differenti. Da un lato, ciò è dovuto al fatto che l'attuale standard PDF/A include una serie di altri standard che non sono necessariamente implementati nei validatori fin in tutti i dettagli. D'altro lato, alcune specifiche dello standard sono formulate in modo tale da poter essere legittimamente implementate in modi diversi. Che tutti gli strumenti rilevanti implementino le specifiche in modo uniforme e completo rimane per il momento un sogno irrealizzabile. KOST-Val offre quindi una doppia validazione come soluzione provvisoria.

¹ In questo senso, la doppia convalida può essere eseguita solo con validatori PDF/A di alta qualità. Tra gli altri, le ultime versioni di 3-Heights™ PDF/A Validator di PDF-Tools e pdfaPilot di callas soddisfano questi requisiti elevati.

5 Installazione di KOST-Val

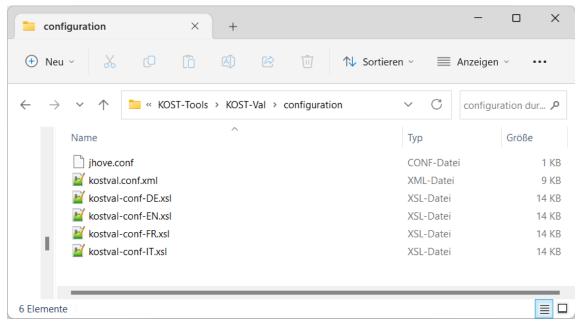
1 KOST-Val (versione 2.1.3.0 e successive) è ora disponibile solo nel pacchetto di installazione a 64 bit KOST-Tools.msi².

https://github.com/KOST-CECO/KOST-Val/releases/latest

Dopo aver scaricato KOST-Tools, il pacchetto di installazione deve essere eseguito con i diritti di amministratore.

KOST-Val è quindi disponibile nel menu di avvio sotto KOST-Tools.

6 Configurazione di KOST-Val



Il file "jhove.conf" è memorizzato nella cartella "configuration" e non deve essere adattato. "jhove.conf" è necessario per la validazione interna di Jhove.

Il file di configurazione "kostval.conf.xml" e i tre fogli di stile vengono copiati nella directory "USERHOME/.kost-val_2x/configuration" se non sono corretti o attualmente disponibili. Tutte le configurazioni di KOST-Val possono essere effettuate tramite GUI.

6.1 Componenti del file di configurazione "kostval.conf.xml"

Il file di configurazione "kostval.conf.xml" è strutturato in diverse parti. I componenti sono descritti brevemente di seguito.

.

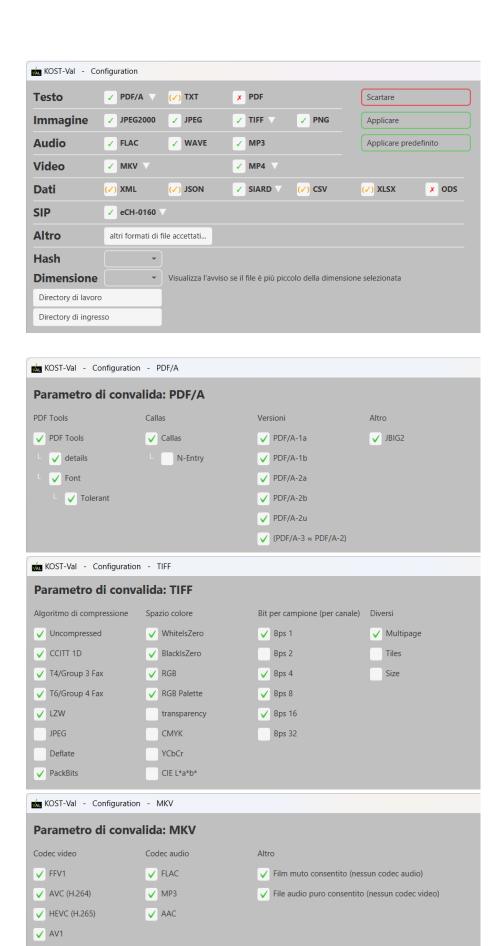
² Istruzioni più dettagliate sull'installazione e sulla sua portata sono disponibili nel manuale di KOST-Tools.

