

POUR DIFFUSION IMMÉDIATE

# Communiqué

## D'importants organismes de financement de la recherche internationaux dévoilent les noms des gagnants du premier concours du Digging into Data Challenge


(Ottawa, le 3 décembre 2009) – Aujourd'hui, l'honorable Gary Goodyear, ministre d'État (Sciences et Technologie), a annoncé le nom des gagnants du Digging into Data Challenge, un nouveau concours international qui consiste en un partenariat entre les quatre importants organismes de recherche suivants : le Conseil de recherches en sciences humaines (CRSH) du Canada; le National Endowment for the Humanities (NEH) et la National Science Foundation (NSF) des États-Unis; le Joint Information Systems Committee (JISC) du Royaume-Uni. Le Digging into Data Challenge fait la promotion d'une recherche en sciences humaines novatrice axée sur l'analyse de données à grande échelle.

« Notre gouvernement appuie les sciences et la technologie afin de créer des emplois, d'améliorer la qualité de vie des Canadiens et de consolider l'économie, a indiqué le ministre Goodyear. Soutenir des partenariats de recherche internationaux aide nos universités à former, à recruter et à conserver les meilleurs chercheurs au Canada. »

« Cette nouvelle initiative conjointe du CRSH, du NEH, de la NSF et du JISC permettra aux chercheurs canadiens de perfectionner les technologies sophistiquées de visualisation de données ainsi que d'exploration de textes et d'images tout en favorisant la création de partenariats de recherche internationaux, a ajouté Chad Gaffield, président du CRSH. Grâce aux vastes ressources numériques actuellement à la disposition des chercheurs, ces projets permettront de produire de nouvelles connaissances interdisciplinaires. »

« Essayer de gérer une avalanche de données et de transformer des parcelles d'information en connaissances utiles est un problème auquel nous avons presque tous à faire face en cette ère numérique, a ajouté Jim Leach, président du NEH. Grâce à ce programme de subventions international, notre organisme espère financer des projets qui non seulement seront avantageux pour les chercheurs en sciences humaines, mais mèneront également à une meilleure compréhension mutuelle sur le plan culturel. »

Les 22 candidats retenus représentent huit équipes de recherche internationales composées de chercheurs et de spécialistes qui utilisent des techniques informatiques avancées dans le cadre de leurs travaux de recherche en sciences humaines. Ces équipes, formées de chercheurs d'au moins deux des pays participants, démontreront comment les outils d'exploration et d'analyse de données scientifiques peuvent servir à



améliorer la recherche menée en sciences humaines. Le financement total accordé par les quatre organismes subventionnaires s'élève à 2 082 926 \$CAN.

La description des huit projets subventionnés est présentée ci-dessous.

-30-

Le Conseil de recherches en sciences humaines est un organisme fédéral qui encourage et appuie la recherche et la formation en milieu universitaire dans le domaine des sciences humaines. Grâce à ses programmes et à ses politiques, il permet d'atteindre les plus hauts niveaux d'excellence en recherche au Canada et favorise la collaboration ainsi que le partage des connaissances entre les disciplines, les universités et tous les secteurs de la société. Pour en savoir plus, consultez son site Web ([www.sshrc-crsh.gc.ca](http://www.sshrc-crsh.gc.ca)).

Créé en 1965 à titre d'organisme fédéral autonome, le National Endowment for the Humanities appuie l'enseignement de l'histoire, de la littérature, de la philosophie et d'autres disciplines des sciences humaines. Ses subventions permettent d'enrichir l'apprentissage en classe, de conserver et de produire des connaissances ainsi que de faire naître de nouvelles idées au moyen de la télévision, de la radio, de nouvelles technologies, d'expositions muséographiques et de programmes lancés dans des bibliothèques et d'autres lieux communautaires. Pour en savoir plus, consultez son site Web ([www.neh.gov](http://www.neh.gov)).

La National Science Foundation est un organisme fédéral autonome qui finance la recherche fondamentale et l'éducation dans toutes les disciplines des sciences et du génie. Elle dispose d'un budget annuel de 6,06 milliards de dollars, qui profite aux 50 États américains par l'intermédiaire de subventions attribuées à plus de 1 900 universités et établissements. Chaque année, elle reçoit environ 45 000 demandes de financement et accorde plus de 11 500 nouvelles subventions. Elle verse également plus de 400 millions de dollars par année en financement de services professionnels et contractuels. Pour en savoir plus, consultez son site Web ([www.nsf.gov](http://www.nsf.gov)).

