Quarterly Progress Report
Covering Period: 10/1/2011 - 12/31/2011
Date of Report: 01/15/2012

Award Number: DE-EE0004449

Project Title: Unconventional and Renewable Energy Research Utilizing Advanced Computer Simulations (UT)

Project Period: 09/23/2010 - 03/22/2012

Recipient Organization: SCI Institute, University of Utah
72 south Campus Drive, Room 3750
Salt Lake City, Utah 84112

Partners: Chris R. Johnson, PhD
Martin Berzins, PhD
Ross Whitaker, PhD
Mike Kirby, PhD
Cost Sharing Partner: University of Utah

Technical Contact: Chris R. Johnson, PhD
72 south Campus Drive, Room 3750
Salt Lake City, Utah 84112
(801) 585-1867
crj@sci.utah.edu

Business Contact: Edward Black
72 south Campus Drive, Room 3750
Salt Lake City, Utah 84112
(801) 585-1867
ed@sci.utah.edu
Project Objective:

The ability to develop science-based and validated computational tools to simulate and facilitate the development of clean, highly efficient energy systems of the future requires innovation in several key computational science technologies, including scientific data management, scientific visualization, scientific software environments, and scientific computing. The overall objective of this work is to leverage our expertise and experience in both scientific visualization and complex science-based simulations toward the accurate and robust simulation of science-based phenomena in the area of unconventional and renewable energy research. This work is aimed at garnering a better understanding of science-based phenomena in energy research and also the advancement of the Uintah software system. The Uintah software system accommodates the massive amounts of data and advanced algorithmic, software, and hardware technologies required to deal with the enormity and complexity of the simulation data in this area of research. To accomplish these goals, we are creating new numerical and visualization techniques needed to assess the uncertainty of the simulation, extend the Uintah scientific problem-solving environment for large-scale simulation of science-based systems, and integrate and extend the data provenance infrastructure of Uintah to systematically capture provenance information and track simulation parameter studies.

Background:

Science-based development of clean and efficient energy systems often involves modeling and simulations of fluid flows, chemical reactions and mechanical properties within heterogeneous media. As part of our DOE-funded (1997-2009) Center for Simulation of Accidental Fires and Explosions (C-SAFE), we created the Uintah scientific problem-solving environment. Uintah is a parallel software environment for solving large-scale computational mechanics and fluid dynamics systems, and has particular strengths when dealing with systems that require large deformations, fire simulation, and fluid-structure interactions. Uintah, general-purpose, fluid-structure interaction code has been used to characterize a wide array of physical systems and processes encompassing a wide range of time and length scales - from microseconds and microns to minutes and meters. Complex simulations require both immense computational power and complex software. Typical
simulations include solvers for structural mechanics, fluids, chemical reactions, and material models, which are efficiently integrated to achieve the scalability required to perform the simulations. Uintah scales to cores by using a novel asynchronous task-based approach for challenging AMR applications. Novel parallel computing algorithms, on both CPUs and GPUs, are needed when simulating large-scale complex science-based energy systems. In moving beyond petascale, it will be necessary to make use of GPU-like architectures as the ongoing convergence between GPUs and multi-core CPUs continues. Task-based codes like Uintah are very well placed to exploit such architectures.

The challenge for finite element type simulations is that the memory access patterns are not well suited for the cache coherency required for efficient operations on streaming architectures. The problem becomes worse for sparse systems associated with large simulations. Thus, the performance improvements over CPU implementations have been limited. An alternative is to take advantage of the geometric configuration of unstructured meshes, and to invent compact, efficient data structures that allow SIMD processing of individual cells and subsequent SIMD assembly of cell computations and mapping onto global degrees of freedom in the solution. The problem becomes more challenging for algorithms that are effective on multi-GPU clusters, such as the NVIDIA cluster at the SCI Institute. We anticipate the need for hierarchical domain decompositions that provide sufficient computational density and efficient communication. This work will pursue GPU and GPU-cluster based algorithms for numerical simulations of combustion using both generic linear solvers and specialized solutions that directly map unstructured and structured domains onto streaming architectures.

The system must also provide data visualization capabilities that allow interaction and analysis of the simulated data. The SCI Institute is an international leader in scientific visualization research. The PI, Chris Johnson, co-leads the DOE Visualization and Analytics Center for Enabling Technology (DOE-VACET). In this work we are leveraging our expertise in large-scale visualization research and development toward the seamless integration of high-end visualization techniques with simulation results of science-based energy systems. Additionally we are exploring the use of higher fidelity visualization with methods based on the use of high-order mesh elements.

