



**Movie Making Finland:
Finnish fiction films
as audiovisual big data, 1907–2017
(MoMaF)**

Valentin Vaala: Vaimoke 1936 © KAVI / Suomi-Filmi Oy

Consortium PI

- Hannu Salmi (hansalmi@utu.fi) University of Turku/School of History, Cultural Research and Art Studies



Subproject PIs

- Filip Ginter (figint@utu.fi) University of Turku/Department of Future Technologies
- Mikko Kurimo (mikko.kurimo@aalto.fi) Aalto University/Department of Signal Processing and Acoustics
- Jorma Laaksonen (jorma.laaksonen@aalto.fi) Aalto University/Department of Computer Science

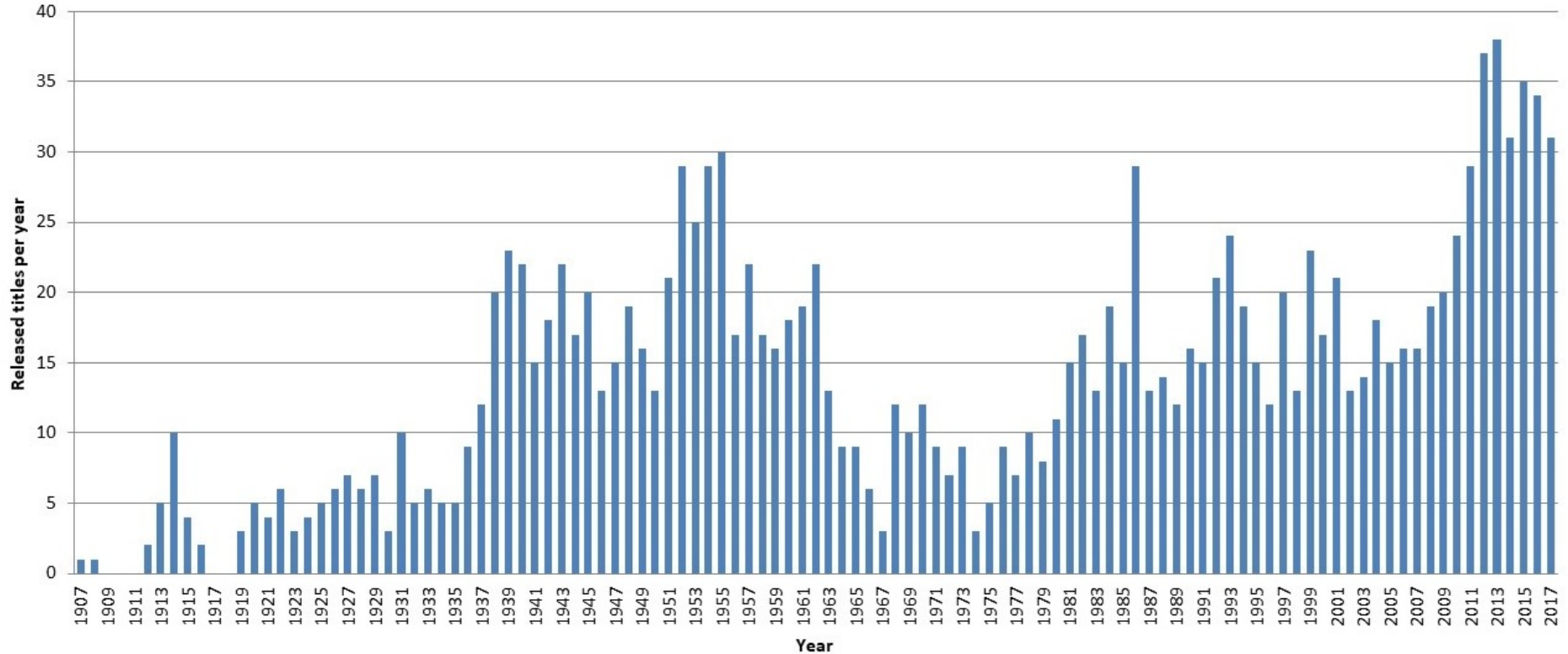


Partner

- National Audiovisual Institute, KAVI



Finnish fiction films as big data, 1907–2017





Research questions

- How has cinema imagined and interpreted Finnish modernization and its discontents?
- How can speech recognition, image analysis and natural language processing be combined in developing a methodological toolbox for the study of historical change in audiovisual cultural heritage, in this case Finnish fiction films, and how can these findings be incorporated into a film historical and cultural historical content analysis?

Cinema and modernization

1. **The changing environment of cinematic narratives.** This theme will be methodologically solved by identifying computationally rural and urban settings, interiors and exteriors, not only between films but within them, and by extracting features of natural surroundings, built and material environment.
2. **The role of technology.** Cinematic screen has been a platform for showing new, and old, technologies for the audience. The project tracks down how such everyday technological devices as telephones and cell phones, cars and airplanes, have been emphasised on screen during the project timespan.
3. **The forms of interaction and sociability.** From the perspective of modernization, it is crucial to explore social structures. In doing so, the project analyzes the language of the films on a large scale, particularly the ways of speech in relation to class, gender and region, and the changes of emotional expressions in dialogue over time.

