Artificial Intelligence:

The Role of Al in Decoding the Evidence

UNLOCKING THE RECORDS: WHERE TO FIND THEM, HOW TO READ THEM

Researching archives worldwide for evidence of Nazi-era cultural plunder is enormously challenging. Knowing where to start and how to determine which archive contains what information is a research project of its own.

Many of the archival repositories lack scanned or digitized copies of their records, making them unavailable online. Viewing the documents to trace the journey of specific looted objects can require costly in-person visits to archives in multiple locations.

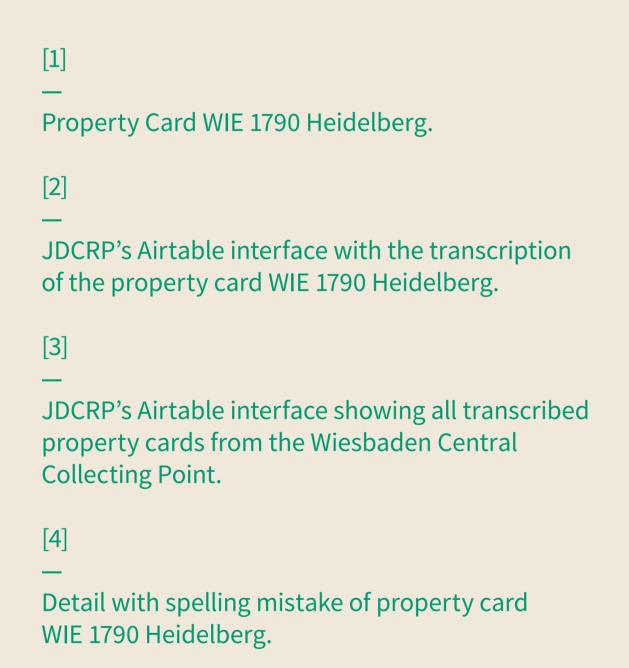
Some archives have posted datasets online, but only with metadata that summarizes the contents of the material. Even the digitized documentation available online often is not machine-readable, due to less advanced computer technology, making document-level content searches impossible.

Once the documents are obtained, it can be challenging to read them. Documents related to a single object are often in different languages, sometimes with difficult-to-decipher handwritten passages. Many records contain faded ink, scribbled notes in margins, stamps, and complex structures that make them difficult to understand.

To address these challenges, JDCRP is bringing together documents from various international archives on one online platform that is being launched with four different document sets:

- Bundesarchiv Koblenz: The property cards from the Wiesbaden Central Collecting Point.
- National Archives and Records Administration: Photographs that were taken at the Wiesbaden Central Collecting Point.
- Bundesarchiv Koblenz: The property cards from the Marburg Central Collecting Point.
- Jeu-de-Paume Database: The records of the looting agency Einsatzstab Reichsleiter Rosenberg.

The archival documents are digitized in high resolution to enable the users to zoom in on details.





The property cards from the Wiesbaden Central Collecting Point, National Archives and Records Administration Photographs of Activities of the Wiesbaden Central Collecting Point, National Archives and Records Administration The property cards from the Marburg Central Collecting Point, Bundesarchiv The property cards from the Marburg Central Collecting Point, Bundesarchiv

THE JDCRP AI-ENHANCED OCR SOLUTION

The next challenge is converting these document scans into searchable text. A specially developed conversion process is applied: Optical Character Recognition (OCR) software incorporates generative AI technology similar to Chat GPT that "reads" the documents and transcribes the content. The AI technology is a language model that enhances the ability of the OCR software to "understand" what it reads. The AI-driven software is programmed to transcribe documents and generate new data, such as texts and images. With the guidance of external AI experts, JDCRP's IT and Data Officer Jonathan Blok coordinated and refined an OCR process that allows a computer to understand the complex structure of documents and produce logical correlations, such as reading the contents of a data field within a chart from top to bottom, instead of from left to right across the document.

With its AI-Enhanced OCR solution, JDCRP has developed computer capabilities to read highly diverse historical documents, transcribed with an unusually low rate of error. The documents are now fully searchable. [1] [2] [3]

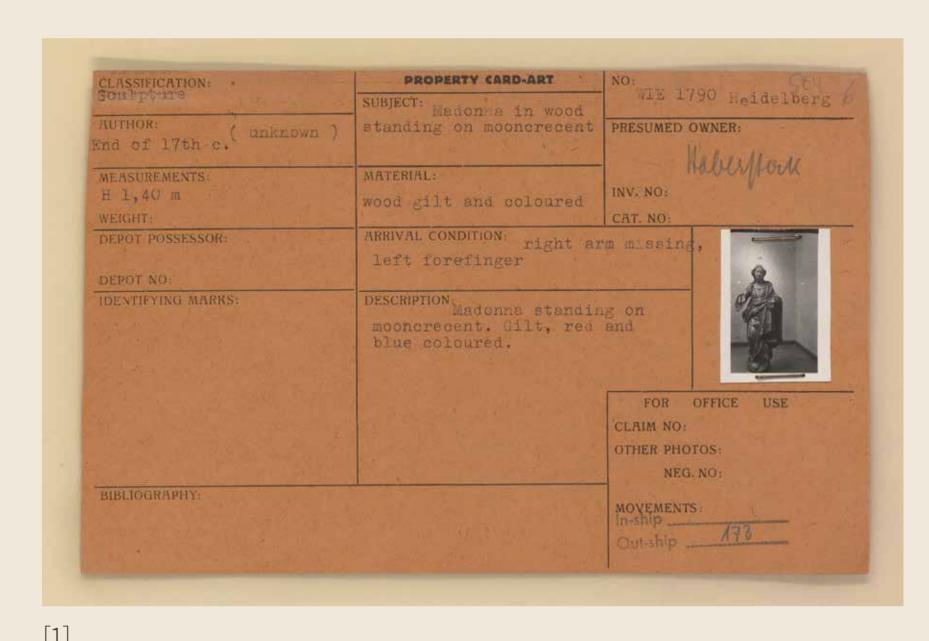
HUMAN REVIEW ENHANCING DIGITAL RECORDS

At each step of this complex process, JDCRP maintains the principle of sticking to the source. Any information that is added, inferred, or curated is always marked as such and must remain clearly distinguishable from OCR data that represents the original source document.

