

di Fabian Bross, Silvia Cramerotti e Alessandra Puglisi



Sono 1,8 miliardi gli utenti di internet in tutto il mondo. Questa cifra, da sola, ci fa capire la portata rivoluzionaria di questo mezzo di comunicazione, nato appena qualche decennio fa e divenuto in brevissimo tempo una componente costante ed essenziale della vita quotidiana di una moltitudine di persone differenti per provenienza, età ed interessi, ma legate dalla loro presenza in “rete”.

La grande sfida cui si vedono confrontate al giorno d'oggi le discipline umanistiche, è quella di riuscire a servirsi dell'enorme mole di informazioni e di utenti che internet mette a disposizione sfruttandole a fini scientifici e didattici.

Il grandissimo vantaggio che la rete propone rispetto ai metodi di ricerca tradizionali è quello di poter raggiungere, in pochissimo tempo, una quantità di persone prima totalmente impensabile.

Il primo a rendersi conto di ciò e a tentare quindi di sfruttare a scopi scientifici il tempo che le persone quotidianamente trascorrono su internet, è stato, nel 2005, l'informatico guatemalteco Luis von Ahn. A lui si deve il concetto di *human computation* con cui si riferisce a metodi che combinano l'intelligenza umana con quella elettronica per risolvere problemi che il computer senza apporto umano non potrebbe risolvere. Secondo il cosiddetto paradosso di Moravec infatti, compiti in media difficili e laboriosi per un essere umano, come per esempio la risoluzione di complicati calcoli matematici, non rappresentano per il computer alcun problema, mentre attività semplicissime di cui anche un bambino sarebbe capace, come il riconoscimento di un volto o di una voce, non possono essere effettuate da un calcolatore e richiedono l'intervento umano.

[\[1\]](#)

Combinando questa premessa con la predisposizione naturale degli esseri umani per il gioco, von Ahn sviluppa i cosiddetti GWAP (*games with a purpose*)

)

[\[2\]](#)

, giochi

crowdsourcing

la cui risoluzione, banale per un essere umano, contribuisce in modo significativo ad importanti questioni di carattere scientifico. La grande innovazione rispetto ai tradizionali metodi di ricerca è la totale anonimità dell'informante di cui non si conosce nulla (né sesso, né età, né livello di istruzione o provenienza ecc...). Questa carenza viene però compensata dalla possibilità sopraccennata di raggiungere in tempi brevissimi, una quantità elevatissima di informanti: la validazione dei dati raccolti avviene quindi grazie alla loro costante conferma da parte degli utenti ed anche eventuali risposte erranee possono essere in tal modo isolate

dalla "massa".

Un innovativo esempio dell'applicazione che i GWAP possono avere in campo umanistico è il progetto Play4Science, finanziato dalla Fondazione Tedesca per la Ricerca Scientifica (Deutsche Forschungsgemeinschaft). Esso vede coinvolti italianisti e storici dell'arte in cooperazione con informatici e linguisti computazionali per la realizzazione di due sottoprogetti: Metropolitatia che mira alla raccolta di dati e metadati nell'ambito della linguistica italiana ed Artigo teso all'etichettatura di quadri nell'ambito della storia dell'arte.

Si tratta in entrambi i casi di operazioni (il riconoscimento e la valutazione di un accento o di un'espressione tipica di un luogo e l'associazione di un'immagine a determinati concetti) che gli esseri umani effettuano quasi automaticamente ma che non possono essere risolte dal computer attraverso l'inserimento di un semplice algoritmo. Esse comporterebbero quindi un investimento cospicuo di tempo da parte di un eventuale programmatore che operasse senza l'appoggio della comunità virtuale.

Nei paragrafi successivi si illustrerà il funzionamento di questi GWAP e le loro implicazioni scientifiche.

