

# An Analysis of the ARTigo Gaming Ecosystem With a Purpose

François Bry, Clemens Schefels

Institute for Informatics, University of Munich  
Oettingenstrasse 67, 80538 Munich, Germany  
bry@lmu.de, schefels@ifi.lmu.de

**Abstract.** ARTigo is both, a gaming platform and an artwork search engine: While playing, ARTigo’s users leaves annotations describing artworks; these annotations are automatically processed by a higher-order latent semantic analysis for building, and continuously improving, an artwork search engine. ARTigo’s search engine has been developed for, and with, art historians and covers more that 65,000 artworks of different styles and epochs. ARTigo is in operation since 2008.

ARTigo’s gaming platform is an ecosystem consisting of different types of games aiming at collecting rich annotations of different type: Description games aim at collecting mere descriptions, diversification games aim at completing artwork descriptions with more specific characterizations, integration games cluster annotations so as to generate more precise descriptions. This article’s analysis points to both, the effectiveness of ARTigo as a gaming ecosystem and the effectiveness of each of the games at collecting specific annotations.

**Keywords:** Serious Games, Games With A Purpose, Artwork Search, Art History, Human Computation

## 1 Introduction

ARTigo (<http://artigo.org>) is an outcome of the five-year research project play4science (<http://play4science.org>). Since the project’s beginning in 2007, ARTigo collects semantic data on artworks from the footprints left by players of games featuring artworks. From these footprints and relying on an original higher-order latent semantic analysis conceived for the purpose [5, 11, 10], ARTigo automatically builds an artwork search engine.

In Art History, there is a considerable interest for search engines that would make it possible to retrieve artworks from their iconographic content, that is, by referring to what they explicitly depict (like an object, a scene, a person, or a group of people) as well as from their implicit content, that is, the emotions (like love or sorrow) or abstractions (like gallantry, freedom, or power). Some art historians expect such search engines to open up new perspectives in art history by unveiling similarities, or discrepancies, or any other so far unnoticed features of artworks, art schools, or epochs.

Since art historians rarely have much funds at their disposal, an inexpensive approach to annotating artworks was necessary. Indeed, annotating a large artwork database can become costly if done for money: low paid workers like students working a few hours a week would for example in Germany cost about €1,500 for a full month work load and would produce on average about three annotations per minute, that is 28,800 annotations per month, yielding an average price for annotation of about €0.05. Annotating 10,000 artworks with an average of 20 annotations per artwork each would therefore cost about €10,000. Requiring further that annotations being proposed by two independent workers and considering, as we have experienced, that only 20% are proposed by two independent workers, increases the cost of annotating 10,000 artworks to €50,000. We were considering an artwork image database of 30,000 to 100,000 artworks from the Prometheus Image Archive (<http://www.prometheus-bildarchiv.de/>) and from museums we were in contact, or collaborating, with, raising the cost for paid annotators to €1.5 to €5 millions. This was by several orders of magnitude out of our reach. We therefore decided to rely on unpaid lay persons, human computation, and gaming as incentivization.

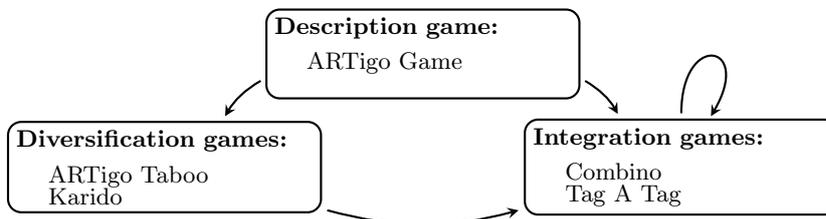
This article briefly describes ARTigo’s gaming ecosystem and reports on an analysis pointing to its effectiveness at collecting annotations well describing the explicit as well as implicit contents of artworks.

This article is structured as follows. Section 1 is this introduction. Section 2 describes the ARTigo gaming ecosystem consisting of description, diversification, and integration games. Section 3 is devoted to an analysis of the artworks annotations collected by ARTigo. Section 4 is a conclusion.