With large computational simulations there is substantial uncertainty inherent in any prediction of science-based systems. A number of factors contribute to uncertainty, including experimental measurements, mathematical formulation, and the way different processes are coupled together in the numerical approach for simulation. Tracking of and analysis of this uncertainty is critical to any work that will truly impact the creation of future energy systems.

Exploration of large-scale scientific systems using computational simulations produces massive amounts of data that must be managed and analyzed. Because of the volume of data manipulated, and the complexity of the simulations and analysis workflows, which are iteratively adjusted as users generate and evaluate hypotheses, it is crucial to maintain detailed provenance (i.e., audit trails or histories) of the derived results. Provenance is necessary to ensure reproducibility as well as enable verification and validation of the simulation codes and results. In order to manage large-scale simulations and the analysis of their results, we will use systems such as the VisTrails software (http://www.vistrails.org), an open-source provenance management and scientific workflow system that was designed to support the scientific discovery process, to guide us in building "hooks" into Uintah for provenance systems.
Accomplishments

Uintah Software System

Uintah is a general purpose, fluid-structure interaction code has been used to characterize a wide array of physical systems and processes. Examples include stage-separation in rockets, the biomechanics of micro-vessels, the effects of wounding on heart tissue, the properties of foam under large deformation, and evolution of transportation fuel fires. Uintah was initially targeted at the heating of an explosive device placed in a large hydrocarbon pool fire and the subsequent deflagration explosion and blast wave. The Uintah core is designed to be general and extensible, and is appropriate for a wide range of PDE algorithms applied to a broad problems class. The broad applicability of Uintah makes it suitable for a broad range of energy problems such as flow through micro and nano-scale porous media as well as combustion examples related to flares, boilers and fires. The aim of the Uintah work is that Uintah may be used to solve challenging energy-related problems on large scale high-performance computers such as the Jaguar XT5 and the new Titan XK6 at Oak Ridge National Laboratory. In doing this we were aided by a DOE INCITE allocation for 2011. A major challenge with Uintah is to extend the dynamic task-based approach to heterogeneous architectures using, for example, GPUs as nodes, as it will be the case with the new Titan architecture. In addressing these two challenges, we will first consider the solution of four energy related problems two of which will be solved at scale on the multicore architectures of Jaguar and Titan and then address using GPUs.

Porous Media - Chemo-Thermo-Poro Mechanical Characterizations

The first problem concerns the storage/movements of fluids and solids in nanoporosity and nano-permeability environments and involves multiphase flow mechanisms including coupled phenomena with poro-thermo-chemo-mechanical behavior. Hydraulic fracturing for hydrocarbon and geothermal energy recovery is growing in importance as a means for providing a significant fraction of our energy requirements. This technique requires a deep and fundamental understanding of various multiphase flow mechanisms in a complicated nanoporous and nanopereable environment involving the storage and movement of fluids and solids. These mechanisms include understanding the coupled porous, thermodynamic, chemical and mechanical behaviors. The Material Point Method (MPM) that is a fundamental component of Uintah can be used as a simulation tool to improve our understanding of the mechanical behavior of realistic rock surrogates undergoing triaxial compression loading. The results of these simulations provide us with a picture of the local failure mechanisms by monitoring the energy releases. From this understanding of the failure mechanisms on realistic rock formations we can minimize the calibration procedures that are necessary out in the field. We have used MPM to perform the Brazilian Test on a Limestone sample based on the geometry from CT image reconstructions. This test allows us to determine the tensile strength with results when compared to experiments. Cohesive zones have been added to the MPM algorithm to account for the highly irregular nature of the rock surrogates. By understanding the cohesive properties of the smaller calibration runs, we will be in a position to formulate a multi-thousand irregularly shaped grain domain. The computational requirements of this problem require using large scale parallel computers and we will be asking for access to DOE and NSF large scale parallel computing resources.