Following the processing of the archival record with JDCRP's unique OCR procedure, a trained team of students carefully reviews the AI-generated results, searching for computer errors that might be missed. This includes reviews for unwanted computer corrections of mistakes in the original sources, in order to preserve the errors in the source material. Their corrections are fed back into the AI system, allowing it to continuously improve its language recognition abilities and reduce the need for manual review over time.

As an example of an error in the source material, the word "mooncrecent" was wrongly corrected by the computer and reversed through manual review to the original error in the source. [4]

This AI-powered OCR process is the foundation of the JDCRP archival platform. By creating accurate, structured, and searchable records, JDCRP has taken a significant step forward in decoding and making accessible the voluminous evidence of Nazi-era theft of Jewish-owned cultural property.



JDCRP OCR « Database »	OCR Reviewer / All Records	- A . W	0			
GCR Reviewer	Alle Karten Rew Cards Revised Cards Approved Cards Feedback Needed	WIE 1790 Heidelberg				
Ranto Farel 300de	The second second	wccp-number-suffix	504			
All Records		dassification	Soulphure			
in Interface		author	End of 17th a curinowny			
		subject	Medonra in wood standing on moontrecent			
	WIE 1790 Heidelberg	mexiconents	H 1,40 m			
	Statys	material	wood git and coloured			
	Portugued.	depenytion	Madenna standing on moonsrecest. Git, red and blue coloured.			
	escondustro suffix 564	presumed owner	Notestada			
		investory number	\$2.5			
	thurtgitipe	catalog-number	14)			
	Soulytire	dam-number	GF C			
		photo-attached	¥			
	euther (red of 17th c. (uninden)	photo :	pen -			
		identifying merks	553			
	product-own	other-photos	100			
	Haberstock	regative-number				

□ Ansichten □ Grid view											
Q Ansicht suchen	@	□ A file	name-front V	≅ wccp-number ∨	≤ catalog-number ∀	AF claim-number V	≦E photo ∨	18 identifying-marks V	△≅ other-photos	∨ № negative-number ∨	15 movements
40200		2588 B_323_	589_1031	WIE/ 1784 [crossed ou				Rests of signature lower lef			Inshipment 5 In-ship
☐ Grid view	90	2589 B_323_	589_1033	WIE 1785/1 Heidelberg						WIE 239	In-ship 2 Out-ship 29
E Clean view		2590 B_323_	589_1035	WIE 1785/2 Heidelberg							In-ship 6 Out-ship 29
		2591 B_323_	589_1037	WIE 1786 Heidelberg							In-ship Out-ship 173
		2592 B_323_	589_1039	WIE 1787 Heidelberg							In-ship 6 Out-ship 173
		2593 B_323_	589_1041	WIE 1788 Heidelberg	Coat of arms from a			Labelled:Innsbruck			In-ship Out-ship 173
		2594 8_323_	589_1043	WIE 1789 Heidelberg	наверном						In-ship Out-ship 173
		2595 8_323_	589_1045	WIE 1790 Heidelberg							In-ship Out-ship 173
		2596 B_323_	589_1047	WIE 1791/92 Heidelberg				Label on back.			In-ship 6 Out-ship 17
		2597 B_323_	589_1049	WIE 1792 Heidelberg							In-shipm. 6 [In-shipme
Wird ersteilt	~	2190 B_323_	589_1051	WIE/1793 Heidelberg 6				Label: Verfügungsberechtig	Photo Marburg	Wie 158	In-shipm. 6, Out-shipr
	+	2509 B_323_	589_1053	WIE/1794/Heidelberg				Label: Dr. Sieger, Berlin-Wil			In-ship 6 Out-ship 298
Malender	+	2600 B_323_	589_1055	WIE/1795 Heidelberg				Label: 32, Gerard Dou Mutt	Photo 529	Wie 155	In-shipm. 6 IS 376
28 Galerie	+	2601 B_323_	589_1057	WIE/1796 Heidelberg				Label: Eigentümer Dr. Siege	Photo Marburg	Wie 110	In-ship 6 Out-ship 124
 ⊕ Kanban ⇔ Zeitleiste ⊯ Liste	+	2602 B_323_	589_1059	WIE/1797 Heidelberg				Signed above left shoulder	Photo Marburg	Wie 112	In-ship 6 Out-ship 298
	+	2608 B_323_	589_1061	WIE 1798/Heidelberg				Label: Verfügungsberech- t	Photo Marburg	Wie 147	In-ship 6 Out-ship 296
	+	2604 B_323_	589_1063	WIE 1799 Heidelberg							In-ship 6 Out-ship 296
€ Gantt	+	2505 B_323_	589_1065	WIE 1800 Heidelberg				none			In-ship 6 Out-ship 298
		2606 B_323_	589_1067	WIE/1801 Heidelberg	CAT. NO			Signed lower right hand co	Photo Marburg	Wie 158 [Handwritten addi	In-ship 6 Out-ship 298
Formular	+	+ 6.289 Karten									

SUBJECT: Madonna in wood standing on mooncrecent

