Stefania Bergamasco - Istat









Giano contro Crono: il caso Istat per la conservazione delle banche dati

con la collaborazione di Simone Coccia e Raffaella Cianchetta - Istat

Stefania Bergamasco - Istat









Giano contro Crono: il caso Istat per la conservazione delle banche dati

con la collaborazione di Simone Coccia e Raffaella Cianchetta - Istat

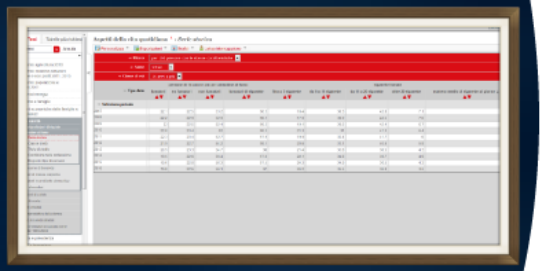
Il quadro di insieme



siti informativi
(news, report, file, tabelle,...



banche dati



sistemi informativi



Strumenti Informazioni Regione



eventi istituzionali & scientifici

info rilevazioni

I.Stat accedi a tutti i dati

notizie

ei consumatori e delle imprese
e indice del clima di fiducia consumatori
114,3; fiducia imprese da 109,1 a 108,8
stampa, lunedì 27 novembre 2017

za sulle donne
Dipartimento delle Pari Opportunità
sponibile un quadro informativo integrato
nza contro le donne in Italia
to 25 novembre 2017

e ordinativi dell'industria
re 2017 il fatturato aumenta del 5,2% e gli
el 4,5% in un anno
stampa, venerdì 24 novembre 2017

dei servizi
estre 2017 l'indice del fatturato dei servizi
2,6% su base annua
stampa, venerdì 24 novembre 2017

io estero extra Ue
merciale ottobre 2017 pari a +4.554
uro (+3.803 milioni a ottobre 2016)
stampa, giovedì 23 novembre 2017

lle imposte ambientali

>> Prezzi al consumo
>> Commercio al dettaglio
>> Occupati
>> Tasso di disoccupazione
>> Prezzi alla produzione
>> Retribuzioni
>> Produzione industriale
>> Prodotto interno lordo

Prezzi al consumo



variazioni % tendenziali, al lordo dei tabacchi

In evidenza

- **Calendario** delle diffusioni e degli eventi
- **Mappa dei rischi** dei comu
- Piattaforma on line per la **formazione statistica**

Quadri informativi

- Sistema informativo **#giov**
- Sistema informativo **#anzi**
- **Immigrati** e nuovi cittadini
- Sistema sulle **professioni**
- **Congiuntura** economica

