

FROM: Hannah Hickey (हान्नह हिक्की)
(206) 543-2580
hickeyh@u.washington.edu
(नोट: शोधक को सम्पर्क करने की सूचना आखरी में ।)

For Immediate Release
Sept. 12, 2007

एक गुलाब अर्थात् rózsza अर्थात् 薔薇 : चित्र - खोज साधन में सैकड़ों भाषाएँ

बाबुल वाले इमारत के गिरने से लेकर एस्पेरान्टो विश्वव्यापी भाषा गतिविधि तक मनुष्य ने सम्पूर्ण जगत में एक सामान्य बोली बोलने के कई स्वप्न देखे हैं । इन्टरनेट द्वारा विश्वव्यापी संचारव्यवस्था के बावजूद भाषा-अवरोध अधिक हैं । वेब पर तस्वीरों भी अनुवाद की भूलभुलैया में खो जाती हैं ।

“चित्र तो सामान्य होते हैं परन्तु उनकी खोज (इमेज सर्च) नहीं ,” ऐसा कहा ओरेन् एटजियोनी , यूनिवर्सिटी आफ वाशिंगटन (यूडब) में कम्प्यूटर विज्ञान तथा यंत्र-विज्ञान के प्रोफेसर ने । “अँग्रेजी में टाईप किया हुआ सर्च चीनी भाषा में चिन्हित चित्र नहीं देता और एक डच व्यक्ति अँग्रेजी में चिन्हित चित्र नहीं ढूँढ पाता । हमने इस समस्या का समाधान करने हेतु एक साधन (टूल) की रचना की है । ”

यूडब के टयूरिंग सेंटर में आयोजित किया यह नवीन बहुभाषी सर्च टूल तस्वीरों की सामान्यता को वास्तविकता प्रदान कर लोगों को अपनी अपनी भाषा में चित्र खोजने का मौका देता है । पैनइमेजिज (PanImages) नामक यह साधन आज कोपनहागन, डेनमार्क के मशीन अनुवाद (एमटी) सम्मेलन (समिट) में प्रस्तुत किया जा रहा है ।

गूगल (Google) जैसे सर्च इंजिन, सर्च टर्म को अनुशीर्षक तथा तस्वीर के आस-पास के अवतरण (लेख) से मिलाकर चित्र की खोज करते हैं । लेकिन यह खोज अक्षरों तक ही सीमित होने के कारण खोज करने वाले की मात्रभाषा तक ही सीमित रह जाता है । इसी जगह पैनइमेजिज सर्च टर्म का अनुवाद लगभग 300 भाषाओं में कर कुछ सुझाव सहित गूगल व फ्लिकर (Flickr) की सहायता से चित्र प्रस्तुत करता है ।

पैनइमेजिज उन लोगों की सहायता करने की चेष्टा करता है जिनकी भाषाओं के वेब-साईट की संख्या बहुत कम है । उदाहरणतः यदि आप जूलू भाषी हैं और फ्रिज (जिसे जूलू में इफ्रिजि कहते हैं) की तस्वीरें सर्च करें तो केवल दो इमेजिज (चित्र) प्राप्त होंगे ।

इसी जगह पैनइमेजिज आपको 472,000 चित्र देगा। ऐसे ही कम वेब मात्रा वाली भाषाओं के जाँच से यह पता चला कि पैनइमेजिज गूगल से 57 गुणा अधिक चित्र खोजता है।

एटजियोनी ने कहा, “हम उन तमाम लोगों की सेवा करना चाहते हैं जो दुनिया की किसी मुख्य भाषा के वक्ता नहीं हैं। जैसे जैसे इन्टरनेट की प्राप्यता संसार के औद्योगिकृत देशों के बाहर बढ़ रही है वैसे ही उन लोगों की सेवा करना आवश्यक है जो अंग्रेजी, फ्राँसीसी, या चीनी भाषा न बोलते हों।”