Sources

Core sources:

- Finnish films throughout the period. In the application phase, we had secured research permissions to such companies as Suomi-Filmi, Suomen Filmitöollisuus, Fennada Filmi, Teuvo Tulio, and Sputnik Ltd.
- By March 2020, we have got permissions also from MTV (Spede Productions), Solar Films, Yellow Film, and Helsinki-filmi. This amounts to 615 full-length films in total.

Other sources:

- Elonet metadata, movie scripts, theatrical plays and novels, digitized scripts, other training materials for speech recognition



Figure 1: Project composition.

Towards Automatic Speech Recognition

The project has started by locating training material for developing automatic speech recognition for historical audiovisual material.

Important finding: KAVI holds around 40 dialogue lists of Finnish fiction films from the years 1934–1962.



Dialogue list of the film
Blue Week (Sininen viikko, 1954)

SININEN VIIKKO

Speaker: On heinäkuinen lauantai-iltapäivä. Liikkeet ja virastot ovat sulkeneet ovensa. Tehtaissa ovat koneet pysähtyneet. Ihmiset ovat kulkeneet koteihinsa, vaihtaneet vaatteita ja peseytyneet. Viikko on ollut helteisen kuuma. Tehtaissa ja konttorisaleissa on ollut tukahduttavaa, vaatteet ovat liimaantuneet ihoon kiinni. Mutta nyt, tällä hetkellä, on ainoa ajatus: maanantaiaamuun on aikaa 40 tuntia. Tietenkin olisi miellyttävintä ajaa omalla Packardilla omalle huvilalleen ja sekoittaa itselleen kevyt viskigroggi terassin aurinkokatoksen alla ja mennä sitten tarkastamaan, onko saunan kuparikatto vielä kiiltävä. Mutta sitä eivät kaikki voi tehdä, asia nyt on kertakaikkiaan sillä tavalla. Mutta sen sijaan voimme päästä naapurin moottoriveneellä tai yhteisesti vuokratulla pirsimoottorilla aivan kaupungin läheisyydessä sijaitsevaan Telttaniemeen. Siitä on Työväenyhdistyksen toimesta ja kaupungin myötävaikutuksella saatu aikaan melkoisen siedettävä leirialue, jonne halukkaat voivat pystyttää telttansa ja viettää viikonlopun yksin, kaksin tai perheensä kanssa. Lauantaina sinne yleensä mennään ja maanantaiaamuna aikaisin tullaan pois ja tavaksi on tullut jättää teltat sinne odottamaan seuraavaa lauantaita. Tanssilavakin sinen saatiin tänä kesänä ja Työväenyhdistyksen torviseitsikko huolehtii tanssi- ja viihdemusiikista.

Usko: Anteeksi vain! Ei kai vain sattunut?

Siiri mie: Ei ollenkaan. Kiitoksia vain paljon.

Usko: Saanko luvan?

Siiri: Kiitos ei.

Speaker: Sunnuntiaamu valkeni kauniina ja aurinkoisena. Heti aamulla tuntui, että päivästä muodostuu pihisevän kuuma. On miellyttävää herätä aikaisin sellaisena aamuna ja todeta, että edessä on vapaapäivä,

Methodological example: visual concepts in *The Unknown Soldier* (1955)