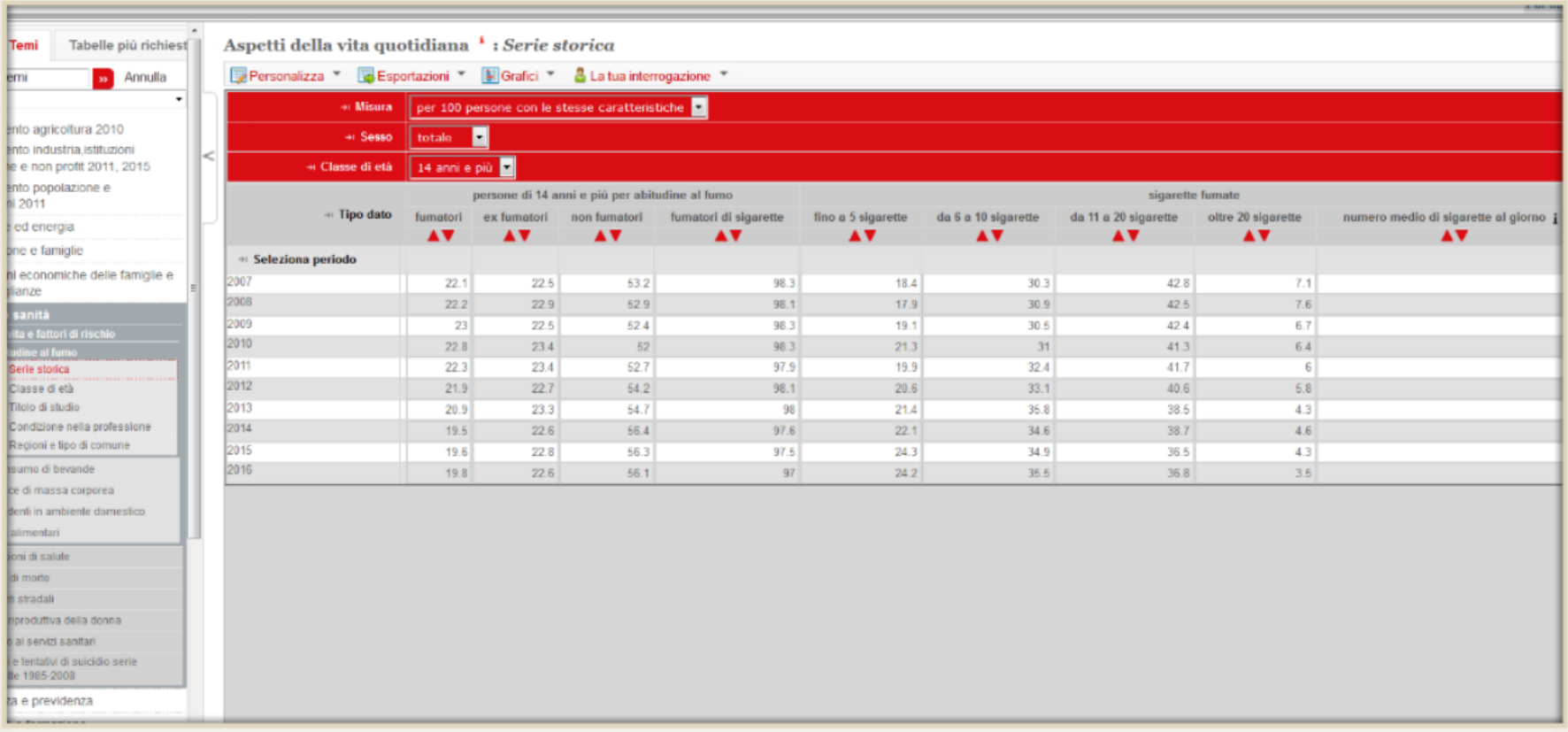
STATISTICA per TUTTI

Indici per aggiorn AFFITTI e ASSEGNI

bes BENESSERE EQUO E SOSTENIBILE

CENSIMEN PERMANE

noi italia edizio



Temi | Tabelle più richiest

Annulla

ento agricoltura 2010
ento industria,istituzioni
e e non profit 2011, 2015
ento popolazione e
i 2011
ed energia
one e famiglie
ni economiche delle famiglie e
ianze
sanità
ita e fattori di rischio
udine al fumo
Serie storica
Classe di età
Titolo di studio
Condizione nella professione
Regioni e tipo di comune
sumo di bevande
ce di massa corporea
denti in ambiente domestico
alimentari
oni di salute
di morte
e stradali
riproduttiva della donna
ai servizi sanitari
e tentativi di suicidio serie
ie 1985-2008
ta e previdenza

Aspetti della vita quotidiana : *Serie storica*

Personalizza | Esportazioni | Grafici | La tua interrogazione

Misura: per 100 persone con le stesse caratteristiche
Sesso: totale
Classe di età: 14 anni e più

Tipo dato	persone di 14 anni e più per abitudine al fumo				sigarette fumate				
	fumatori	ex fumatori	non fumatori	fumatori di sigarette	fino a 5 sigarette	da 6 a 10 sigarette	da 11 a 20 sigarette	oltre 20 sigarette	numero medio di sigarette al giorno
Seleziona periodo									
2007	22.1	22.5	53.2	98.3	18.4	30.3	42.8	7.1	
2008	22.2	22.9	52.9	98.1	17.9	30.9	42.5	7.6	
2009	23	22.5	52.4	98.3	19.1	30.5	42.4	6.7	
2010	22.8	23.4	52	98.3	21.3	31	41.3	6.4	
2011	22.3	23.4	52.7	97.9	19.9	32.4	41.7	6	
2012	21.9	22.7	54.2	98.1	20.6	33.1	40.6	5.8	
2013	20.9	23.3	54.7	98	21.4	35.8	38.5	4.3	
2014	19.5	22.6	56.4	97.6	22.1	34.6	38.7	4.6	
2015	19.6	22.8	56.3	97.5	24.3	34.9	36.5	4.3	
2016	19.8	22.6	56.1	97	24.2	35.5	36.8	3.5	


in cifre
istat.it
Guida al sito
Contenuti
Settori
Cerca
rrogazione dati
duzione libraria
tecipazione culturale
videnza
Annuari delle statistiche
turali: 2010-2011
tadini e nuove tecnologie
Italia. 100 statistiche
capire il Paese - Cultura
atrimonio museale non
tale
duzione e lettura di libri
talia
talia in 150 anni.
editoria
biblioteche
archivi
musei
sport
spettacoli
cinema
radio e TV
varie
Cultura in cifre
un sistema
informativo dedicato
alle statistiche
culturali
che contiene dati e
informazioni sulla
produzione, la
distribuzione e la

basi statiche (non più aggiornate)



banche dati dinamiche (in continuo aggiornamento)



Istat

DaWinci.istat.it

Censimento 2001

14° Censimento Generale della Popolazione e delle Abitazioni



Dati definitivi

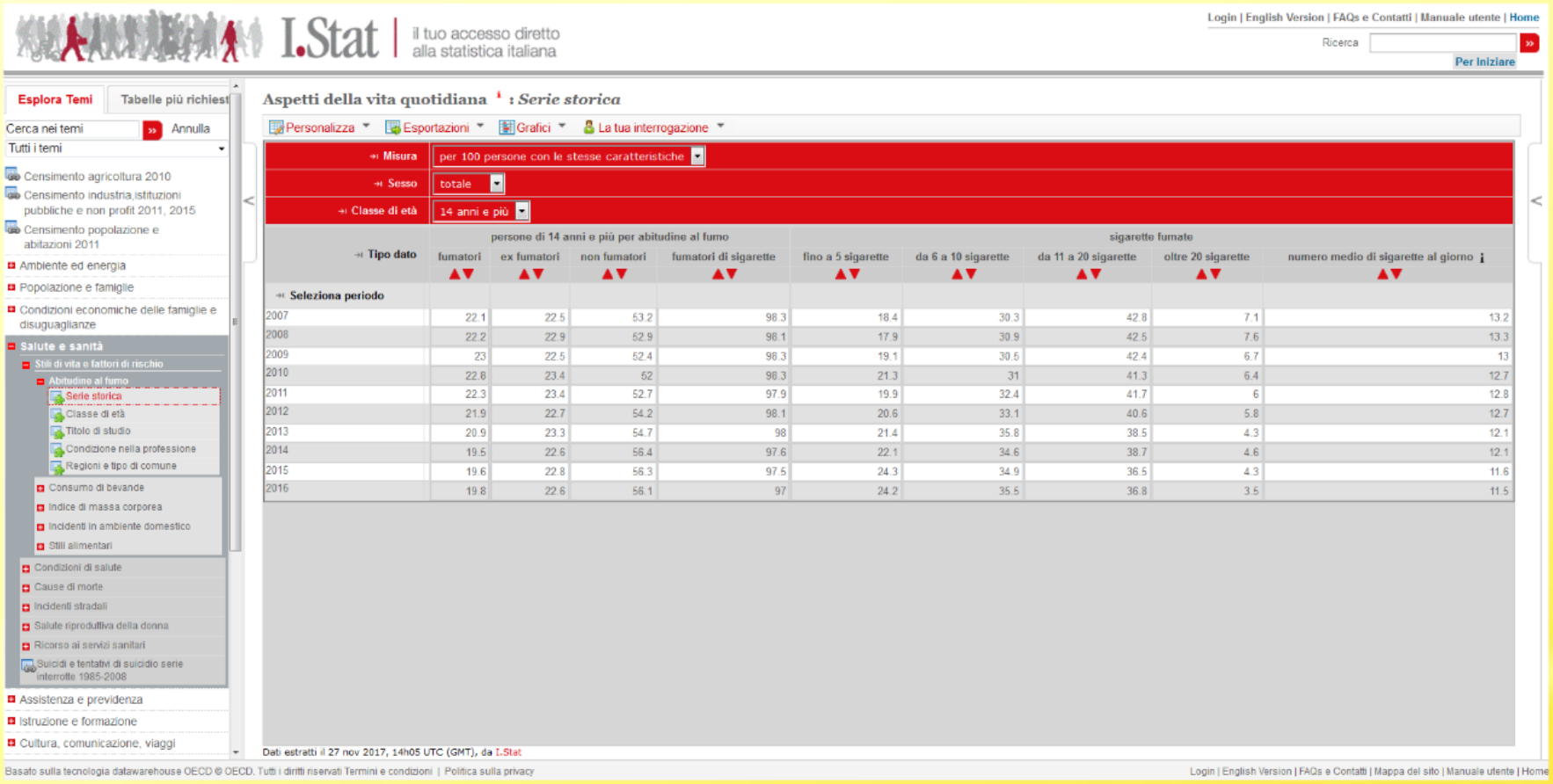
Rilascio del 21 luglio 2005 - comprendente anche le informazioni sui Sistemi Locali del Lavoro.