## 2 The Gaming Ecosystem

Figure 1 [10] shows how the annotations collected on the ARTigo platform flow from game to game providing the “seed data” necessary for some games to be playable and for collecting rich annotations.

Rich annotations is used throughout this article for denoting annotations of different kinds collected by different games: mere descriptions collected by so-called “description games”, more specific characterizations collected by so-called “diversification games”, and annotation clusters collected by so-called “integration games”. The different types of games are introduced in the following.



**Fig. 1.** Seed data flow in the ARTigo ecosystem.



**Fig. 2.** The ARTigo Game, Franz Marc: The World Cow, 1913, MoMA (NY).

Whether the games of each type collect complementary annotations they have been conceived for is the core issue of the analysis this article reports about.

*Description Games* are games the players of which describe an artwork by proposing annotations related to anything referring to the artwork: objects or characters it depicts, its colors, the materials it is made of, etc. ARTigo's has a single description game, a variant of the ESP Game [9], called *ARTigo Game* (see Figure 2), with the following gameplay: When two randomly selected players enter the same annotation for the same artwork, they both score and the annotation is validated. Validation is necessary to ensure annotation correctness (among others exclude malicious annotations). This form of validation has been found sufficient [9]. The annotations collected by description games provide “surface semantics”. They are needed both as basic descriptions and as seed data for games collecting rich annotations. Like other researchers [3, 2]), we noticed a limitation of the ESP game: its tendency to collect very general annotations. To overcome this limitation, we deployed complementary games as follow.

*Diversification Games* collect annotations that are in general more precise or of a deeper semantic than most annotations collected by description games. Diversification games use annotations collected by description games as seed data. The ARTigo ecosystem has two diversification games:

*ARTigo Taboo* is an ESP Game [9] and like the ARTigo Game except that some annotations already entered by former players can no longer be entered by any player: they are taboo. ARTigo Taboo forces players to enter novel annotations. Therefore, ARTigo Taboo yields more sophisticated descriptions than the ARTigo Game. Taboo-ing annotations can be seen as a form of “scripting” [1] discussed below since taboos are instructions for players. The annotations collected by the ARTigo Game on the one hand and by ARTigo Taboo on the other hand are complementary.

*Karido* [7, 6] is the second diversification game of the ARTigo ecosystem. Its gameplay is as follows: Nine similar artworks are randomly selected and

displayed to randomly paired players in two  $3 \times 3$  grids such that the artworks are differently ordered in each grid. Image similarity is determined from the annotations, especially from “surface semantics” annotations, so far collected for these artworks [6]. One player is the “describer”, the other the “guesser”. In the next round, they exchange roles. The describer selects one of the nine artworks on her grid and starts annotating it in such a way that the guesser can recognize and select it. The guesser can ask yes/no questions. Since the guesser’s grid and the player’s grid are differently ordered, locational annotations like “South-East” do not help recognizing the selected artwork. The sooner the guesser selects the right artwork, the higher the scores for both players. Karido’s gameplay incites players to enter more specific, or rich, annotations. Karido can be played in two modes called in the following “time” and “turn” in which the time, the number of turns, respectively, are bounded.

*Integration Games* cluster annotations yielding more precise descriptions than the unstructured sets of annotations collected by other games. Annotation clusters are often more difficult for players to suggest since they require a deeper understanding of an artwork, in some cases even some specific knowledge. Integration games are therefore sometimes more challenging than other games. Two integration games have been designed for the ARTigo ecosystem:

*Combino* makes its players bring formerly collected annotations into relation. A same artwork together with a same set annotations formerly collected for this artwork are displayed to randomly paired players. Both players score when they select the same pairs or sets of annotations from the displayed annotations. Combino is, like ARTigo Game and ARTigo Taboo, a variation of the ESP Game [9]. By “squaring” Combino, that is running it with phrases it formerly collected, long phrases like “old man sitting” can be stepwise collected from “old”, “man”, and “sitting”, that is, reconstruct descriptive phrases.