Food – true positives 1

0.999489 f10:113528:136 113528:18881

lscm-346-c.in12.m152.pool5o.d.a@lscm3141



Food – true positives 3

0.958982 f10:113528:1709 113528:165232

lscm-346-c.in12.m152.pool5o.d.a@lscm3141



Food – true positives 2

0.991315 f10:113528:1211 113528:56704

lscm-346-c.in12.m152.pool5o.d.a@lscm3141



Food – true positives 4

0.956435 f10:113528:1830 113528:210788

lscm-346-c.in12.m152.pool5o.d.a@lscm3141



Examples by
Jorma Laaksonen

Horse – true positives 1

0.999924 ff10:113528:t439 113528:99146

lscom-346-c.in12_rn152_pool5o.d.a#lscom322



◁ □ ▷ ⏪ ⏩ 🔍 🔄

Horse – true positives 7

0.915316 ff10:113528:t895 113528:222638

lscom-346-c.in12_rn152_pool5o.d.a#lscom322



◁ □ ▷ ⏪ ⏩ 🔍 🔄

Horse – true positives 2

0.999855 ff10:113528:t880 113528:220577

lscom-346-c.in12_rn152_pool5o.d.a#lscom322



◁ □ ▷ ⏪ ⏩ 🔍 🔄

Horse – true positives 8

0.631946 ff10:113528:t65 113528:31982

lscom-346-c.in12_rn152_pool5o.d.a#lscom322



◁ □ ▷ ⏪ ⏩ 🔍 🔄

trench – true positives 1

0.999842 ff10:113528:t748 113528:184327

sun-397-c.plpl.g.cls3pool.d.aA3#trench



◁ □ ▷ ⌂ 🔍

trench – true positives 3

0.999137 ff10:113528:t962 113528:237063

sun-397-c.plpl.g.cls3pool.d.aA3#trench



◁ □ ▷ ⌂ 🔍

trench – true positives 2

0.999401 ff10:113528:t747 113528:184304

sun-397-c.plpl.g.cls3pool.d.aA3#trench



◁ □ ▷ ⌂ 🔍

trench – true positives 4

0.999022 ff10:113528:t403 113528:93662

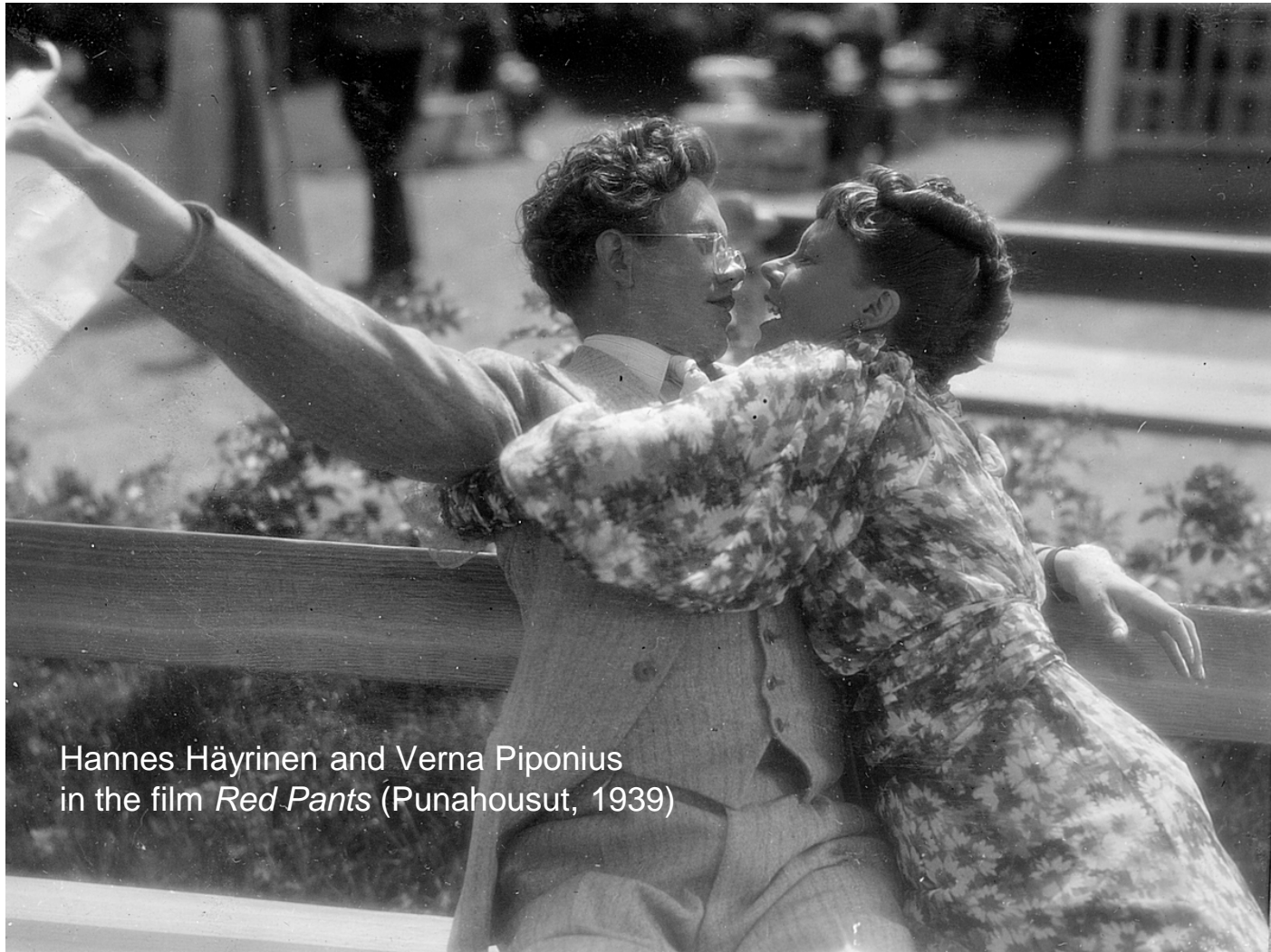
sun-397-c.plpl.g.cls3pool.d.aA3#trench



◁ □ ▷ ⌂ 🔍


Expected outcomes

- A digital index of all analyzed films
- Transcriptions of Finnish fiction films that can be later used as raw material for subtitling and thus for the internationalization of Finnish cinema
- Enrichment of metadata: named entities, topics, environments, objects
- Publications on research questions and methods
- Toolbox for analyzing audiovisual cultural heritage



Hannes Häyrinen and Verna Pipónius
in the film *Red Pants* (Punahousut, 1939)

Thank you!

 @MoMaF_digihum
www: momaf.github.io