Popolazione legale (English version)

Pubblicati sulla G.U. Serie Generale n.81 del 7 Aprile 2003, Supplemento Ordinario n.54. Elaborati a seguito della completa acquisizione delle informazioni desumibili dai modelli di rilevazione.

Primi risultati

Dati **provvisori** elaborati prima del completamento delle operazioni censuarie
(rilasciati il 27 Marzo 2002).


I.Stat
il tuo accesso diretto alla statistica italiana
Login | English Version | FAQs e Contatti | Manuale utente | Home
Ricerca
Per Iniziare
Esplora Temi
Tabelle più richieste
Cerca nei temi
Annulla
Tutti i temi
Censimento agricoltura 2010
Censimento industria,istituzioni pubbliche e non profit 2011, 2015
Censimento popolazione e abitazioni 2011
Ambiente ed energia
Popolazione e famiglie
Condizioni economiche delle famiglie e disuguaglianze
Salute e sanità
Stili di vita e fattori di rischio
Abitudine al fumo
Serie storica
Classe di età
Titolo di studio
Condizione nella professione
Regioni e tipo di comune
Consumo di bevande
Indice di massa corporea
Incidenti in ambiente domestico
Stili alimentari
Condizioni di salute
Cause di morte
Incidenti stradali
Salute riproduttiva della donna
Ricorso ai servizi sanitari
Suicidi e tentativi di suicidio serie interrotte 1985-2008
Assistenza e previdenza
Istruzione e formazione
Cultura, comunicazione, viaggi
Aspetti della vita quotidiana : Serie storica
Personalizza
Esportazioni
Grafici
La tua interrogazione
Misura
per 100 persone con le stesse caratteristiche
Sesso
totale
Classe di età
14 anni e più
Tipo dato
persone di 14 anni e più per abitudine al fumo
sigarette fumate
fumatori
ex fumatori
non fumatori
fumatori di sigarette
fino a 5 sigarette
da 6 a 10 sigarette
da 11 a 20 sigarette
oltre 20 sigarette
numero medio di sigarette al giorno
Seleziona periodo
2007 22.1 22.5 53.2 98.3 18.4 30.3 42.8 7.1 13.2
2008 22.2 22.9 52.9 98.1 17.9 30.9 42.5 7.6 13.3
2009 23 22.5 52.4 98.3 19.1 30.5 42.4 6.7 13
2010 22.6 23.4 52 98.3 21.3 31 41.3 6.4 12.7
2011 22.3 23.4 52.7 97.9 19.9 32.4 41.7 6 12.8
2012 21.9 22.7 54.2 98.1 20.6 33.1 40.6 5.8 12.7
2013 20.9 23.3 54.7 98 21.4 35.8 38.5 4.3 12.1
2014 19.5 22.6 56.4 97.6 22.1 34.6 38.7 4.6 12.1
2015 19.6 22.8 56.3 97.5 24.3 34.9 36.5 4.3 11.6
2016 19.8 22.6 56.1 97 24.2 35.5 36.8 3.5 11.5
Dati estratti il 27 nov 2017, 14h05 UTC (GMT), da I.Stat
Basato sulla tecnologia datawarehouse OECD © OECD. Tutti i diritti riservati Termini e condizioni | Politica sulla privacy
Login | English Version | FAQs e Contatti | Mappa del sito | Manuale utente | Home

Il quadro di insieme



nella definizione di **regole**, **metodi** e **strumenti** per l'archiviazione delle **banche dati** (che rappresentano la specificità dell'Istat) **statiche** (considerando che il dinamismo è una sequenza di attimi statici

L'archiviazione ha l'obiettivo di conservare la *"memoria della cultura e della vita sociale italiana" (Legge 15 aprile 2004, n. 106)* ... per l'Istat, che produce *Informazione* questo significa non solo archiviare i dati ma anche il modo in cui tali dati vengono diffusi cioè ... archiviare tre differenti oggetti:

- ✔ i dati,
- ✔ i piani di presentazione dei dati,
- ✔ le funzionalità con le quali l'utente interagisce con la banca dati

Accanto alle riflessioni relative al come poter archiviare una banca dati ai fini "documentativi", l'Istat si è posto la domanda sulla opportunità o meno di mantenere on-line le banche dati (o in genere i sistemi informativi) non più aggiornate.