Metropolitatia: un esempio dell'impiego di GWAP nella linguistica italiana

La piattaforma linguistica Metropolitatia – <http://www.metropolitatia.org/> – nasce allo scopo di sfruttare il sapere metalinguistico di ciascun parlante per investigare l'universo delle diverse varietà dell'italiano. Se prima di internet gli specialisti della lingua erano costretti a raccogliere i loro materiali autentici spostandosi 'sul campo', da un luogo all'altro ed investendo ingenti quantità di tempo e di denaro, oggi, grazie alla rete, ogni parlante può inserire liberamente il suo contributo stando comodamente seduto sul divano di casa sua. Ciò significa che si possono raggiungere contemporaneamente utenti provenienti da diverse parti d'Italia sfruttando la loro naturale propensione al gioco e il loro sapere vissuto sull'italiano in generale e sulle varietà da loro parlate in particolare.

Lo scenario che fa da cornice a Metropolitatia è quello di una profonda ristandardizzazione della lingua italiana basata sulla progressiva accettazione di tratti tipici della lingua parlata [3] per forze più o meno regionali: non si tratta sicuramente di uno sviluppo unitario, che investe alla stessa misura l'intero spazio linguistico nazionale ma, piuttosto, di tendenze concorrenti e divergenti. Lo stesso tratto della lingua parlata può, infatti, assumere un diverso status a seconda delle regioni, sia in base alla sua marcatezza che al suo grado di standardizzazione. Per esempio, un costrutto totalmente accettato al Nord, come la collocazione

Al giocatore vengono presentate diverse espressioni (5 per ogni turno) e il suo compito è quello di valutarne la diffusione geografica (a livello nazionale o regionale) e di localizzarle servendosi di una cartina sviluppata sulla base di Google Maps. La sua localizzazione può essere effettuata sia su un livello generico (Italia settentrionale, Centrale o Meridionale) che su livelli più specifici (regione, provincia, città) come esemplificato nelle figure 2a, b e c.

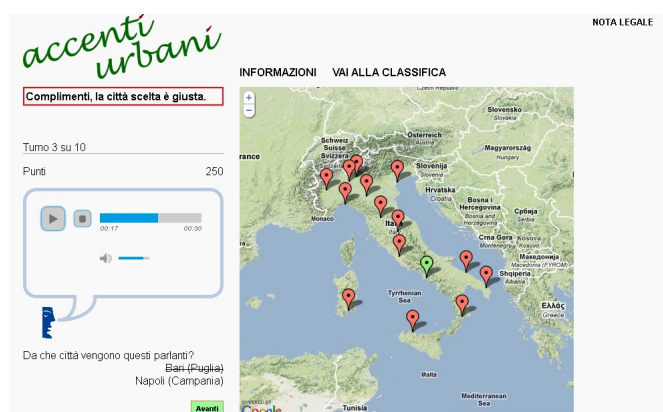


Facoltativamente, l'utente può poi fornire di volta in volta anche valutazioni diastatiche dell'espressione in questione, ipotizzando il sesso, l'età e il livello di istruzione di chi potrebbe averla usata. Infine, egli ha la possibilità di aggiungere alla banca dati nuove espressioni che, a suo parere, sono tipiche di una determinata zona.

In questo modo si ha la possibilità di raccogliere sia dati linguistici autentici, che dati metalinguistici emergenti dalla valutazione soggettiva del giocatore.

L'idea alla base è infatti quella che per la ricostruzione di varietà linguistiche sia fondamentale analizzare, oltre alla produzione di dati, anche la consapevolezza che il parlante ha del suo patrimonio linguistico e del modo di parlare della sua comunità.

Il progetto attualmente è ancora in fase sperimentale e non è accessibile agli utenti in rete. Non è quindi possibile prevedere del tutto eventuali risultati nè tantomeno l'impatto sul pubblico. Fin ad oggi, solo un sottoprogetto di Metropolitaitalia, il gioco Accenti Urbani è andato online lo scorso 22 giugno ed è tuttora fruibile in versione demo (con un numero limitato di prove acustiche). Esso ha riscosso grande interesse da parte degli utenti del web, facendo sperare simili risultati anche per l'intero GWAP Metropolitaitalia.



Accenti Urbani

In Accenti Urbani (<http://labs.play4science.org/accentiurbani/>), l'utente si trova confrontato, non con espressioni scritte ma con dati acustici estratti dalla banca dati online CLIPS (Corpora e Lessici Italiano Parlato e Scritto).