Generic Fluid-Structure Interaction Problems

Solving fluid-structure interactions has long been a feature of Uintah. Making good use of large
scale parallel architectures in doing so has proven to be more challenging than we anticipated in software engineering terms. We were using Uintah on Jaguar to do very large scale fluid structure calculations and have demonstrated scaling of up to 198K cores on Jaguar (XT5) using both an MPI only scheduler as well as a hybrid threads/MPI scheduler. We moved to the hybrid threads/MPI to help reduce the memory footprint of the global data structures as well as improve load balancing and scheduling. The hybrid/threads model has worked well with improvements in scaling results with AMR Fluid-Structure simulations run on the XT5 system. These improvements in scaling are shown below for a case in which a solid is transported through fluid. The four solid lines represent runs with a fixed problem size in each case (strong scaling) and the dashed lines represent a fixed problem size per core.

![AMR MPMICE: Scaling](image.png)

**Figure 1:** Scalability of Uintah on Fluid-Structure Interaction Problem.

In this work we focused on the scaling performance on the Titan system using incompressible flow calculations. We performed two weak scaling runs with two different CFD components that are part of Uintah. Both used the HYPRE linear solver (version 2.70b). For both simulations, we used an MPI only scheduler. Using these parameters we found nice improvement in the overall performance, and better weak scaling characteristics up to 131K cores on the XK6 system.
We attribute the performance improvements to the improved communication infrastructure of the XK6. Shown in Figures 2 and 3 are two plots showing XT5 vs XK6

A significant advance realized in this work is the demonstration of good weak scaling with the HYPRE solver for implicit calculations using the new Uintah component, "Wasatch". This required careful use of HYPRE to limit initialization and preservation of data structures. The weak scaling results are shown in Figure 2. This result is very encouraging in terms of presenting the ability to use an implicit solver to accelerate calculation on large core counts. This work saw a performance difference on the XK6 resulting in speeds of roughly 2x per core relative to the XT5 and also showed superior weak scaling on the XK6.
The second simulation scenario concerns the validation of the Large Eddy turbulence closure model used by O’Hern et.al and is performed by simulating a buoyant helium plume. The scenario consists of a 3m\(^3\) domain with a 1m opening that introduces the helium into a quiescent atmosphere of air. The sides and top of the cube are modeled using pressure and outlet boundary conditions. The CFD solution procedure exercises major components of the overall algorithm, including the modeling of small, sub-grid turbulence scales. Additionally, the coupled problem combines the effects of fluid flow and turbulent scalar mixing for a full spectrum of length and time scales without introducing the complications of combustion reactions. Data from the CFD calculation are compared to the 1m helium plume data collected at the Sandia FLAME facility in Albuquerque, NM. CFD results are compared to time-averaged velocity of both vertical and horizontal components, mixture fraction variables, and turbulence statistics. For this more complex problem the XK6 does not show an advantage over the XT5 per core unless we factor in the smaller number of floating point units on the XK6.

The third simulation scenario concerns DOE-funded work on the large quantities of natural gas produced in the oil refining process. The large quantities of natural gas pose a serious explosion hazard and are disposed of by open air burning; a method commonly referred to as flaring. Incomplete combustion of the natural gas results in the release of hazardous organic compounds, which pose serious risks to the environment and to public health. Real-time monitoring and control of flare combustion efficiency is needed to mitigate the release of these chemicals into the environment. The overall goal of the Uintah calculations is to help improve flare combustion efficiency. This is being accomplished through the use of flare combustion and spectral simulation models to evaluate system design concepts and spectral analysis algorithms. These calculations are being done in collaboration with the Institute of Clean and Secure Energy (ICSE) at the University of Utah. In preliminary experiments centered on these simulations the current calculations run approximately three times faster on the new XK6 relative to the XT5.

**Using Uintah on Hybrid Architectures with Accelerators**

In preparation for hybrid architectures such as the final form of the DOE Titan system, we have begun an investigation of GPU approaches. In June 2011 we undertook a GPU porting effort for Uintah under the auspices of the University of Utah Nvidia CUDA Center of Excellence. This has involved extending the existing hybrid threaded-MPI scheduler (through which the barrier to scalability beyond O(100K) cores due to an MPI-only approach was overcome). Uintah's hybrid threaded-MPI scheduler significantly reduces the memory footprint of the code per core and has demonstrated scalability of up to 196k cores on the DOE Jaguar system. The threaded-MPI scheduler has an adaptive controller that sends ready tasks to execute on available CPU cores. The controller and worker threads all share a single data warehouse. In our modification of this approach, instead of tasks being sent to just available CPU cores, a task may also be sent to an available GPU on node.