*Tag A Tag* is a “squared” [1] and “scripted” [1] ESP Game [9]. “Squaring” has been already introduced in the former paragraph: “Squaring” a game means running it with annotations it had formerly collected. “Scripting” means giving the players a short instruction on the annotations, or pairing, or whatever other input they are expected to enter. Tag A Tag is displaying to its player an artwork AW, an annotation A formerly collected for this artwork AW and asking “Describe A in artwork AW”. Thus, Tag A Tag collects annotations on an annotation A in the context of an artwork AW.

Most games of the ARTigo ecosystem are variations of the ESP game. Is this a good feature? The authors believe it is for the following reasons. First, it makes all games “casual games”, that is, games that can be played in sessions of any duration, especially in short sessions, and requires neither a learning of rules nor much training. Second, keeping the gameplays of all games similar contributes to the “casual” nature of the gaming platform. Finally, it reduces the effort needed for implementing and maintaining the games thus contributing to the cheapness of the platform. Recall that cheapness is a concern in building a semantic artwork search engine.

### 3 Analysis of the Effectiveness of the Ecosystem

This section reports on an analysis of the raw data collected by the gaming ecosystem consisting in both quantitative and qualitative analyses. The analyses' main finding is that, as intended, the ecosystem's various games collect different type of data, thus achieving its purpose.

The analyses are limited to the German Annotations collected by ARTigo for the following reasons: First, thanks to an excellent press coverage over at least 5 years in Austria and Germany, the two main German speaking countries, over 6.8 million annotations in German have been collected during 123,000 game sessions from on average 150 players per day. Second, much less annotations have been so far collected in English and French, the two other languages supported on ARTigo. Third, it is more sensible to analyze annotations all in a same language.

The survey used in the qualitative analyses has been performed in 2014. For consistency reasons, only annotations collected up till 2014 are considered in the analyses reported about in the following.

#### 3.1 Quantitative Analysis of the Collected Annotations

The games' performances at collecting annotations vary considerably, as Table 1 shows, ARTigo Game having collected almost 96% of the annotations and each of the other games less than 2%, one of them, Tag A Tag having collected only 0.07% of the annotations.

There are three obvious reasons for this: First, while ARTigo has been played since 2008, the other gamers have been introduced later, Karido in 2011, ARTigo Taboo, Combino and Tag A Tag in 2012. Second, ARTigo Game is listed first on the ARTigo main Web page. Third, and more importantly, ARTigo has a simpler gameplay than the other games and allows for a stronger focus at the artworks displayed during game sessions. Assuming that an interest in artworks is a core motivation for using the ARTigo gaming platform, it is reasonable to conclude that ARTigo Game is likely to attract players more than the other games.

We argue that the aforementioned performance differential of the various games does not harm the overall gaming ecosystem's performance. Indeed, the ecosystem's purpose is to collect annotation and games contributing with small numbers of annotations would be useful provided their contributions are not

**Table 1.** Annotations collected from 2008 to 2014.

	ARTigo Game (2008)	ARTigo Taboo (2012)	Karido Time Turn (2011)		Combino (2012)	Tag A Tag (2012)	all games (2008)
Per year	1,140,683	66,900	28,400	26,200	23,300	2,600	1,327,000
Totals	6,844,100	133,900	85,100	78,700	46,500	5,300	7,141,900
Contributions	95.83%	1.87%	1.19%	1.10%	0.65%	0.07%	100.00%

**Table 2.** Game complementarity in annotation collecting (Explanation: 92% of the annotations collected by Karido have not been collected by ARTigo Game)

Game	ARTigo Game	ARTigo Taboo	Karido	Game	Combino	Tag A Tag
ARTigo Game	–	90%	92%	Combino	–	97%
ARTigo Taboo	16%	–	65%	Tag A Tag	89%	–
Karido	68%	83%	–			

redundant, that is, are not also made by other games. Table 2 shows that this is the case.

Table 2 considers on the one hand-side the non-integration games ARTigo Game, ARTigo Taboo and Karido, the integration games Combino and Tag A Tag on the other hand-side. Indeed, non-integration games collect almost only one-word annotations while the integration games have been designed to collect many-words annotations.