Oltre alla localizzazione geografica degli accenti di volta in volta proposti (10 per ogni turno), questo gioco si distingue per la possibilità di associare ad ogni accento un grado di apprezzamento tramite degli smiley. In questo modo, è possibile analizzare quali accenti godono di un maggiore prestigio e quali sono invece stigmatizzati dagli utenti.

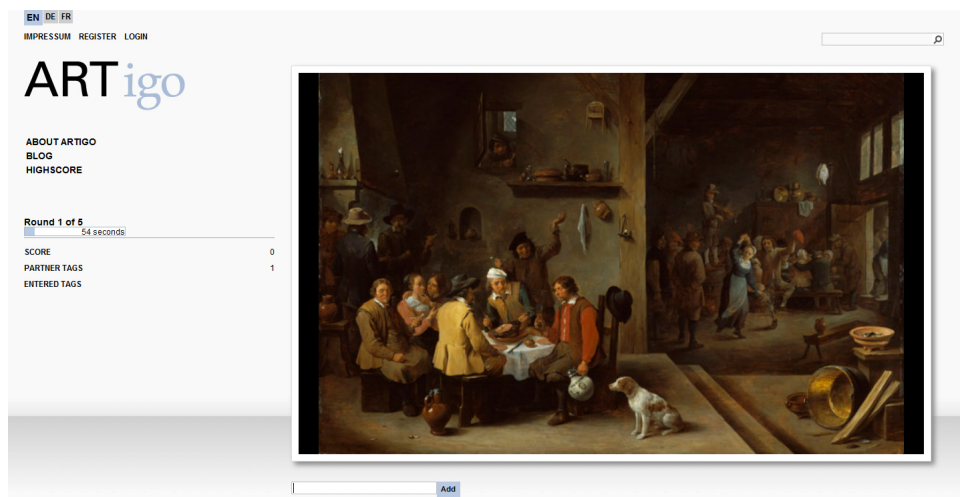
Il numero più che soddisfacente di giocatori raggiunto è dovuto anche alla presentazione ludica del progetto basata sull'assegnazione di punti per ogni accento correttamente identificato. Facoltativamente, gli utenti possono inoltre registrarsi indicando sesso, data di nascita e provenienza con l'ausilio di una cartina geografica interattiva. Il vantaggio per loro è rappresentato dalla possibilità di salvare il proprio punteggio ed entrare eventualmente a far parte della classifica dei migliori. Per la ricerca queste informazioni sono invece importanti in quanto possono fornire risposte a domande come:

- il sesso influisce sulla valutazione di un determinato accento?

-le persone provenienti dal sud, valutano più negativamente gli accenti del nord e viceversa?

-le persone provenienti dal nord, riconoscono con più facilità gli accenti del nord?

Le potenzialità di GWAP come Metropolitania ed Accenti Urbani sono dunque molteplici e si spera che in futuro potranno essere incrementate anche con l'aiuto ed il supporto degli utenti di internet a cui, naturalmente, si chiederà non solo di giocare ma anche di esprimere un loro parere su eventuali miglioramenti e cambiamenti da apportare alla piattaforma.



Artigo e Karido – Esempi di giochi *crowdsourcing* nella storia dell'arte

In rete esistono moltissime banche dati che permettono agli utenti di cercare opere d'arte e che rappresentano, senza alcun dubbio, un grande aiuto anche per gli storici dell'arte. Spesso, tuttavia, questi strumenti presentano dei limiti, rappresentati, ad esempio, dalle ridotte funzioni di ricerca. La raccolta di numerosi metadati per le riproduzioni digitali di opere d'arte richiede infatti un investimento di tempo considerevole in quanto questa operazione non può essere eseguita automaticamente dal computer attraverso un algoritmo. Per la riproduzione digitale questo significa, che mentre un essere umano associa quasi automaticamente un'opera d'arte al medioevo dopo una breve occhiata, trovare un algoritmo che permetta al computer di fare lo stesso appare molto difficile. ^[4]

E' proprio in queste situazioni che, come illustrato nei paragrafi precedenti, l'utilizzo di GWAP offre un importante supporto.