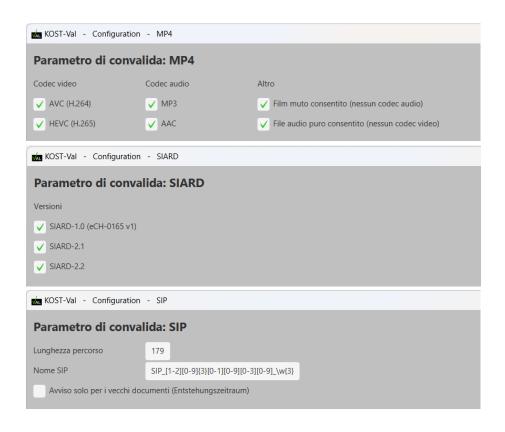


 \checkmark = accettato e convalidato

(✓) = accettato x = non accettato

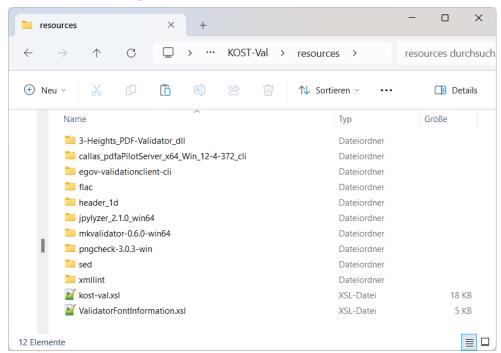
DDE/A: Accettations a convolida [/]	/	
PDF/A: Accettazione e convalida [/]		
Convalida PDF/A con PDF Tools [yes]:	yes	
- PDF Tools anche errori dettagliati in inglese [yes]:	yes	
- Convalida (ricercabilità ed estraibilità) [tolerant]:	tolerant	
Convalida PDF/A con callas [yes]:	yes	
- Errore (E) / avviso (W) se la voce N non corrisponde [W]:	W	
Versioni PDF/A permesse [1A, 1B, 2A, 2B, 2U]:	1A 1B 2A 2B 2U	
Convalida di PDF/A-3 in PDF/A-2 ed emissione di un avviso invece di un errore [yes]:	yes	
È ammessa la compressione JBIG2 [yes]:	yes	
TXT: Accettazione [(✓)]	(✓)	
PDF: Accettazione [X]	X	
JPEG2000: Accettazione e convalida [✓]	✓	
JPEG: Accettazione e convalida [✓]	✓	
TIFF: Accettazione e convalida [✓]	✓	
Algoritmi di compressione ammessi [uncompressed, CCITT 1D, CCITT Group 3, CCITT Group 4, LZW, PackBits]:	uncompressed CCITT 1D CCITT Group 3 CCITT Group 4 LZW PackBits	
Spazi colore ammessi [white is zero, black is zero, RGB, palette color]:	white is zero black is zero RGB palette color	
Bits per Sample ammessi [1, 4, 8, 16]:	1 4 8 16	
Sono ammessi i TIFF multipagina [yes]:	yes	
È consentita la struttura a piastrelle (tile) [no]:	no	
Sono consentiti file di dimensioni pari o superiori a 1000MB (~1GB) [no]:	no	
PNG: Accettazione e convalida [✓]	✓	
FLAC: Accettazione [✓]	✓	
WAVE: Accettazione [/]	✓	
MP3: Accettazione [✓]	✓	
MKV: Accettazione e convalida [✓]	✓	
- Codec video ammessi [FFV1, AVC, HEVC, AV1]:	FFV1 AVC HEVC AV1	
- Codec audio ammessi [FLAC, MP3, AAC]:	FLAC MP3 AAC	
- Film muto consentito (nessun codec audio) [Warning]:	Warning	
- File audio puro consentito (nessun codec video) [Warning]:	Warning	
MP4: Accettazione e convalida [✓]	✓	
- Codec video ammessi [AVC, HEVC]:	AVC HEVC	
- Codec audio ammessi [MP3, AAC]:	MP3 AAC	
- Film muto consentito (nessun codec audio) [Warning]:	Warning	
- File audio puro consentito (nessun codec video) [Warning]:	Warning	
XML: Accettazione e convalida [(✓)]	(✓)	
JSON: Accettazione [(✓)]	(✓)	
SIARD: Accettazione e convalida [✓]	✓	
Versioni SIARD ammesse [1.0, 2.1, 2.2]:	1.0 2.1 2.2	
CSV: Accettazione [(✓)]	(✓)	
XLSX: Accettazione [(\(\sigma \)]	(v)	
ODS: Accettazione [x]	×	
SIP: Validierung [✓]:	<i>,</i>	
Numero massimo di caratteri consentiti nella lunghezza dei percorsi [179]:	179	
Specifiche per la struttura del nome SIP [SIP_[1-2][0-9] $\{3\}$ [0-1][0-9][0-3][0-9]_\w{3}]:	SIP_[1-2][0-9]{3}[0-1][0-9][0-3][0-9]_\w{3}	
Avviso solo per i vecchi documenti (Entstehungszeitraum) [no]:	no	
Altri formati di file accettati [WARC, HTML, DWG]:	HTML WARC DWG	
Calcola ed emette il valore hash dei file. Vuoto significa nessun calcolo		
e nessun output []: Rilascia un avviso per i file di piccole dimensioni. Vuoto significa nessun avviso []:		
Directory di lavoro []:		
Directory di ingresso []:		





