

Version 1.0

List of the authors

Short title	SWHAP
Full title	Software Heritage Acquisition Process
	Roberto Di Cosmo, Software Heritage, Inria and University of Paris <roberto@dicosmo.org> Carlo Montangero, Dept. of Computer Science, University of Pisa <carlo@montangero.eu> Guido Scatena, Dept. of Computer Science, University of Pisa <guido.scatena@unipi.it>
Date	November 16, 2021
Contact	Roberto Di Cosmo <roberto@dicosmo.org>

Abstract

The presentation of acquired source code of landmark legacy software is particularly important. This document presents the first version of the integration of WikiMedia and ScinceStories.io presentation into SWHAP, the Software Heritage Acquisition Process: a protocol for the collection and preservation of software of historical and scientific relevance. SWHAP results from a fruitful collaboration of the University of Pisa with Software Heritage in this area of research, under the auspices of UNESCO, and has been validated on a selection of software source code produced in the Pisa area over the past 50 years.

Acknowledgments TBD

License This work is distributed under the terms of the Creative Commons license CC-BY 4.0

Contents

1	Introduction	2
2	TBD : Presentation designer might be abstract, and instantiated as SWH-stories designer in this version.	2
2.1	Resources in the process	2
2.2	SWHAP Repository structure	3
2.3	Story Inventory	3
2.4	People Moments	3
2.5	Roles in the process	3
3	A walkthrough on a running example	3

1) Introduction

Create a SWH-story

The suggested process to document the recovered source code with a story in the SWH-stories website <https://swh.stories.k2.services/stories/> has two phases: - Collect, where the Presentation designer collects the images, videos, documents to be published. For each item, he also gathers in a suitably structured *inventory* in the Workbench the information needed to insure that once uploaded in the appropriate Wikimedia data base the item is satisfactorily self documenting. The Presentation designer should also take care that the item can be granted a public domain license, as required by Wikimedia policies. - Publish, where the Web engineer, in this case the Wikimedia expert, uploads the items in such a way that they are best presented exploiting the STORIES SERVICE at [\[http://stage.stories.k2.services/publisher/\]](http://stage.stories.k2.services/publisher/).

The story will be available at [{.underline}](https://swh.stories.k2.services/stories/). Since the story is dynamically constructed at each access, the above process can be freely iterated, to add new elements or to correct the presented information.

TBD : Choose between inserting the sentence *The process is described in detail in ???* or elaborate here the description above. The presentation process, abstract view `{#sec:processabs}`

2) TBD : Presentation designer might be abstract, and instantiated as SWH-sto

Present in SWH-stories This way of presenting the recovered source code is inspired by the <https://sciencestories.io> website.

TDB : decide between an *embedded* or an *external* approach. In the embedded case we insert here a description as abstract as possible of the process to create a SWH-story, and see how to put other information in the rest of the document. In the external case, we refer to a new document, based on Morane and Kat's report, completed with a section related to the support in the adjourned template of SWHAPPE.

2.1) Resources in the process

- SWHAP Repository
- Wikidata
- Wikimedia Commons
- Software Heritage
- Publisher Workspace (?)

2.2) SWHAP Repository structure

2.3) Story Inventory

Wikidata Entities in this section we annotate the entities to be created ### Presentation Moments

2.4) People Moments

as a software story is also a story about people ... ##### People Media Gallery contains all images related to people

People Library contains all documents related to people

Software Moments

Media Gallery Moment

Libray Moment

Media Gallery Moment

2.5) Roles in the process

- Curator

Presentation designer and Web engineer

Though most of the presentations of the archived software will be on line, the abilities to design the contents of a presentation should be considered separately from the technical ones. For instance, in the case of the SWH-stories, the presentation designer should be competent in the topic addressed by the code, to be able to search and select the items to be inserted in the story. On the other side, the web engineer should be knowledgeable of Wikimedia and the other tools involved in creating the stories.

- Visitor ### Collector

3) A walkthrough on a running example