ARTigo (<http://www.artigo.org>) è un gioco competitivo per browser realizzato come JEE application. Lo scopo di questo GWAP è quello di ottenere descrizioni semantiche di opere d'arte digitalizzate che serviranno poi come base per un ampio motore di ricerca di opere d'arte online accessibile a tutti. Si tratta quindi da una parte di un gioco basato sul principio del *crowdsourcing*, dall'altra parte di un motore di ricerca per l'arte.

Una volta entrato nel sito di ARTigo, il giocatore ha la possibilità di effettuare il *login* per poter accumulare punti con il suo nome utente e poter accedere alla classifica dei migliori. Ovviamente, è possibile giocare anche senza essere registrati. Il *layout* del gioco è rappresentato in figura 4. ARTigo prevede sempre due giocatori alla volta. Agli utenti online, collegati anonimamente, vengono presentate cinque opere d'arte (ciascuna per un minuto). Il compito del giocatore è quello di trovare quante più parole chiave possibili (i cosiddetti *tag*) per il quadro a lui di volta in volta presentato. Contemporaneamente, il giocatore vede il numero dei *tag* inseriti dal suo avversario ma non sa quali siano. Ogni qualvolta entrambi i giocatori inseriscono lo stesso *tag* si crea un *match* e vengono assegnati dei punti. Un altro modo per accumulare punti è quello di digitare un *tag* già assegnato a quel quadro in partite precedenti da altri giocatori. Come per Metropollitalia, anche in ARTigo il *matching* assicura l'affidabilità dei dati raccolti. In tre anni sono stati generati dagli utenti circa 6 milioni di *tag* con un totale di 800.000 *match*. Poiché un collegamento diretto degli utenti in alcuni casi potrebbe causare lunghe attese, se i giocatori online sono troppo pochi, si utilizzano vecchie partite per simulare il confronto. Può quindi capitare che nascano dei *match* anche se una parola chiave era stata fornita tempo addietro. Ogni partita si conclude con una panoramica delle opere *taggate* dai fini pedagogici. I giocatori scoprono infatti autore, luogo di esposizione, titolo e anno dei quadri. Questa fase ha un senso anche da un punto di vista psicologico dell'apprendimento poiché dati di questo tipo sono più facili da memorizzare se ci si è prima dedicati ad i quadri in maniera intensiva, come succede in ARTigo.

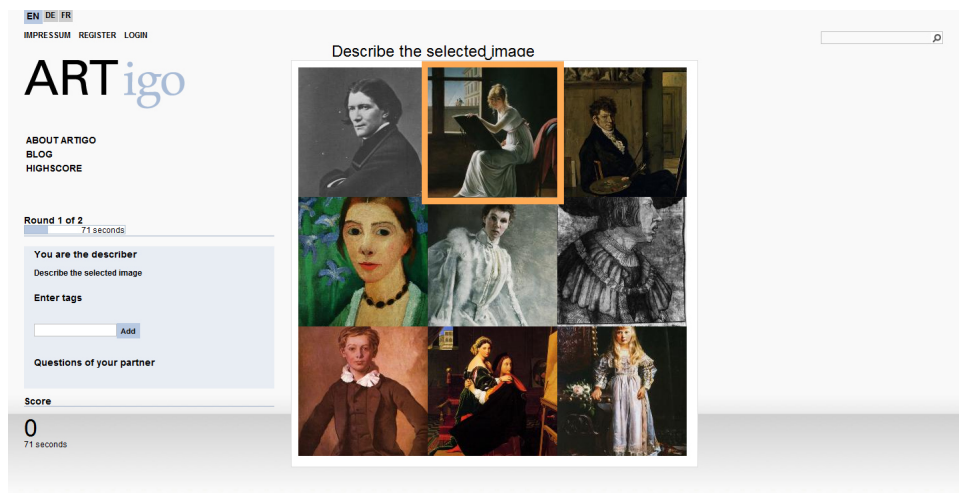


Figure 6: Layout of Karido. The application is designed to be used by two players. The describer is responsible for describing the selected image, while the partner is responsible for providing questions. The application is designed to be used by two players. The describer is responsible for describing the selected image, while the partner is responsible for providing questions. The application is designed to be used by two players. The describer is responsible for describing the selected image, while the partner is responsible for providing questions.