



UNIVERSITÀ
DI SIENA
1240



Towards a protocol for the curation and dissemination of vulnerable people archives

Silvia Calamai, Aleksei Kelli, Chiara Kolletzek

Rationale of the talk

- How to define a protocol for the curation and dissemination of speech archives, which appear to have – *de jure* – the highest restrictions on their curation and dissemination
- How to find a balance between the rights of the recorded people (and their heirs) and the right of information and the preservation of memory

The speech archive of Anna Maria Bruzzone

(1925-2015)

A.M. Bruzzone, *Ci chiamavano matti.*
Storie da un ospedale psichiatrico,
Torino, Einaudi, 1979



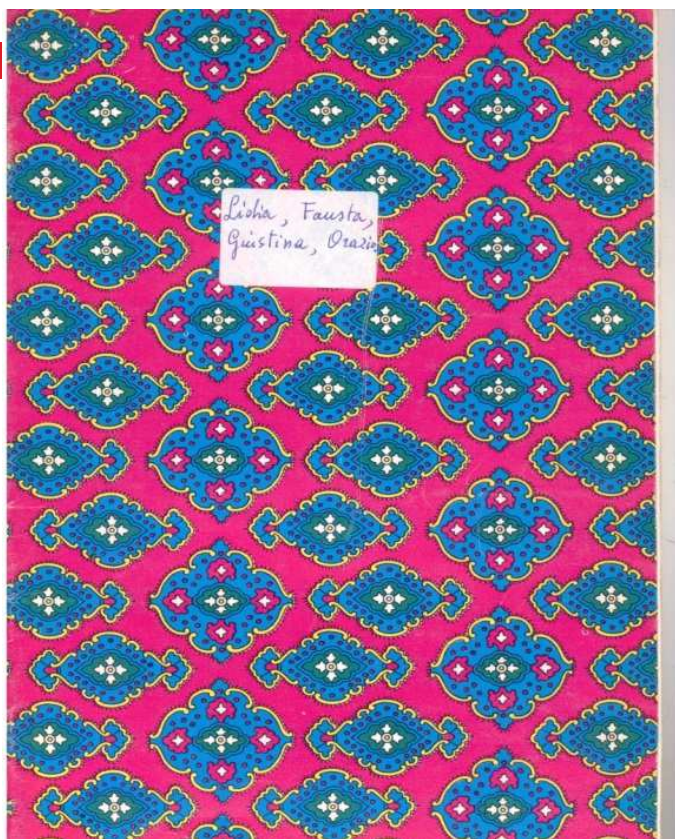
What we discovered

- 36 tapes accompanied by the handwritten and the typewritten transcriptions of all the interviews.
- In addition to the complete transcriptions, there are different versions that show all the work of editing made by A.M. Bruzzone so that the interviews could be suited for publishing.

What is now possible to investigate

- ❖ understand, and examine the changes undergone by an interview from the moment it was recorded on tape to its publication in the book, through the comparative study of all the available documents: the original audio recording, the first, handwritten transcription, the typewritten transcription, the edited version and, finally, the one published in the book
- ❖ associate the oral life stories with the medical diagnosis of every single inpatient (preserved in the Historical Archive of the Arezzo psychiatric hospital), since the real name and not the pseudonym has been found in the box of every single tape.

The handwritten transcriptions



Lidia, Fausta,
Giustina, Grazia

17

[Le ha detto: prova...]
| ^{allora} io tornavo al fasciglione, poi mi misero
in ~~vigilanza~~, recai volontaria al fasciglione,
poi mi misero in vigilanza, volontaria, poi
mi riportarono dopo un mese e venti giorni,
[così è la prima volta...]
e poi ~~non~~ hanno messo a tradimento, ancora
sono qui.
] Non sapui che cosa avrebbe. Se ha qualcosa
da raccontarmi delle vite dell'ospedale, non
lo so, delle differenze che c'è ^{che} adesso a
alcuni occhi fa, per esempio, in ospedale...]
ma... Come punto di vista è molto meglio, ora, in osp
dale, si ~~decano~~ ^{osserva ribatte} di più, per chi non ha nessuno
è molto meglio,
] cioè si accostano... si ha qualche soldo]
il lavoro ~~se~~
] viene pagato]
viene ricambiato, è, molto meglio, si ~~ne~~