Le Joint Information Systems Committee est un comité mixte qui est composé des organismes britanniques de financement de la formation complémentaire et de l'enseignement supérieur. Il est responsable d'appuyer l'utilisation novatrice des technologies de l'information et des communications dans le cadre de l'apprentissage, de l'enseignement et de la recherche. Il est surtout connu pour offrir au Royaume-Uni un réseau d'infrastructures nationales, une gamme de services d'appui, de contenu et de consultation ainsi qu'un large éventail de ressources de grande qualité. Pour en savoir plus, consultez son site Web ([www.jisc.ac.uk](http://www.jisc.ac.uk)).

*This document is also available in English.*

Pour obtenir des renseignements supplémentaires sur ce communiqué et d'autres projets de recherche subventionnés par le CRSH, veuillez communiquer avec :

**Trevor Lynn**  
Gestionnaire, Communications  
Tél. : 613-992-7302  
Tél. cell. : 613-302-9879  
Courriel : [Trevor.Lynn@sshrc-crsh.gc.ca](mailto:Trevor.Lynn@sshrc-crsh.gc.ca)



## Fiche d'information

Le Digging into Data Challenge est un concours international lancé en janvier 2009 par les quatre importants organismes de recherche suivants : le Conseil de recherches en sciences humaines (CRSH), du Canada; le National Endowment for the Humanities (NEH) et la National Science Foundation (NSF), des États-Unis; le Joint Information Systems Committee (JISC), du Royaume-Uni. Le présent événement vise à annoncer le nom des gagnants du premier concours, qui ont été sélectionnés dans le cadre d'un rigoureux processus d'évaluation par les pairs.


L'avènement de ce qui a été appelé la « recherche axée sur les données » ou la « cyberrecherche » a modifié la nature de la recherche de renseignements dans de nombreuses disciplines – y compris les sciences humaines –, ce qui a offert de nouvelles possibilités de collaboration interdisciplinaire sur des questions d'intérêt commun. La production d'une grande quantité de données numériques accessibles dans Internet ainsi que l'élaboration de techniques de consultation et d'analyse de données de grande envergure ont mené à de nouvelles découvertes remarquables en génétique, en astronomie et dans d'autres domaines, mais surtout à une collaboration entre diverses disciplines universitaires. De nouvelles techniques d'analyse de données de grande envergure permettent aux chercheurs de découvrir des relations, de déceler des différences et d'effectuer des calculs sur des ensembles de données qui sont si vastes qu'on ne peut les traiter qu'à l'aide d'outils informatiques et de méthodes de calcul qui ont été conçus et rendus abordables au cours des dernières années. Maintenant que les livres, les journaux, les revues, les films, les œuvres d'art et les documents audio font l'objet d'une numérisation de grande envergure, il est possible d'utiliser des techniques d'analyse de données pour de vastes collections de ressources patrimoniales et culturelles diversifiées ainsi que des données scientifiques. Comment ces techniques peuvent-elles aider les chercheurs à utiliser ces outils afin de poser de nouvelles questions sur notre monde et de mieux le comprendre?

Pour favoriser l'adoption d'approches novatrices à ce sujet, en janvier 2009, les quatre organismes internationaux ont lancé le présent concours conjoint – visant à attribuer des subventions – afin d'orienter les activités des sciences humaines vers l'analyse de données de grande envergure et son utilisation possible pour une grande diversité de ressources de recherche.

Cette initiative comporte les objectifs suivants :

- promouvoir l'élaboration et l'utilisation de techniques de recherche novatrices dans le cadre de l'analyse de données de grande envergure;
- favoriser la collaboration interdisciplinaire chez les chercheurs de certains domaines – notamment les sciences humaines, l'informatique et les sciences de l'information – en ce qui concerne des questions d'analyse de textes et de données;
- promouvoir la collaboration internationale;
- travailler avec de vastes sources de données comportant d'importantes collections numériques afin que les chercheurs puissent facilement y avoir accès.