7 Risorse KOST-Val

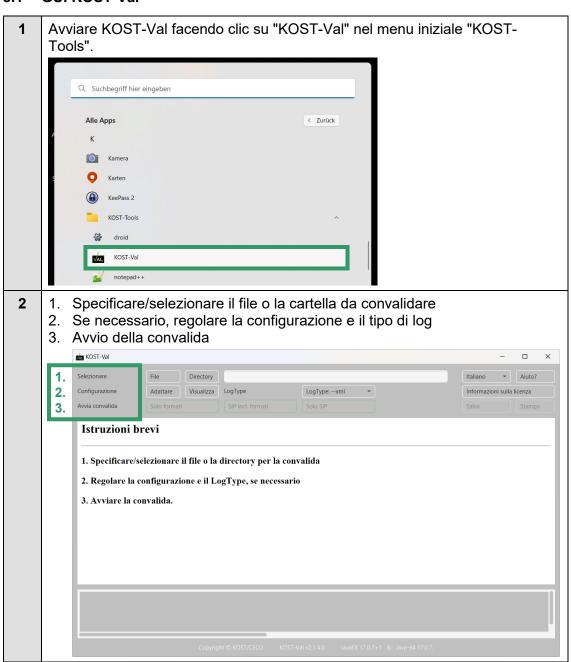
Tutte le risorse di KOST-Val sono memorizzate nella sottocartella "risorse".

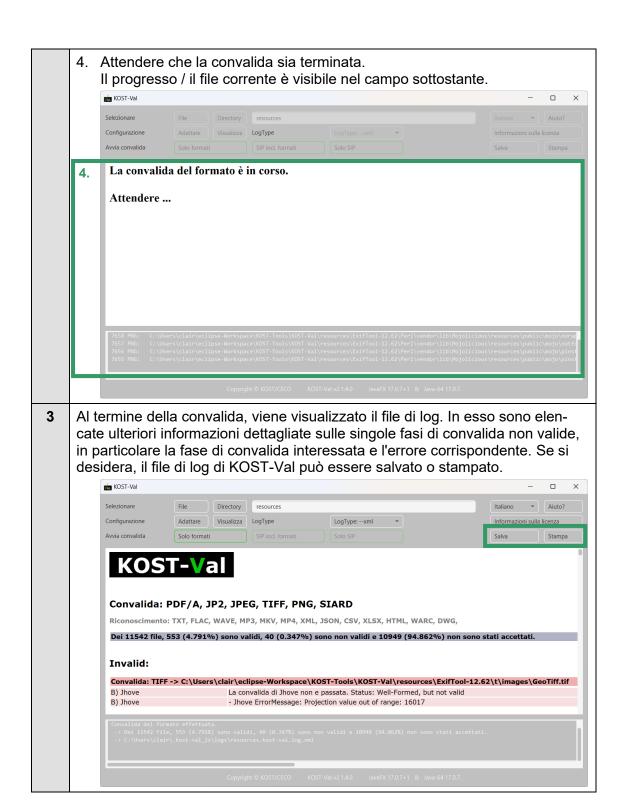


8 Avvio della convalida

! KOST-Val non è thread-safe!
Ciò significa che più istanze di KOST-Val non possono essere eseguite contemporaneamente senza interferire l'una con l'altra. Se KOST-Val viene eseguito contemporaneamente, possono verificarsi errori come la mancanza di una copia di lavoro.