The central idea of the multi-threaded scheduler is knowing that all 'ready' tasks can be executed at the same time, the multi-threaded scheduler can dispatch these tasks to different task threads. Each of these task threads will run on different processing units dynamically. Its runtime system has one management thread and several task worker threads. The control thread hosts the data warehouse, processes communication and manages shared memory and task queues. The task threads bind to actual processing cores and can directly access the data warehouse. The management thread hosts the data warehouse, processes communication and manages share memory and
task queues.

The threaded-MPI task scheduler provides the basis for our newest work, extending the multi-threaded scheduler by adding an accelerator-ready queue and GPU task controller to its existing runtime system, shown in Figure 4. The control thread will additionally dispatch tasks from the accelerator-ready queue to available GPUs. This multi-threaded GPU scheduler will allow Uintah to make use of all available execution hardware present in current and emerging architectures. Our design includes the ability to maintain and use CPU versions of GPU-able tasks for further flexibility and remains broad enough to make use of other accelerator designs such as the Intel MIC chip.

![Uintah CPU/GPU Scheduler](image)

**Figure 4:** Uintah CPU/GPU Scheduler.

For our GPU development we are using Nvidia CUDA 4.0 C/C++ extensions to implement all GPU kernels. This work began by first adding CUDA support to the existing Uintah build system. This portion of the work is in place and has been tested on several local and remote GPU clusters. Initial analysis was then done to determine what tasks might readily lend themselves to SIMD parallelism and also exhibited properties that would yield a performance increase. For highly regular Uintah grid variables present in Uintah’s stencil computations, we have exposed the underlying representation (raw memory), essentially a flat array representation of an abstract three-dimensional data structure. This exposed, raw memory is easily mapped onto a GPU where individual cells are then SIMD processed.

Our initial focus has been on migrating portions of the ICE component (the multi-material CFD code) by profiling and analyzing and implementing kernels for several key ICE algorithms, including its first and second order advection operators to estimate and then measure the theoretical and actual performance we might see on Nvidia Tesla C20 series GPUs, typical in most many heterogeneous systems. These types of operations show a high FLOPs to word ratio and in practice we have observed roughly a 2x speedup with naive implementations on larger patch sizes 323 (Figure
5). These comparisons were based on a single Nvidia Tesla C20 series GPU and a single CPU similar to that found in Cray XT5 systems such as Kraken and Jaguar; a dual core AMD Opteron Processor 285, running at 2.6 MHz.

![Figure 5: CPU and GPU Performance on a Uintah Task.](image)

The stencil computations typically found in Uintah components are bandwidth bound in that data must be copied from the host to the device for computation. Results are then copied from device to host. In the case where data may be read and written only once, faster, pinned host memory can be mapped onto the GPU. These data copies are across the PCIe bus, which has a maximum theoretical bandwidth of 8.0GB/s (PCI Express Gen2 x16 for the Nvidia Tesla C20 series cards). Actual bandwidth for the GPU used in our tests has been determined, and in practice is more on the order of 3.3GB/s when using paged memory, and 5.3GB/s using pinned (page-locked) memory. These bandwidth limitations can be overcome to a larger extent by using CUDA Streams and asynchronous memory copies and additionally keeping some data on the GPU, eliminating a portion of the costly transfers. These techniques will serve to overlap the computation and communication and will also allow the multi-threaded GPU scheduler to execute multiple GPU tasks (kernels) concurrently, allowing several smaller kernel launches as opposed to a single larger one. The final work on the GPU scheduler will be to add support so the Uintah infrastructure handles the details of the device memory allocation and asynchronous data copies automatically, in essence having the memory ready on the device prior to task execution. We expect to achieve more significant speedups using this strategy.
Algorithm Development

Smoothness-Increasing Accuracy-Conserving (SIAC) Filtering for discontinuous Galerkin Solutions: Improved Errors Versus Higher-Order Accuracy

Mike Kirby continued his work with James King and finalized a publication on their results. Smoothness-increasing accuracy-conserving (SIAC) filtering has demonstrated its effectiveness in raising the convergence rate for discontinuous Galerkin solution from order $k + 1/2$ to order $2k + 1$ for specific types of translation invariant meshes. Additionally, it improves the weak continuity in the discontinuous Galerkin method to $k-1$ continuity. Typically this improvement has a positive impact on the error quantity in the sense that it also reduces the absolute errors. However, not enough emphasis has been placed on the difference between superconvergent accuracy and improved errors. This distinction is particularly important when it comes to understanding the interplay between geometry and filtering introduced through meshing. The underlying mesh over which the DG solution is built is important because the tool used in SIAC filtering - convolution - is scaled by the geometric mesh size. This heavily contributes to the effectiveness of the post-processor. In this work, we present a study of this mesh scaling and how it factors into the theoretical errors. This work was submitted to the Journal of Scientific Computing.