la complessità operativa e l'alto grado di effort necessari per la manutenzione nel tempo dei prodotti

la necessità di rispondere alla richiesta dell'utenza di governo del "diluvio di dati" piuttosto che "dell'affollamento informativo"

le nuove opportunità offerte dalle tecnologie in temini di archiviazione delle banche dati e sitemi informativi

stanno portando l'Istat a scegliere la strada della **dismissione** dei prodotti piuttosto che quella della **manutenzione "live"** (dopo opportuna archiviazione e indicazione all'utenza sul come e dove reperire le informazioni) ... modificando un po' il concetto di pubblicazione dei dati e del rapporto con l'utenza .... pubblicazioni web <> pubblicazioni su stampa

Contemporanemante ...

la difficoltà tecnologica nella archiviazione dei sistemi pubblicati alcuni anni fa

la necessità di rispondere alla esigenza di archiviazione in maniere completa ma semplice

sta portando l'Istat a scegliere di archiviare ciascuna componente delle banche dati (dati, funzioni, i piani di presentazione") secondo logiche precise ... *vediamo quali*

I dettagli dell'archiviazione

I dati

L'idea di base, relativamente alla componente dati, non è quella di **archiviare** quanto quella di "**spostare** i dati nelle **banche dati live**"





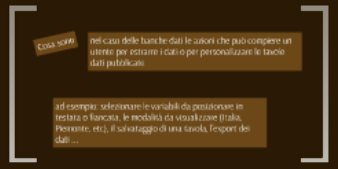
I piani di presentazione dei dati

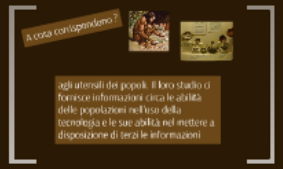






Le funzionalità







I dettagli dell'archiviazione

I dati

L'idea di base, relativamente alla componente dati, non è quella di **archiviare** quanto quella di "**spostare** i dati nelle **banche dati live**"





I piani di presentazione dei dati

Cosa sono ?

L'organizzazione in "topics" della banca dati e la struttura di default, se esiste, delle tavole dati

Alle locandine, alla moda, al packaging ... cioè al "confezionamento" ed al loro stile



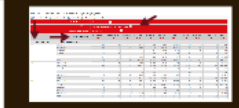


I piani di presentazione dei dati

Cosa sono ?

L'organizzazione in "topics" della banca dati e la struttura di default, se esiste, delle tavole dati





A cosa corrispondono ?

Alle locandine, alla moda, al packaging ... cioè al "confezionamento" ed al loro stile



Alcuni esempi

"I nati vivi" sono collocabili in popolazione o in sanità?

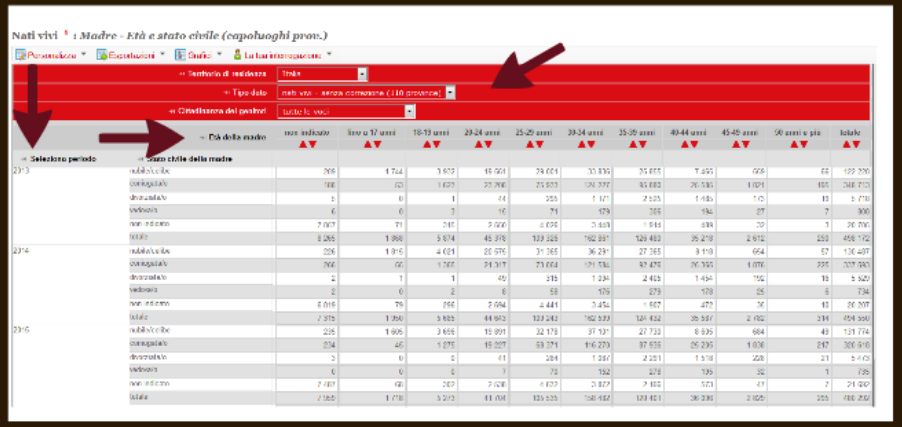
Perchè nella tabella relativa ai matrimoni tra le caratteristiche della sposa in testata si pone l'età e in fiancata si pone lo stato civile piuttosto che la "cittadinanza della coppia"?

Le funzionalità

Cosa sono ?

L'organizzazione in "topics" della banca dati e la struttura di default, se esiste, delle tavole dati

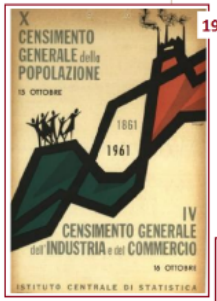




A cosa corrispondono ?

Alle locandine, alla moda, al packaging ...
cioè al "confezionamento" ed al loro stile


1911



1961



1936



2001

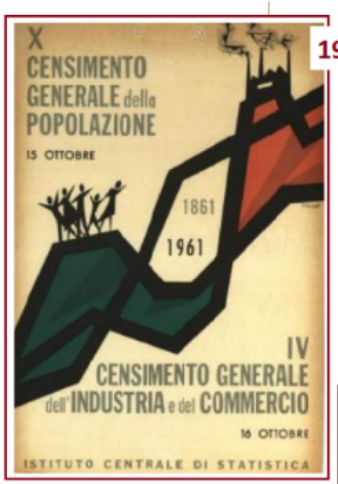


2011

Alle locandine, alla moda, al packaging ...
cioè al "confezionamento" ed al loro stile



1911



1961



1936



2001



2011

Alcuni esempi

"I nati vivi" sono collocabili in popolazione o in sanità?

Perchè nella tabella relativa ai matrimoni tra le caratteristiche della sposa in testata si pone l'età e in fiancata si pone lo stato civile piuttosto che la "cittadinanza della coppia"?







Perchè nella tabella relativa ai matrimoni tra le caratteristiche della sposa in testata si pone l'età e in fiancata si pone lo stato civile piuttosto che la "cittadinanza della coppia"?