The typewritten transcriptions

[E il babbo e la mamma erano poi già morti?]
No, no no, poi morì il babbo e dopo morto il babbo io mi son trovata, insomma, sentita male dal servizio, mi son sentita male. Un anno sono stata a casa, mi fecero ... m'hanno portata in ospedale, però avevo una malattia abbastanza forte, mi sentivo molto, molto male, perché io sentivo proprio le voci, insomma, sentivo per sone che parlavano e dicevo ai miei: Sono qui, sono qui, sentite come cantano! Ma non è vero, non c'è nessuno! Ma a me sembrava proprio che venissero ^{dei costei,} insomma.

[Sì sì, lo so.]
Mi si avvicinava in continuazione queste voci.
Le voci in continuazione, ^{ci} avevo queste voci che ci ammazzavano, e il giorno dopo, insomma, non ci sarà più, non ci sarà più, e mi sono trovata in ospedale.

[Qui ad Arezzo?]
Ad Arezzo, sì, m'hanno portato in vigilanza. Le pastioche non le volevo, perché mi faceva effetto, insomma, pigliare le pastioche in quell'ostia e si vedeva quei colori, blu, rosso, insomma, quei colori non li potevo buttar giù e ~~mi~~ mi facevano le punture perché non prendevo le pastioche, non per condizioni...
Senonché dopo mi hanno fatto l'insulina perché ...

[E il babbo e la mamma erano poi già morti?]
No, no no, poi morì il babbo e dopo morto il babbo io mi son trovata, insomma, sentita male dal servizio, mi son sentita male. Un anno sono stata a casa, mi fecero ... m'hanno portata in ospedale, avevo una malattia abbastanza forte, mi sentivo molto, molto male, perché io sentivo proprio le voci, ^{insomma,} sentivo per sone che parlavano e dicevo ai miei: Sono qui, sono qui, sentite come cantano! Ma non è vero, non c'è nessuno! Ma a me sembrava proprio che venissero ^{dei costei,} insomma.

[Sì sì, lo so.]
Mi si avvicinava in continuazione queste voci,
Le voci in continuazione, ^{ci} avevo queste voci che ci ammazzavano, e il giorno dopo, insomma, non ci sarà più, non ci sarà più, e mi sono trovata in ospedale, ad

[Qui ad Arezzo?]
Ad Arezzo, sì, m'hanno portato in vigilanza. Le pastioche non le volevo, perché mi faceva effetto, insomma, pigliare le pastioche in quell'ostia e si vedeva quei colori, blu, rosso, ^{insomma,} quei colori non li potevo buttar giù e ~~mi~~ mi facevano le punture ...
Senonché dopo mi hanno fatto l'insulina perché ...

The legal side of the coin

- The curation and dissemination of vulnerable people archives is subject to the personal data regulation, since they contain personal data.

Personal data



- personal data = “any information relating to an identified or **identifiable** natural person (‘data subject’); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person” (GDPR)

Special categories of personal data

Speech archives

- Human voice is considered biometric data (González-Rodríguez et. al. 2008; Jain et. al. 2004).

Speech archives with vulnerable subjects

- health data

Biometric and health data both belong to the special categories of data (sensitive or delicate data).

The challenge



- **How to strike a fair balance between research, dissemination, and protection of privacy**

Two socially relevant interests:

- the protection of personal data
- the transmission of knowledge and freedom of research.

Any solution?

- The GDPR does not apply to the personal data of deceased persons (Recital 27).
- The GDPR also does not apply to anonymous data (= the natural person is not identifiable (Recital 26)).

But...

- Right-holders (e.g. relatives)
- Anonymisation may not correspond to the needs and scope of historical research (historians are interested in “names and faces”).

The 'legal chain' for the re-use of the Archive

- **identification of the real names** of the patients and their matching with the pseudonyms, as attested in the volume (Bruzzone 1979).
- **go back to the interviewees**, also engaging the network of all the people – physicians, nurses, social workers, ordinary citizens – involved in the recent history of the Psychiatric hospital.