A row of Table 2 gives the proportions of annotations collected by the row’s game that have not been also collected by the column’s game: 90% (92%, resp.) of annotations collected by ARTigo Taboo (Karido, resp.) have for example not been collected by ARTigo Game.

All but one figures of Table 2 are above 80% showing that all but one game of the ARTigo ecosystem has its *raison d’être*. The possibly redundant game is ARTigo Game: Only 16% of ARTigo Game’s annotations have not been collected by ARTigo Taboo. This is surprising considering that it is the most popular game. ARTigo Game’s high popularity is, more than its contribution with unique annotations, its *raison d’être* in the gaming ecosystem.

### 3.2 Quality Analysis of the Collected Annotations

In 2014, a survey [4] among humans has been performed aiming at determining whether the ARTigo gaming ecosystem does collecting rich annotations in the sense that the collected annotations are not only mere descriptions (like “woman”) but also include more specific characterizations like compound annotations (like “young woman”).

**Survey Participants.** The survey was run on-line, in German, and open to everyone over four months in 2014. It had been advertised in the ARTigo blog and in social media. No experience with the ARTigo gaming platform was required for participating. The participants were not paid. As an incentive to participate, twelve Amazon vouchers à €25 were drawn as prizes. 159 people fully answered the survey of whom 77 were randomly assigned to a first group, the remaining 82 to a second group. 60% of the participants were female, 2.5% younger than 20 years, 47% 20 to 29 years old, 20% 30 to 39 years old, 22% 40 to 59 years old, and 5.6% 60 years old or older. Among the survey participants, there were almost as

**Table 3.** Part 2: Comparison of sets of annotations.

Game	Artwork Set 1	Artwork Set 2	Artwork Set 3	Artwork Set 4
Karido	88.31%	45.12%	<b>86.59%</b>	<b>28.57%</b>
ARTigo Game	<b>93.90%</b>	<b>51.95%</b>	15.58%	3.66%

many ARTigo players as non-players: about 55% participants had already played on the ARTigo platform. 92% of the players had expressed an interest in art, 30% of the non-players had no or only some interest in art, 28% a high interest in art. 45% of the participants told that art was related to their profession.

**Survey Tasks, Questions, and Results.** The survey consists of five tasks or questions about artworks’ annotations have been conceived for two groups of participants. The survey parts were identical for all participants of a same group but slightly different between the two groups. Only the parts two, three, and five are relevant for this analysis and focused at the following questions:

1. Do diversification and description games collect different annotations?
2. Do integration and description games collect different annotations?
3. Do description and diversification games collect different annotations?

*Part 1* was devoted demographic data like age and gender (see Paragraph “Survey Participants”).

*Part 2* was devoted to Question 1. Four groups of four similar artworks were used in this part. Each artwork group was build as follows from a so-called “initial artwork” randomly selected from the ARTigo database. The five most frequent annotations of an initial artwork were used as search query for selecting three further artworks, so-called “test artworks”, from the ARTigo database. Because of their sharing of annotations, the three test artworks can be considered as similar to the initial artwork they have been determined from. Each group of four artworks, an initial artwork with its three associated test artworks, were presented to each group of participants together with a set of the five most frequent annotations collected for the initial artwork by the diversification game Karido, to the second group together with a set of five randomly selected annotations collected by the description game ARTigo Game. The participants were asked to select the artwork best corresponding to the five annotations. Table 3 gives the results. The game from which the annotation set had been collected is given in Column 1. The percentages in bold indicate the annotations that helped most of the participants to correctly identify the initial artwork.

*Part 3* was devoted to Question 2: Do integration and description games collect different annotations? Three artworks, each artwork together with six of the annotations collected for it on the ARTigo platform, were presented to the participants who were asked to rank these annotations after their relevance for the artworks. The six annotations associated to each artwork included both,

**Table 4.** Part 3: Comparison of one-word and many-words annotations. In bold are the percentage of annotations considered by the survey participants as among the most relevant for describing the artworks they have been assigned to.