8.1 GUI KOST-Val





8.2 Avviare la convalida manualmente

Aprire il prompt dei comandi e passare alla directory di lavoro desiderata (cd "C:\Program Files\KOST-CECO\KOST-Tools\KOST-Val")³.

Avviare la chiamata al programma KOST-Val (separare i singoli componenti di input con spazi).

```
..\Liberica_JRE\bin\java.exe -jar cmd_KOST-Val.jar --format resources --it --xml
```

```
C:\Program Files\KOST-CECO\KOST-Tools\KOST-Val>..\Liberica_JRE\bin\java.exe -jar cmd_KOST-Val.jar --format resources --it --xml
```

Note:

L'input java -jar è possibile solo se il Java Runtime Environment (JRE) desiderato è la versione standard.

Se necessario, è possibile regolare l'impostazione della memoria virtuale di Java. Si consiglia di regolare -Xmx per gli errori di "Out of Memory" e -Xss per gli errori di "Stack Overflow" (java -xmx1024m -xss128m -jar).

Se un componente di input contiene spazi, deve essere inserito con caratteri iniziali e finali.

KOST-Val può anche essere richiamato da qualsiasi posizione. Tuttavia, ciò richiede l'inserimento di percorsi assoluti.

Struttura del comando KOST-Val:

Il file è stato convalidato non appena nella finestra cmd viene visualizzato "Valid" o "Invalid". La cartella è stata convalidata non appena è stato visualizzato il prompt

(C:\Program Files\KOST-CECO\KOST-Tools\KOST-Val>).

```
Convalida del formato effettuata.

-> Dei 11542 file, 553 (4.791%) sono validi, 40 (0.347%) sono non validi e 10949 (94.862%) non sono stati accettati

-> C:\Users\clair\.kost-val_2x\logs\resources.kost-val.log.xml

C:\Program Files\KOST-CECO\KOST-Tools\KOST-Val>
```

I risultati dettagliati sono visibili nel file cost-val.log.xml.

Anche il risultato della convalida complessiva (file corretto/errato) viene emesso ed è visibile nell'exit status (o valore di uscita) del programma, in modo che la convalida possa essere integrata in una catena di elaborazione automatizzata. L'exit status può assumere i seguenti valori:

0 tutto OK

1 errore nella chiamata al programma

2 la convalida non è stata superata

³ L'unità viene modificata, ad esempio con c:.

9 Copyright

KOST-Val è uno sviluppo di CECO. Tutti i diritti sono di proprietà di CECO. KOST-Val è stato pubblicato nel 2012 da CECO sotto la GNU General Public License v3+.

ı	Avviso:	Questo prodotto include software sviluppato dalla Apache Software Foun-
		dation (http://www.apache.org/).

KOST-Val si basa sui seguenti componenti non modificati di altri produttori, integrati direttamente nel codice sorgente di KOST-Val:

Programma/componente di terzi	Ver- sione	Licenza
3-Heights™ PDF/A Validator API		
http://www.pdf-tools.com	6.27.2.1	Vedere capitolo 9.1
Apache Commons http://commons.apache.org/		Licenza Apache 2.0
- commons-logging-1.2.jar	1.2	
- commons-io-2.13.jar	2.13	
Apache Xerces2 http://xerces.apache.org/	2.12.2	Licenza Apache 2.0
BadPeggy http://coderslagoon.com/	2.0	Licenza GPL v3
Jdom http://www.jdom.org/	2.0.6.1	Licenza jdom
Jhove https://jhove.openpreservation.org/	1.28	Licenza LGPL v2.1
API del framework Spring		
https://spring.io/projects/spring-framework	5.3.19	Licenza Apache 2.0
zip64 <u>http://sourceforge.net/projects/zip64file/</u>	1.02	Licenza GPL v2+