Mike Kirby also submitted a paper with his student Blake Nelson to Eurovis 2012 and worked on his own on a paper for the Uncertainty Journal of Uncertainty Visualization. Both of these two paper covers two aspects of research, respectively Algorithm development & Visualization, and Uncertainty Quantification & Visualization, and they are both listed below under Scientific Visualization.
Scientific Visualization

Accurate and Interactive Direct Volume Rendering of Spectral/hp Finite Element Fields

We describe a method for the direct volume rendering of scalar fields produced by spectral/hp finite element methods that is both accurate and interactive. While the use of these methods is becoming increasingly common, there are few visualization methods specifically designed for high-order fields. Existing methods that operate on high-order data directly are generally limited to specific cell types and polynomial degrees, require significant hardware resources, or are not interactive. Consequently, visualizations of high-order data are generally created by first approximating the high-order field with low-order primitives and then generating the visualization via traditional methods based on linear interpolation. This approximation introduces error into the visualization pipeline, which requires the user to balance image quality, interactivity, and resource consumption. To address these issues, we have developed a GPU-based raycasting method for direct volume visualization that uses the high-order field directly. Our method evaluates the volume rendering integral using high-order quadrature when the field is smooth and low-order quadrature when it is not. The accuracy and performance of this method is demonstrated in the context of several high-order fluid flow simulations. This work was submitted to Eurovis 2012.

Visualization of covariance and cross-covariance fields

We present a numerical technique to visualize covariance and cross-covariance fields of a stochastic simulation. The method is local in the sense that it demonstrates the covariance structure of the solution at a point with its neighboring locations. When coupled with an efficient stochastic simulation solver, our framework allows one to effectively concurrently visualize both the mean and (cross-)covariance information for two-dimensional (spatial) simulation results. Most importantly, the visualization provides the scientist a means to identify interesting correlation structure of the solution field. This work was submitted to the International Journal of Uncertainty Quantification.
Uncertainty Visualization

Uncertainty Workshop at IEEE VisWeek 2011

Chris Johnson co-chaired the full day workshop: Working with Uncertainty Workshop: Representation, Quantification, Propagation, Visualization, and Communication of Uncertainty during the IEEE VisWeek Conference in October in Providence (http://www.sci.utah.edu/uncertainty2011)

Uncertainty in the Development and Use of Equation of State Models

This work was presented as posters by Kristi Potter at the Workshop during IEEE VisWeek Conference. In this work we present the results from a series of focus groups on the visualization of uncertainty in Equation-Of-State (EOS) models. The initial goal was to identify the most effective ways to present EOS uncertainty to analysts, code developers, and material modelers. Four prototype visualizations were developed to presented EOS surfaces in a three-dimensional, thermodynamic space. Focus group participants, primarily from Sandia National Laboratories, evaluated particular features of the various techniques for different use cases and discussed their individual workflow processes, experiences with other visualization tools, and the impact of uncertainty to their work. Related to our prototypes, we found the 3D presentations to be helpful for seeing a large amount of information at once and for a big-picture view; however, participants also desired relatively simple, two-dimensional graphics for better quantitative understanding, and because these plots are part of the existing visual language for material models. In addition to feedback on the prototypes, several themes and issues emerged that are as compelling as the original goal and will eventually serve as a starting point for further development of visualization and analysis tools. In particular, a distributed workflow centered around material models was identified. Material model stakeholders contribute and extract information at different points in this workflow depending on their role, but encounter various institutional and technical barriers which restrict the flow of information. An effective software tool for this community must be cognizant of this workflow and alleviate the bottlenecks and barriers within it. Uncertainty in EOS models is defined and interpreted differently at the various stages of the workflow. In this context, uncertainty propagation is difficult to reduce to the mathematical problem of estimating the uncertainty of an output from uncertain inputs.

