Abstract
The ESP Game tends to generate “low effort” or “surface semantics” tags. This paper presents two variations of the ESP Games called “squaring” and “scripting” that trim the ESP Game to collect “deep semantics” tags. The approaches do not require players to get used to, and for the GWAP operators to deploy, new games. First experiments point to the efficiency of squaring and scripting the ESP Game at collecting “deep semantic” tags.

Motivation
The ESP Game has been proposed in (von Ahn and Dabbish 2004) for collecting image descriptions. The ESP Game is good at its purpose but it tends to generate tags that are not image specific (Weber, Robertson, and Vojnović 2009; Law et al. 2011; Steinmayr et al. 2011). A theoretical study points to the tendency of players of the ESP Game to produce “low effort” tags (S. Jain and D. Parkes 2012).

New games have been proposed in (Weber, Robertson, and Vojnović 2009; Law et al. 2011; Steinmayr et al. 2011) so as to overcome this trait of the ESP Game.

This short paper suggests rather simple and natural variations of the ESP Game requiring neither new game logics nor new implementations that, in spite of their simplicity are capable of collecting “deep semantic” tag.

Squaring the ESP Game
A squared ESP Game is an ESP Game that asks players to tag a pair consisting of an image and of a tag that has been formerly collected for that image. We call this variation of the ESP Game a “squared ESP Game” in analogy to arithmetics because the game is applied to one of its former outcomes.

Squaring the ESP Game has appealing properties. It is possible without any selection of the tags selected for the pairs (image, tag) to be presented to players. It does not require for the gaming platform operators to develop a new game and for players to learn how to perform well at a new game. And it seems to be quite efficient at generating deep semantics tags – see Section below.

Scripting the ESP Game
Scripting the ESP Game consists in giving the players a context for their annotating of the images presented. The ESP Game has been proposed in (von Ahn and Dabbish 2004) for collecting image descriptions. The ESP Game is good at its purpose but it tends to generate tags that are not image specific (Weber, Robertson, and Vojnović 2009; Law et al. 2011; Steinmayr et al. 2011). A theoretical study points to the tendency of players of the ESP Game to produce “low effort” tags (S. Jain and D. Parkes 2012).

New games have been proposed in (Weber, Robertson, and Vojnović 2009; Law et al. 2011; Steinmayr et al. 2011) so as to overcome this trait of the ESP Game.

This short paper suggests rather simple and natural variations of the ESP Game requiring neither new game logics nor new implementations that, in spite of their simplicity are capable of collecting “deep semantic” tag.

Some pairs (image, tag) turn out to be rather tricky to play well with. This, however, contributes to the attractiveness of the gaming platform. Furthermore, we observed that pairs (image, tag) difficult to tag turn out to collect very interesting tags.

The ESP Game can be squared in more sophisticated manners than the one mentioned above.

First, automatically generated questions can be posed. A pair consisting of a tag W and an image can be presented with one of the following questions, depending whether W is a noun, an adjective or a verb: How is W in this image? What is W in this image? What is W-ing in this image?

Second, squaring can be realized using a set of tags instead of a single tag. The approach is especially promising if the set of tags –automatically– selected are semantically related.

Third, squaring can be realized with one tag T and several images so as collect annotations reflecting common traits of T in the images presented.

Fourth, the principle of squaring the ESP Game can be repeated. A squared ESP Game can itself be squared once again in whatever manner deemed convenient.
An essential requirement for a convenient approach to scripting is that it does require human work, or, if it does, that the human work needed be provided for by (other) games with a purpose (von Ahn 2006).

Also, combinations of the aforementioned forms of scripting can be thought of.

Like squaring the ESP Game, scripting it is appealing because it does not require for the gaming platform operators to develop new games and for players to learn how to perform well at new games.

**Squaring and Scripting Combined**

Squaring and scripting can both be combined, provided that the resulting combination is clearly understandable and not overloaded with explicit or implicit instructions.

**Experimental Evaluation**

A squared ESP Game called “TagATag” and a scripted ESP Game called “Sentiment” are tested since mid January 2012 on an extended version of the gaming platform ARTigo (http://artigo.org).

The purpose of ARTigo is a semantic search engine for art works (the search engine which so far performs well in German only is available at http://artigo.org). ARTigo has 14.169 registered users and additional 26.428 unregistered users of which on every day at least 73 play on the platform during visits lasting in average 5.5 minutes. Within 4 years, ARTigo collected 4.03 millions of tags, 0.855 million (21%) of which have been validated.

In a first experiment ten players have been selected. Between 17 January 2012 and 12 April 2012 they tested “TagATag” and “Sentiments”. During this period of time, the ten players spend together only 106 minutes testing “TagATag” and “Sentiment” and generated 195 pairs (image, tag), 85 (9%) of which were new.

The players rapidly got used to the non-standard versions of the ESP Game. In average less tags are collected with “TagATag” and “Sentiment” than with the plain ESP Game suggesting that the new versions of the ESP Game are more difficult to play well with. The addition of more demanding games to the the gaming platform has been positively received by the selected players.

An evaluation with non-selected players is to follow in the forthcoming weeks.

**Conclusion**

The great ability of the ESP Game to collect simple, concrete and rather immediate descriptions is for two reasons not to be wrongly interpreted as a weakness.

First, such description are needed for many applications.

Second, with squaring and scripting, the ESP Game can easily be trimmed to collecting deep semantics tags.

Refer to (Bry and Wieser 2012) for a longer and more detailed version of this short paper.

**Acknowledgements**

The authors are thankful to Hubertus Kohle of the Art History department of the Ludwig-Maximilian University of Munich for his many suggestions and explanations.

This research has been founded in part by the German Foundation for Research (DFG) within the project Play4Science number 578416.

**References**