Le funzionalità

Cosa sono

nel caso delle banche dati le azioni che può compiere un utente per estrarre i dati o per personalizzare le tavole dati pubblicate

ad esempio: selezionare le variabili da posizionare in testata o fiancata, le modalità da visualizzare (Italia, Piemonte, etc), il salvataggio di una tavola, l'export dei dati ...

A cosa corrispondono ?





agli utensili dei popoli. Il loro studio ci fornisce informazioni circa le abilità delle popolazioni nell'uso della tecnologia e le sue abilità nel mettere a disposizione di terzi le informazioni

Come possono essere archiviate ?

- conservando la guida utente (l'equivalente di un libretto di istruzioni)
- conservando un video tutorial che spieghi le funzioni disponibili
- conservando un "archivio interattivo" che consenta all'utente una simulazione dell'uso delle funzioni disponibili

Cosa sono

nel caso delle banche dati le azioni che può compiere un utente per estrarre i dati o per personalizzare le tavole dati pubblicate

ad esempio: selezionare le variabili da posizionare in testata o fiancata, le modalità da visualizzare (Italia, Piemonte, etc), il salvataggio di una tavola, l'export dei dati ...

A cosa corrispondono ?





agli utensili dei popoli. Il loro studio ci fornisce informazioni circa le abilità delle popolazioni nell'uso della tecnologia e le sue abilità nel mettere a disposizione di terzi le informazioni

Come possono essere archiviate ?

conservando la guida utente (l'equivalente di un libretto di istruzioni)



conservando un video tutorial che spieghi le funzioni disponibili



conservando un "archivio interattivo" che consenta all'utente una simulazione dell'uso delle funzioni disponibili



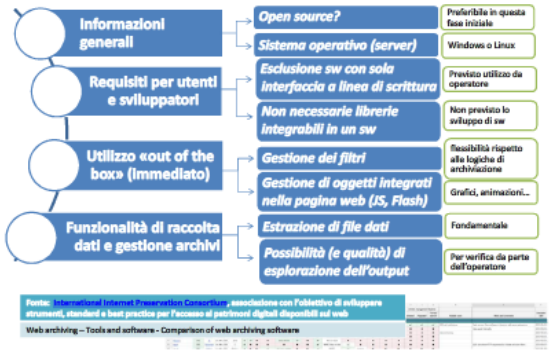
C'è differenza in ciò che si può comprendere di un popolo studiando i suoi manufatti attraverso dei disegni o dal vivo. La sperimentazione degli oggetti offre informazioni circa la qualità del manufatto, la pesantezza, la levigatezza e, quindi, informazioni anche sulla capacità artigianale di un popolo. Analogamente l'esperienza diretta delle funzioni offrirà informazioni, ad esempio, sulle abilità concrete degli utenti nell'interazione macchina/uomo nella ricerca dei dati



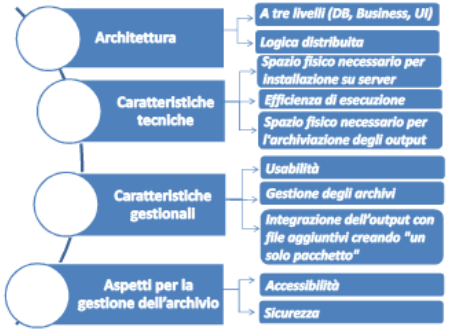


Le soluzioni tecnologiche

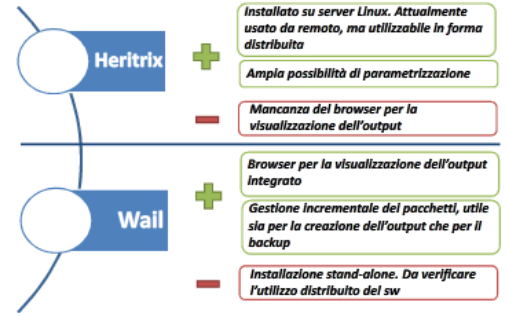
I criteri selezionati



I criteri aggiuntivi individuati

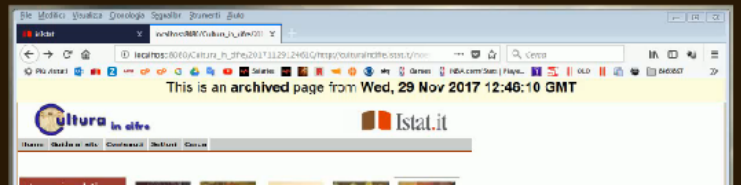


Lo stato dell'arte



Gli step seguiti ed I primi risultati

I nostri step	Acquisizione e studio della griglia comparativa
	Attribuzione di un «peso» ai criteri di valutazione
	Individuazione dei software rispondenti alle nostre esigenze
	Aggiunta di criteri per la valutazione dei software