AW Annot. Type	1st	2nd	3rd	4th	5th	6th
1 Many-words	<b>64.78%</b>	49.67%	<b>51.41%</b>	42.06%	47.86%	46.02%
1 One-word	35.22%	<b>50.33%</b>	48.59%	<b>57.94%</b>	<b>52.14%</b>	<b>53.98%</b>
2 Many-words	<b>50.94%</b>	47.74%	47.10%	<b>53.79%</b>	<b>56.67%</b>	41.23%
2 One-word	49.06%	<b>52.26%</b>	<b>52.90%</b>	46.21%	43.33%	<b>58.77%</b>
3 Many-words	<b>66.04%</b>	47.22%	<b>54.69%</b>	<b>57.14%</b>	38.61%	36.46%
3 One-word	33.96%	<b>52.78%</b>	45.31%	42.86%	<b>61.39%</b>	<b>63.54%</b>

one-word annotations collected by ARTigo Game and many-words annotations collected by Tag A Tag. The one-word annotations of one survey group were parts of the many-words annotations of the other group, for example, if the annotation set of the first group contained an annotation like “woman”, then the set of the second group contained a many-words annotation like “woman + dress”. Table 4 gives the rankings the participants gave to the annotations, “1st” indicating the annotation considered most adequate, etc. The annotations most often selected are highlighted in bold. There is no clear preference for either one-word annotations or many-words annotations.

*Part 5* was devoted to Question 3: Do description and diversification games collect different annotations? The participants were asked to assess the relevancy, the specificity, and the level of detail of annotations collected for six artworks by the game description game ARTigo Game and the diversification games ARTigo Taboo and Karido. Three of the six artworks were presented to each group of the participants with eight annotations: four of the ten most frequently collected by ARTigo Game and four of the ten most frequently collected by ARTigo Taboo. The participants had to qualify each annotation on five-level Likert scale as “important”, “visible at first sight”, “broad”, and “generic” with respect to their associated artworks. One group of participants had to evaluate annotations collected by the ARTigo Game and ARTigo Taboo. The other group had to evaluate annotations collected by the ARTigo Game and Karido. Table 5 gives the results. The table shows that, for five of the six artworks (abbreviated “AW”), the annotations collected by the description game ARTigo Game were considered more important, better visible at first sight, broader, and more generic than the annotations collected by the diversification games ARTigo Taboo and Karido. For one of the six artworks, interestingly an abstract painting, the annotations collected by the diversification game Karido were rated as more important, better visible at first sight, broader, and more generic.

To avoid so-called primacy-recency [8], or serial position, effects, the artworks respectively annotations were presented to each participant in random orders.

**Table 5.** Results of Part 5: Comparison of annotations collected by description and diversification games. (“AW” stands for artwork.)

Dimension	Game Type	AW 1	AW 2	AW 3	AW 4	AW 5	AW 6
Relevancy	Description	76.62%	71.43%	67.99%	71.64%	49.35%	75.92%
	Diversification	29.55%	31.17%	55.18%	32.32%	59.09%	17.99%
Sem. depth	Description	88.31%	77.60%	78.66%	73.48%	41.56%	85.67%
	Diversification	19.48%	38.63%	67.68%	29.57%	67.53%	24.09%
Detail	Description	59.41%	55.19%	44.21%	43.60%	30.20%	58.23%
	Diversification	21.10%	16.88%	44.82%	20.73%	56.17%	11.59%
Specificity	Description	69.80%	58.12%	57.71%	63.11%	39.61%	66.16%
	Diversification	28.57%	37.34%	55.79%	49.39%	53.90%	53.05%

**Interpretation.** Summing up, the survey points to the following: *The different games of the ARTigo ecosystem collect different types of annotations.*

According to the results of Part 5, the annotations collected by the ARTigo Game, the single description game of the ecosystem, were mostly judged as the most important, generic, detailed, and highly visible at first sight. In addition, the diversification games ARTigo taboo and Karido complemented the good results of the ARTigo game. For some artworks their annotations have been seen as more useful to the participants, especially in respect to their detail level and specificity. Thus, the game diversity can be considered beneficial to a good annotating of artworks.