This work has been submitted to the International Journal on Uncertainty Quantification.
From Quantification to Visualization: A Taxonomy of Uncertainty Visualization Approaches


Quantifying uncertainty is an increasingly important topic across many domains. The uncertainties present in data come with many diverse representations having originated from a wide variety of domains. Communicating these uncertainties is a task often left to visualization without clear connection between the quantification and visualization. In this paper, we first identify frequently occurring types of uncertainty. Second, we connect those uncertainty representations to ones commonly used in visualization. We then look at various approaches to visualizing this uncertainty by partitioning the work based on the dimensionality of the data and the dimensionality of the uncertainty. We also discuss noteworthy exceptions to our taxonomy along with future research directions for the uncertainty visualization community.
QuizLens: A Multi-Lens Uncertainty Visualization Environment

The work performed last year by Drs. Joel Daniels and Kristi Potter was submitted to Eurovis 2012. In this work, we propose a multi-lens system that allows users to easily combine a variety of uncertainty-based visualization techniques within a single context. While each lens describes a unique visualization method, they share several components, including the placement, orientation, sizing and references to specific scene geometries. This abstraction enables efficient toggling between the different visualization tasks, as well as scene geometries, in order to better understand data uncertainty and compose hierarchical focus-plus-context visualizations. We demonstrate the use of our system through two different volumetric scalar field examples, including: probabilities for multiple tissue types compiled from segmentations of Magnetic Resonance Image (MRI) data and simulated probabilities of Diffusion Tensor Image (DTI) fiber bundle intersections.

Figure 7: The QuizLens system uses a collection of flippable lenses for the exploration and visualization of uncertainty. Using a lens abstraction, the systems allows for global display, focus-plus-context drill down, and multiple lenses within a single context to show multiple focus points or layers of focus.
Provenance

The VisTrails provenance management system has been extended to capture provenance during a user session with the Uintah problem-solving environment. Currently, a small number of VisTrails user modules are in place to allow the system to execute Uintah-based simulations from within the VisTrails environment. Additionally, VisTrails modules have been created to allow scientists to change the inputs to their simulations. By enabling this functionality within VisTrails, robust provenance-enabled parameter studies are possible using the native properties of VisTrails.

Progress and Status:

- Personnel - Kristi Potter, who recently was hired as a Research Computer Scientist after she finished her Post-Doc at the SCI Institute, was added to the DOE NETL Project. Dr. Potter works closely with Chris Johnson in many projects in Uncertainty Visualization.

- Equipment - No equipment was purchased during this reporting period.

Scope issues:

There are no scope changes.

Budget and Schedule Status:

There are no budget nor schedule status changes.
**Patents:**

There are no patent applications attached to this award.

**Publications / Presentations:**

**Publications**


**Presentations**


Chris Johnson *Visual Computing: Making Sense of a Complex World*, University of Minnesota, Minneapolis, October 2011 (Cray Distinguished Lecture)

Chris Johnson, Meet the Editors, IEEE Visualization Conference, Providence, October 2011 (Panel Presentation)

Chris Johnson *Representation, Quantification, Propagation, Visualization, and Communication of Uncertainty* Working with Uncertainty Workshop, Providence, October 2011 (Panel Presentation).

**Plans for Next Quarter:**

**Uintah Software System**

Major challenges remain with regard to showing that Uintah can be used to solve large scale multi-physics energy problems on the new hybrid architectures such as Titan.

**Algorithm Development**

Ross Whitaker and Mike Kirby are currently working with a student (who was hired as of January 1st, 2012 to the DOE NETL project) to prepare a paper on solving of the linear system by geometrically-informed algebraic multigrid obtained by discretization via FEM of elliptic operators on two dimensional and three dimensional unstructured meshes. This work will be submitted to Journal of Scientific Computing or SIAM Journal of Scientific Computing.

**Provenance Enabling Uintah**

The preliminary VisTrails modules are now in place and user testing and feedback are now required to begin the process of refining how provenance is used with the Uintah system.
Uncertainty Visualization

The continuation of uncertainty visualization work includes investigating measures of uncertainty for categorical data. This work will focus on exploring measures of entropy described by the field of information theory to elucidate uncertainty from data which cannot be described by mean and standard deviation. This work will be submitted as a Eurovis Short Paper in March, 2012.