On a demandé aux candidats de former des équipes internationales dont les membres devaient provenir de plusieurs des pays participants. Les équipes gagnantes recevront une subvention d'au moins deux des organismes subventionnaires et seront invitées à



présenter leurs travaux au cours d'une conférence spéciale. On demandera à ces équipes – pouvant être composées de chercheurs et de scientifiques – de montrer en quoi les outils actuels d'exploration et d'analyse de données peuvent contribuer à l'amélioration de la recherche en sciences humaines. On espère que ce concours permettra à ces projets de servir de modèles dans leur domaine et de favoriser l'établissement de nouveaux partenariats internationaux entre des chercheurs, des informaticiens, des documentalistes, des bibliothécaires et d'autres intervenants.

Les huit projets subventionnés sont les suivants :

- 1. Structural Analysis of Large Amounts of Music Information** (Stephen Downie, University of Illinois at Urbana-Champaign, NSF; David De Roure, University of Southampton, JISC; Ichiro Fujinaga, Université McGill, CRSH) – Ce projet servira à amasser près de 23 000 heures de musique numérique regroupant un vaste éventail de styles, de régions et d'époques. L'objectif : mettre au point des outils permettant d'analyser et de marquer les structures sous-jacentes des morceaux de musique retenus afin de révéler un ensemble de musique du monde, ce qui permettrait aux musicologues d'avoir accès à des analyses et à des données qui n'étaient pas disponibles de façon interactive jusqu'à présent.
- 2. Digging into the Enlightenment: Mapping the Republic of Letters** (Dan Edelstein, Stanford University, NEH; Chris Weaver, University of Oklahoma, NSF; Robert McNamee, University of Oxford, JISC) – En s'appuyant sur un corpus de 53 000 lettres datant du 18<sup>e</sup> siècle, ce projet permettra d'analyser le point auquel l'effet du Siècle des lumières transparait dans la correspondance de gens issus de diverses professions.
- 3. Using Zotero and TAPoR on the Old Bailey Proceedings: Data Mining with Criminal Intent** (Daniel Cohen, George Mason University, NEH; Tim Hitchcock, University of Hertfordshire, JISC; Geoffrey Rockwell, University of Alberta, CRSH) – Ce projet visera à considérer et à comparer les circonstances entourant des crimes afin de déterminer si ceux-ci relèvent de la violence criminelle ou non criminelle.
- 4. Towards Dynamic Variorum Editions** (Gregory Crane, Tufts University, NEH; John Darlington, Imperial College (London), JISC; Bruce Robertson, Mount Allison University, CRSH) – Ce projet servira à créer une importante ressource pour les chercheurs en études classiques, car il permettra de mettre au point des outils de comparaison dynamique, de produire des lexiques, de sélectionner des sujets et d'extraire des citations dans plus de 10 000 textes écrits en grec et en latin.
- 5. Digging into Image Data to Answer Authorship Related Questions** (Dean Rehberger, Michigan State University, NEH; Peter Bajcsy, University of Illinois at Urbana-Champaign, NSF; Peter Ainsworth, University of Sheffield, JISC) – En s'appuyant sur trois ressources particulières, ce projet permettra de concevoir des outils visant à analyser et à déterminer la paternité des supports visuels.
- 6. Harvesting Speech Datasets for Linguistic Research on the Web** (Mats Rooth, Cornell University, NSF; Michael Wagner, Université McGill, CRSH) – Ce projet regroupera les données sonores et transcrites de fichiers balados, de nouvelles diffusions, de conférences publiques ou éducatives et d'autres sources afin de constituer un répertoire complet de discours. Par la suite, on créera des outils permettant d'analyser ces communications.

- 
- 7. Railroads and the Making of Modern America—Tools for Spatio-Temporal Correlation, Analysis, and Visualization** (William Thomas, University of Nebraska-Lincoln, NEH; Richard Healey, University of Portsmouth, JISC) – Ce projet servira à rassembler une vaste collection de données textuelles, géographiques et numériques sur les chemins de fer de divers siècles, à commencer par les Grandes Plaines et le Nord-Est des États-Unis.
  - 8. Mining a Year of Speech** (Mark Liberman, University of Pennsylvania, NSF; John Coleman, University of Oxford, JISC) – Ce projet permettra de créer des outils offrant un accès facile et rapide à plus de 9 000 heures de fichiers audio tirés de certains des principaux recueils sonores britanniques et américains.

De plus amples renseignements sont disponibles à l'adresse suivante :  
[www.diggingintodata.org](http://www.diggingintodata.org).