KOST-Val si basa sui seguenti componenti non modificati di altri produttori, che vengono forniti con KOST-Val:

Programma/co	omponente di terzi	Ver- sione	Licenza
egov-validatio	nclient-cli https://www.bit.admin.ch/	1.0.10	Vedere capitolo 9.3
flac	https://xiph.org/flac	1.4.3	Licenza BSD
Jpylyzer	http://jpylyzer.openpreservation.org/	2.1.0	Licenza LGPL v3.0
mkvalidator	https://www.matroska.org/	0.6.0	Licenza BSD
pdfaPilot CLI	https://www.callassoftware.com	12.4.372	Vedere capitolo 9.2
pngcheck http	o://libpng.org/pub/png/apps/pngcheck	3.0.3	Licenza GPL v2
GNU sed	https://www.gnu.org/software/sed	4.4	Licenza GPL v3+
Xmllint	https://xmllint.com/	20630	Licenza MIT

Gli utenti di KOST-Val sono tenuti a rispettare le condizioni di licenza di tutti questi componenti, disponibili nella directory delle licenze di KOST-Val.

9.1 Licenza API 3-Heights™ PDF/A Validator

Per l'utilizzo della versione limitata di 3-Heights™ PDF/A Validator di PDF Tools, CECO ha concordato con PDF Tools il seguente accordo individuale sui termini di licenza generali:

2. Individuelle Vereinbarung

Dieses Vertragsverhältnis regelt die Client-Lizenz zwischen der PDF TOOLS als Lizenzgeber und der KOST als Lizenznehmer gemäss nachfolgenden Spezialbestimmungen:

- PDF Tools AG erteilt für KOST eine kostenfreie OEM-Lizenz für das 3-Heights™ PDF/A Validator API als Zusatzfunktion ihrer eigenen Validator-Software (KOST-Val).
- Die Lizenz schliesst den Gebrauch der Software (KOST-Val) durch Gedächtnisinstitutionen, bestehend aus Archiven oder Bibliotheken, deren Zulieferer und der KOST selbst, ein.
- Der OEM-Lizenzschlüssel, welcher fest in KOST-Val eingebunden ist, darf nicht ausserhalb der Applikation (KOST-Val) verwendet werden.
- Die Lizenz ist zeitlich unbegrenzt, jedoch bezüglich Durchsatz pro Installation begrenzt (72'000 Seiten pro Jahr).
- Für die Verteilung der Software (KOST-Val) an den Anwender ist die KOST zuständig.
- Der First Level Support der Anwender erfolgt durch KOST. Second Level Support Fälle leitet KOST an PDF Tools AG weiter.
- Wenn der Anwender weitergehende Bedürfnisse hat, z.B. höherer Durchsatz, Integration in andere Applikationen etc. kauft er die Software (3-Heights™ PDF/A Validator API) direkt bei PDF Tools AG.
- Die KOST darf weiterhin den Quellcode von KOST-Val Open Source publizieren und KOST-Val gratis und ohne Registrierung abgeben.

I seguenti punti sono decisivi per gli utenti:

- La licenza comprende l'uso del software (KOST-Val) da parte di istituzioni di memoria costituite da archivi o biblioteche, dei loro fornitori e della stessa CECO.
- La chiave di licenza OEM, integrata in modo permanente in KOST-Val, non può essere utilizzata al di fuori dell'applicazione (KOST-Val).
- La licenza è illimitata nel tempo, ma limitata in termini di capacità per installazione (72 000 pagine all'anno).
- Il supporto di primo livello per gli utenti è fornito da CECO. I casi di assistenza di secondo livello vengono inoltrati da CECO a PDF Tools AG.
- Se l'utente ha ulteriori esigenze, ad esempio una maggiore produttività, l'integrazione in altre applicazioni e così via, acquista il software (3-Heights™ PDF/A Validator API) direttamente da PDF Tools AG. Questa licenza viene attivata con "LicenseManager.exe", che esiste già in "KOST-Val\resources\3-Heights_PDF-Validator dll".