जो लोग दुनिया की कोई मुख्य भाषा बोलते हैं वे भी पैनइमेजिज द्वारा दूसरी भाषाओं का लाभ उठा सकते हैं। एक शब्द के कई अर्थ होते हैं जो अवांछित चित्र इमेज सर्च में देता है। जैसे अंग्रेजी में ‘स्प्रिंग’ शब्द का सर्च अनेक प्रकार की तस्वीरें लौटाता है — वसंत ऋतु के रूप में हरे भरे मैदान, मेटल के लच्छे, और अफ्रीका की सिवर स्प्रिंग नामक जगह। अब अगर आपको मेटल के लच्छे वाला स्प्रिंग चाहिए तो पैनइमेजिज द्वारा फ्राँसीसी शब्द ‘रेस्सोर्ट’ का सर्च लाभदायक सिद्ध होगा। वैसे ही अंग्रेजी में ‘बार’ शब्द के सर्च में, जब आपको आयतकारी रूप वाला अर्थ चाहिए नकि शराब पीने की जगह, पैनइमेजिज द्वारा रूसी शब्द “бар” का खोज फायदेमंद होगा। परीक्षण से यह पता चला कि सामान्य भाषाओं की इमेज सर्च में पैनइमेजिज पहले पंद्रह पेजों पर सही चित्रों की संख्या दुगुना कर देता है।

पैनइमेजिज की रचना 350 से अधिक मशीन पठनीय शब्दकोशों को स्केन कर की गई है। इनमें से कुछ विकशनरी (Wiktionaries) नामक बहुभाषी शब्दकोश हैं जिसे वालंटियरों ने लिखा है। पैनइमेजिज साफ्टवेअर इन शब्दकोशों को स्केन कर एक कलनविधि द्वारा सर्च के परिणामों की सत्यता जाँच करता है। इसके उपरान्त यह एक मैट्रिक्स में एकत्रित कर गुजराती से लीथुएनियन सर्च टर्म जैसे अद्वितीय अनुवाद का प्रयास करता है।

एटजियोनी ने कहा, “यह एक अभूतपूर्व शाब्दिक संसाधन है।” उन्होंने यह भी कहा कि, “पैनइमेजिज की सबसे विशेष कुशलता प्रचुर मात्रा में भाषानुवाद है। हमारा अंतिम लक्ष्य है वे सारी भाषाओं को पैनइमेजिज में लाना जिन्में जन की रुचि हो।”

एटजियोनी ने कहा कि याहू (Yahoo!) व गूगल जैसे मुफ्त आनलाइन अनुवाद सेवाओं के पास केवल एक या दो दर्जन भाषाएँ हैं तथा अफ्रीकी अनुवादी अनुसंधान अरबी व चीनी जैसी सैन्य उपयोगी भाषाओं पर गौर करता है। पैनइमेजिज के पास साल

के आरम्भ में 50 भाषाएँ, जून 2007 तक 100 भाषाएँ, और अब 300 भाषाएँ, 25 लाख शब्द तथा लाखों अनुवाद हैं ।

भविष्य में पैनइमेजिज कई अधिक आनलाइन शब्दकोशों को एकत्रित कर अपने शब्दों व भाषाओं की संख्या बढ़ाने की चेष्टा करेगा । शोधकों की आशा है कि डेलिशियस (del.icio.us) जैसे टैगिंग वेबसाईट (जहाँ दर्शक सम्पूर्ण पेज का वर्णन एक शब्द में करते हैं) में उपस्थित शब्दों का भी अनुवाद पैनइमेजिज की सहायता करे ।

एटजियोनी ने कहा कि हमारा उद्देश्य समस्त भाषाओं के अनुवाद को बढ़ावा देना है और पैनइमेजिज रूपी इस प्रथम चरण से हम आकांक्षा करते हैं कि यह वेब सेवा अति लाभप्रद होगी ।

###

अधिक जानकारी के लिए प्रो० एटजियोनी को सम्पर्क करें —

(206) 685-3035 अथवा etzioni@cs.washington.edu

पैनइमेजिज को आजमाने के लिए जाएँ —

www.PanImages.org

यूडब के ट्यूरिंग सेंटर के बारे में जानने के लिए जाएँ —

<http://turing.cs.washington.edu/>