Dissemination via TV news

https://www.rainews.it/tgr/toscana/video/2018/09/tos-arezzo-manicomio-bruzzone-88ee6948-a3ab-41fdec303b90cd.html?wt_mc=2.www.fb.tgra3ab-4135-a034-fdec303b90cd.&wt



Acquiring informed explicit consent

- The 'legal chain' could help to assess the possibility and feasibility to obtain detailed and clear informed consents, describing
 - ▣ the aims,
 - ▣ the scope
 - ▣ the positive spill-over effects of the dissemination of such an oral archive.
- If the consent form will be obtained by the right-holders, the oral archive could be finally enjoyed by the research communities and the entire civil society.

The relationship with CLARIN Infrastructure



- Anna Maria Bruzzone Archive will be part of the CLARIN Infrastructure and metadata description will be pursued according to COALA (Schiel, Jackl 2014).
- A mobility grant under the H2020 project CLARIN-PLUS allowed the first author to prepare a feasibility study on this topic (Bayerisches Archiv für Sprachsignale c/o Institut für Phonetik, Universität München; 4-7 December 2017).

SELECTED REFERENCES

- [Data Protection Directive] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. Official Journal L 281, 23/11/1995 p. 0031 – 0050. Available at <http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:31995L0046&qid=1522340616101&from=EN> (29.3.2018);
- [GDPR] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). OJ L 119, 4.5.2016, p. 1-88. Available at <http://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1515793631105&uri=CELEX:32016R0679> (29.3.2018);
- Joaquín González-Rodríguez, Doroteo Torre Toledano, Javier Ortega-García (2008). Voice Biometrics. In Handbook of Biometrics edited by Anil K. Jain, Patrick Flynn, Arun A. Ross. Springer;
- Italian Personal Data Protection Code. Legislative Decree 30.06.2003 No. 196. English version available at: <http://194.242.234.211/documents/10160/2012405/Personal+Data+Protection+Code+-+Legislat.+Decree+no.196+of+30+June+2003.pdf> (11.4.2018);
- Anil K. Jain, Arun Ross, Salil Prabhakar (2004). An Introduction to Biometric Recognition. - IEEE Transactions on Circuits and Systems for Video Technology 14(1). Available at https://www.cse.msu.edu/~rossarun/BiometricsTextBook/Papers/Introduction/JainRossPrabhakar_BiometricIntro_CSVT04.pdf (31.3.2018);

- Aleksei Kelli, Kadri Vider, Krister Lindén (2015). The Regulatory and Contractual Framework as an Integral Part of the CLARIN Infrastructure. 123: Selected Papers from the CLARIN Annual Conference 2015, October 14–16, 2015, Wrocław, Poland. Ed. Koenraad De Smedt. Linköping University Electronic Press, Linköpings universitet, 13–24. Available at <https://www.ep.liu.se/ecp/article.asp?issue=123&article=002> (20.8.2018);
- Manon Oostveen. Identifiability and the applicability of data protection to big data. *International Data Privacy Law*, 2016, Vol. 6, No. 4, 299- 309;
- G. Spindler, P. Schmechel (2016). Personal Data and Encryption in the European General Data Protection Regulation. - *JIPITEC* 7 (2), 163-177. Available at https://www.jipitec.eu/issues/jipitec-7-2-2016/4440/spindler_schmechel_gdpr_encryption_jipitec_7_2_2016_163.pdf (14.4.2018);
- WP29. Guidelines on Consent under Regulation 2016/679. Adopted on 28 November 2017 [adopted, but still to be finalized]. Available at http://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=615239 (2.4.2018);
- WP29. Opinion 05/2014 on Anonymisation Techniques. Adopted on 10 April 2014. Available at http://collections.internetmemory.org/haeu/20171122154227/http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf (20.8.2018).
- WP29. Opinion 4/2007 on the concept of personal data. Adopted on 20th June. Available at http://collections.internetmemory.org/haeu/20171122154227/http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf (20.8.2018).