I criteri selezionati

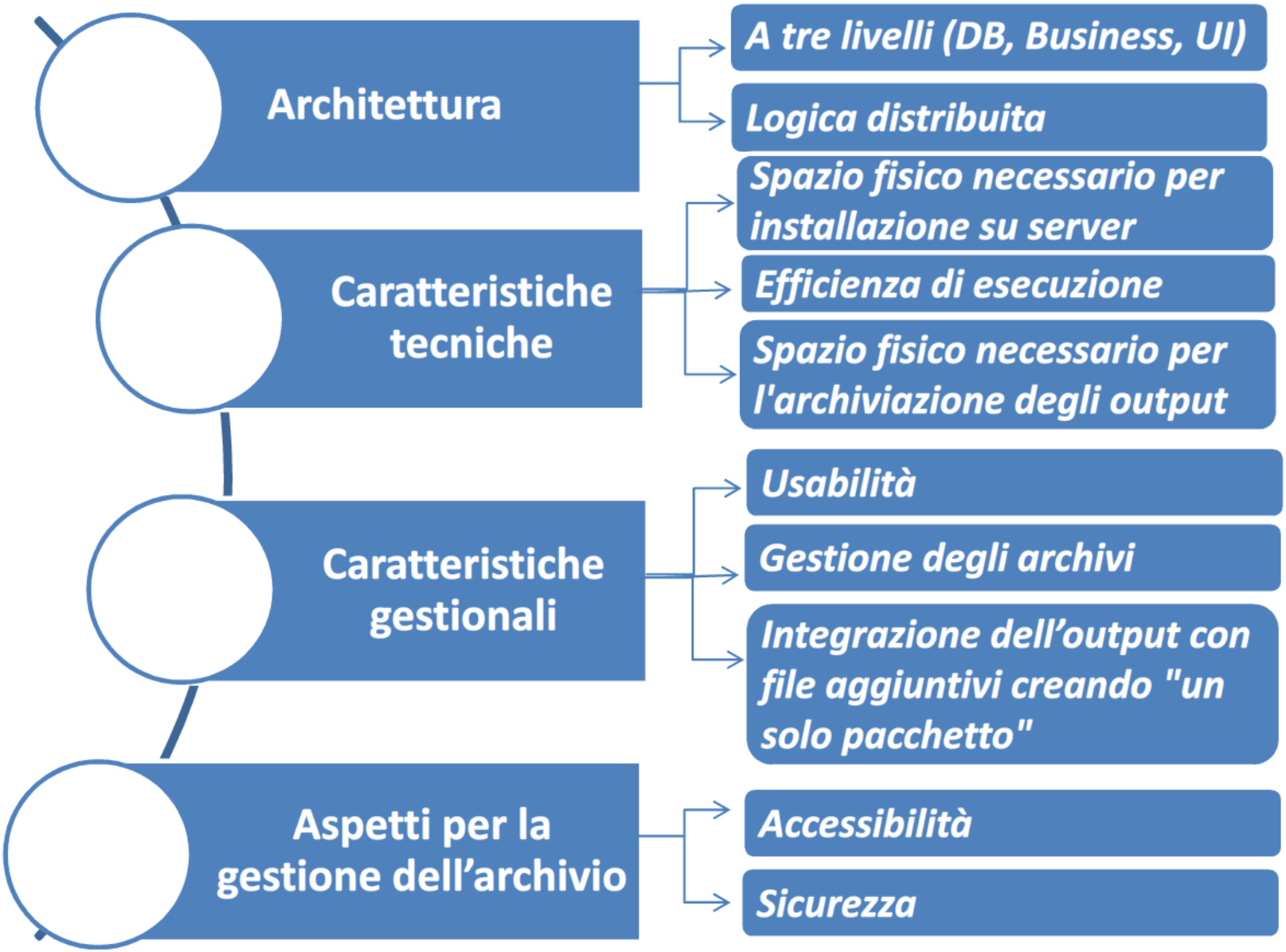
Informazioni generali	***Open source?***	Preferibile in questa fase iniziale
	Sistema operativo (server)	Windows o Linux
Requisiti per utenti e sviluppatori	***Esclusione sw con sola interfaccia a linea di scrittura***	Previsto utilizzo da operatore
	Non necessarie librerie integrabili in un sw	Non previsto lo sviluppo di sw
Utilizzo «out of the box» (immediato)	***Gestione dei filtri***	flessibilità rispetto alle logiche di archiviazione
	Gestione di oggetti integrati nella pagina web (JS, Flash)	Grafici, animazioni…
Funzionalità di raccolta dati e gestione archivi	***Estrazione di file dati***	Fondamentale
	Possibilità (e qualità) di esplorazione dell'output	Per verifica da parte dell'operatore

Fonte: International Internet Preservation Consortium, associazione con l'obiettivo di sviluppare strumenti, standard e best practice per l'accesso ai patrimoni digitali disponibili sul web

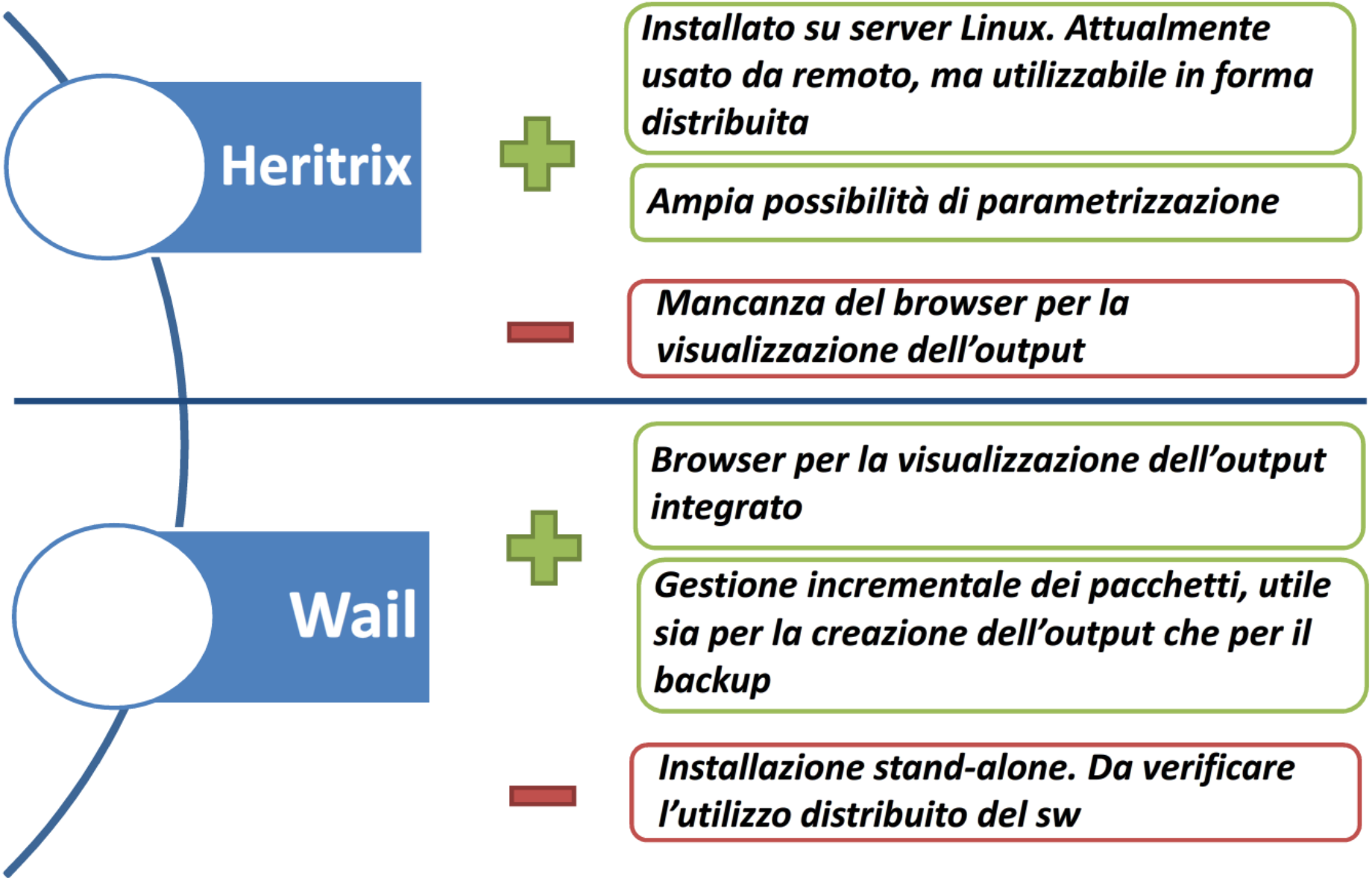
Web archiving – Tools and software - Comparison of web archiving software