Gli utenti di KOST-Val sono tenuti a rispettare il presente contratto di licenza.

9.2 Licenza pdfaPilot CLI

Per l'utilizzo della versione limitata di pdfaPilot CLI di callas, CECO ha concordato con callas il seguente accordo individuale sulle condizioni generali di licenza:

2. Individuelle Vereinbarung

Dieses Vertragsverhältnis regelt die Lizenz zwischen der callas software als Lizenzgeber und der KOST als Lizenznehmer gemäss nachfolgenden Spezialbestimmungen:

- callas software erteilt für die KOST eine kostenfreie Lizenz für callas pdfaPilot CLI für Windows zur innerbetrieblichen Nutzung und zur Integration in ihren eigenen Validator "KOST-Val".
- Die Lizenz schliesst die Distribution von KOST-Val an "Anwender" (Gedächtnisinstitutionen, Archive oder Bibliotheken und deren Zulieferer) ein.
- Für die Distribution von KOST-Val an diese Anwender ist die KOST zuständig und darf KOST-Val auch gratis und ohne Registrierung an diese abgeben.
- Die Lizenz ist zeitlich unbegrenzt, jedoch bezüglich Durchsatz pro Installation begrenzt auf 72'000 Seiten pro Jahr.
- Die KOST darf den eigenen Quellcode von KOST-Val Open Source publizieren, callas pdfaPilot CLI ist hiervon ausgenommen.
- First Level Support der Anwender erfolgt durch die KOST. Second Level Support leistet callas software gegenüber der KOST.

I seguenti punti sono decisivi per gli utenti:

- La licenza include la distribuzione di KOST-Val agli "utenti" (istituzioni della memoria, archivi o biblioteche e loro fornitori).
- La licenza è illimitata nel tempo, ma limitata in termini di produttività per installazione a 72 000 pagine all'anno.
- CECO può pubblicare il proprio codice sorgente di KOST-Val Open Source. La callas pdfaPilot CLI ne è esclusa.
- Il supporto di primo livello per gli utenti è fornito da CECO. Il supporto di secondo livello è fornito dal software callas a CECO.

Gli utenti di KOST-Val sono tenuti a rispettare il presente contratto di licenza.

9.3 Licenza egov-validationclient-cli

Per la convalida delle firme elettroniche nei file PDF/PDF/A è necessaria una licenza per il validatore di firme discrete del governo federale. Questa licenza non è inclusa in KOST-Val. Ulteriori informazioni sono disponibili all'indirizzo https://www.bit.ad-min.ch/bit/it/home/themen/elektronische-signatur/anzeigen-und-validieren/der-diskrete-validator-web-service.html.

KOST-Val 2.2.0.0 non esegue ancora la convalida della firma. Finora è stato programmato solo il riconoscimento, che non richiede una licenza.

10 Appendice

10.1 Struttura del programma

KOST-Val è stato sviluppato in base ai seguenti requisiti:

Requisiti funzionali:

I risultati (compresi i messaggi di incongruenza o di errore) vengono emessi per ogni fase e scritti in un file di log di convalida.

Il risultato della convalida complessiva (file corretto/incorretto) è anche emesso e visibile nell'exit status (o valore di uscita)

del programma, in modo che la convalida possa essere integrata in una catena di elaborazione automatizzata. L'exit status può assumere i seguenti valori:

0 tutto OK

1 errore nella chiamata al programma

2 la convalida non è stata superata

Le singole fasi di convalida/controllo vengono eseguite una dopo l'altra. Se possibile, la convalida continua anche in caso di errori, per ridurre il numero di cicli di correzione.

Requisiti non funzionali

Per compiti speciali si utilizzano programmi esterni o framework Java corrispondenti.

L'applicazione è modulare, in modo da poter aggiungere uno o più moduli di convalida senza grandi sforzi.

L'uscita del registro/programma consente di leggere facilmente il risultato della singola convalida e quindi di utilizzare lo strumento in una catena di processi,

L'output della console è limitato al nome del tipo di convalida, al risultato complessivo "valido" o "non valido" e al percorso del file. Tutte le informazioni aggiuntive sono elencate nel file di log.

10.2 Convalida del formato del principio funzionale