The benefit of integration games on the ecosystem is shown by the results of Part 3. Annotations consisting of many-words are only slightly more valuable than one-word annotations. Consequently, it is useful to have both many-words annotation generating games as well as one-word annotation generating game in the ecosystem.

The results of Part 2, however, show that the annotations collected by diversification games were judged as more specific in contrast to the annotations collected by the description game. The “initial artwork” was often identified with the help of annotations collected from the diversification game Karido what demonstrates that this diversification game fully fulfills its purpose. Furthermore, the results show the usefulness of the game Karido, which has been designed to complement the ARTigo game, in half of the cases.

## 4 Conclusion

This article has first briefly described the ARTigo gaming ecosystem aiming at collecting descriptions of artworks, that are not only mere descriptions but also more specific characterizations and annotation clusters. To the best of the knowledge of the authors, ARTigo is the first ecosystem of data-collecting games,

or games with a purpose [9]. Ecosystem means here that some of the data-collecting games collect seed data necessary for other data-collecting games to be playable.

This article has reported on an analysis pointing to the effectiveness of the ARTigo gaming ecosystem at collecting rich annotations. The analysis of the collected data indicates that each of the games of ARTigo collect different artwork descriptions. A survey-based qualitative analysis has shown that the ecosystem as a whole performs better than its parts.

**Acknowledgments.** The research reported about in this article as been founded in part by the German Foundation for Research (DFG) within the project play4science.

The authors are thankful to Corina Schemainda for preparing and conducting the survey, to Gerhard Schön for the implementation of a first prototype, to Christoph Wieser for his essential contributions to the conception and implementation of ARTigo’s gaming platform and search engine, to Hubertus Kohle, art historian at the University of Munich, and to Klaus U. Schulz, Center for Information and Language Processing Munich, for useful suggestions.

## References

1. Bry, F., Wieser, C.: Squaring and scripting the ESP game. In: Proc. 4th Human Computation Workshop (HCOMP). AAAI Press, Palo Alto, CA, USA (2012)
2. Jain, S., Parkes, D.C.: A game-theoretic analysis of the ESP game. *ACM Trans. on Economics and Computation* 1(1), 3:1–3:35 (January 2013)
3. Robertson, S., Vojnovic, M., Weber, I.: Rethinking the ESP game. In: CHI '09 Extended Abstracts on Human Factors in Computing Systems. pp. 3937–3942. CHI EA '09, ACM, New York, NY, USA (2009)
4. Schemainda, C.: Qualitative Analysis of the ARTigo Gaming Ecosystem. Bachelor Thesis, Institute for Informatics, University of Munich (2014)
5. Shah, P., Wieser, C., Bry, F.: Parallel Higher-Order SVD for tag-recommendations. In: Proc. of the Inti. Conf. WWW/Internet 2012. Lisbon, Portugal (2012)
6. Steinmayr, B.: Designing Image Labeling Games For More Informative Tags. Diploma Thesis, Institute for Informatics, University of Munich (2010)
7. Steinmayr, B., Wieser, C., Kneißl, F., Bry, F.: Karido: A GWAP for telling artworks apart. In: Proc. of the 16th Int. Conf. on Computer Games (CGAMES). pp. 193–200. IEEE, Washington, DC, USA (2011)
8. Tourangeau, R., Couper, M.P., Conrad, F.: Spacing, position, and order: Interpretive heuristics for visual features of survey questions. *POQ* 68(3), 368–393 (2004)
9. Von Ahn, L., Dabbish, L.: Labeling images with a computer game. In: Proc. SIGCHI Conference on Human Factors in Computing Systems (CHI). pp. 319–326. ACM, New York, NY, USA (2004)
10. Wieser, C.: Building a Semantic Search Engine with Games and Crowdsourcing. Doctoral Thesis, Institute for Informatics, University of Munich (2014)
11. Wieser, C., Bry, F., Bérard, A., Lagrange, R.: Artigo: Building an artwork search engine with games and higher-order latent semantic analysis. In: In Proc. of Disco 2013, Workshop on Human Computation and Machine Learning in Games (HCOMP). AAAI Press, Palo Alto, CA, USA (2013)