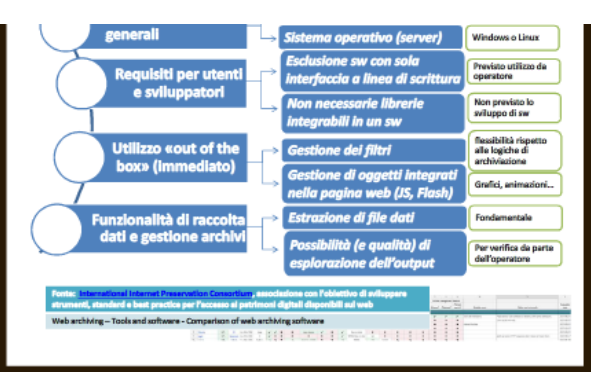


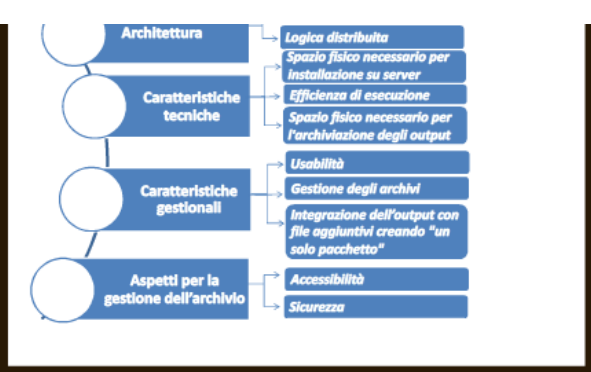
I criteri aggiuntivi individuati

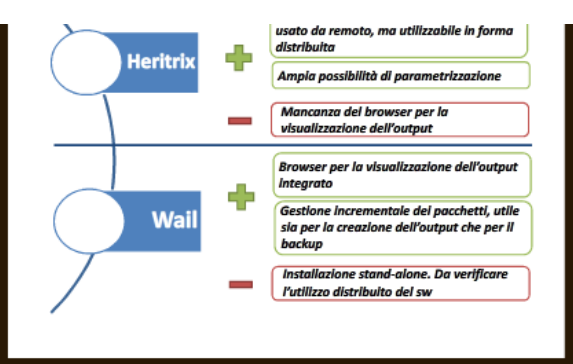


Lo stato dell'arte



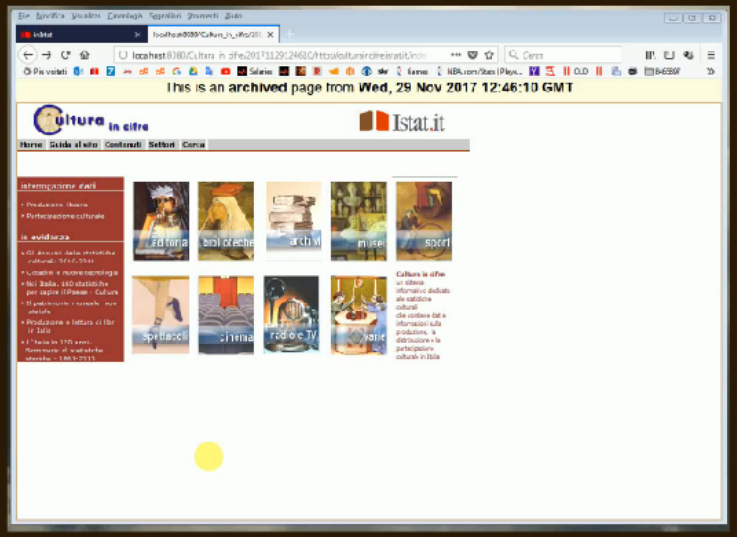






Gli step seguiti ed I primi risultati

I nostri step	Acquisizione e studio della griglia comparativa
	Attribuzione di un «peso» ai criteri di valutazione
	Individuazione dei software rispondenti alle nostre esigenze
	Aggiunta di criteri per la valutazione dei software
	Individuazione dei software
Risulati	Creazione di pacchetti (warc, log) per culturaincifre.istat.it tramite i sw Heritrix e Wail
	Navigazione dei warc prodotti tramite browser integrato nel sw Wail e tramite sw Webrecorderplayer (electron)
	Creazione video della navigazione del warc prodotto tramite Wail effettuata con il sw apowersoft



I nostri step

- Acquisizione e studio della griglia comparativa
- Attribuzione di un «peso» ai criteri di valutazione
- Individuazione dei software rispondenti alle nostre esigenze
- Aggiunta di criteri per la valutazione dei software
- Individuazione dei software

Risulati

- Creazione di pacchetti (warc, log) per culturaincifre.istat.it tramite i sw Heritrix e Wail
- Navigazione dei warc prodotti tramite browser integrato nel sw Wail e tramite sw Webrecorderplayer (electron)
- Creazione video della navigazione del warc prodotto tramite Wail effettuata con il sw apowersoft

File Modifica Visualizza Cronologia Segnalibri Strumenti Aiuto

inIstat | localhost:8080/Cultura_in_cifre/201

localhost:8080/Cultura_in_cifre/20171129124610/http://culturaincifre.istat.it/inde

Cerca

Più visitati | Salaries | Games | NBA.com/Stats | Playe... | OLD | 8465897

This is an **archived** page from **Wed, 29 Nov 2017 12:46:10 GMT**

Cultura in cifre

Istat.it

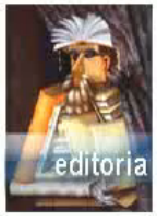
Home | Guida al sito | Contenuti | Settori | Cerca

interrogazione dati

- Produzione libraria
- Partecipazione culturale

in evidenza

- Gli Annuari delle statistiche culturali: 2010-2011
- Cittadini e nuove tecnologie
- Noi Italia. 100 statistiche per capire il Paese - Cultura
- Il patrimonio museale non statale
- Produzione e lettura di libri in Italia
- L'Italia in 150 anni. Sommario di statistiche storiche - 1861-2010



















Cultura in cifre
un sistema informativo dedicato alle statistiche culturali
che contiene dati e informazioni sulla produzione, la distribuzione e la partecipazione culturale in Italia

Conclusioni

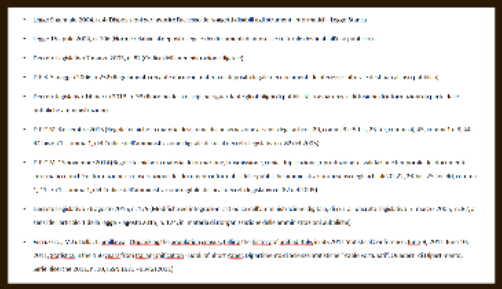
- conservare le banche dati significa non solo prestare l'attenzione alla conservazione del dato ma anche all'informazione che è stata diffusa ed alle modalità con le quali è stata diffusa

- conservare le banche dati significa organizzare e disporre di strumenti che facilitino la lettura storiografica dell'informazione (dati +funzioni+piani di presentazione)

- si ravvisa la necessità di norme e standard che definiscano le regole per produrre software e banche dati che siano facilmente archiviabili

- si ravvisa la necessità di inserire gli aspetti di accessibilità e sicurezza all'interno delle riflessioni sullo sviluppo di tecnologie per l'archiviazione delle banche dati

Cenni bibliografici e riferimenti



Stefania Bergamasco - bergamas@istat.it
Simone Coccia - sicoccia@istat.it
Raffaella Cianchetta - cianchet@istat.it

Grazie per l'attenzione

Conclusioni

- conservare le banche dati significa non solo prestare l'attenzione alla conservazione del dato ma anche all'informazione che è stata diffusa ed alle modalità con le quali è stata diffusa

- conservare le banche dati significa organizzare e disporre di strumenti che facilitino la lettura storiografica dell'informazione (dati +funzioni+piani di presentazione)

- si ravvisa la necessità di norme e standard che definiscano le regole per produrre software e banche dati che siano facilmente archiviabili

- si ravvisa la necessità di inserire gli aspetti di accessibilità e sicurezza all'interno delle riflessioni sullo sviluppo di tecnologie per l'archiviazione delle banche dati

Cenni bibliografici e riferimenti

- Legge 9 gennaio 2004, n. 4 (Disposizioni per favorire l'accesso dei soggetti disabili agli strumenti informatici) - Legge Stanca
- Legge 15 aprile 2004, n. 106 (Norme relative al deposito legale dei documenti di interesse culturale destinati all'uso pubblico)
- Decreto Legislativo 7 marzo 2005, n. 82 (Codice dell'amministrazione digitale)
- D.P.R. 3 maggio 2006, n.252 (Regolamento recante norme in materia di deposito legale dei documenti di interesse culturale destinati all'uso pubblico)
- Decreto legislativo 14 marzo 2013, n. 33 (Riordino della disciplina riguardante gli obblighi di pubblicità, trasparenza e diffusione di informazioni da parte delle pubbliche amministrazioni)
- D.P.C.M. 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)
- D.P.C.M. 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)
- Decreto Legislativo 26 agosto 2016, n. 179 (Modifiche ed integrazioni al Codice dell'amministrazione digitale, di cui al decreto legislativo 7 marzo 2005, n. 82, ai sensi dell'articolo 1 della legge 7 agosto 2015, n. 124, in materia di riorganizzazione delle amministrazioni pubbliche)
- Ferruzzi C., M.G. Bella, I. Smilkova - Organizing the population census, telling the history of unified Italy, in SIS 2011 Statistical Conference, June 8, 2011-June 10, 2011, Statistics in the 150 years from Italian Unification - Book of Short Paper, Dipartimento di Scienze Statistiche "Paolo Fortunati", Quaderni di Dipartimento, Serie Ricerche 2011, n. 10, ISSN 1973 - 9346 (2011)

Stefania Bergamasco - bergamas@istat.it
Simone Coccia - sicoccia@istat.it
Raffaella Cianchetta - cianchet@istat.it

Grazie per l'attenzione

Stefania Bergamasco - Istat









Giano contro Crono: il caso Istat per la conservazione delle banche dati

con la collaborazione di Simone Coccia e Raffaella Cianchetta - Istat