Many algorithms for nonlinear optimization, including those used in computer vision, rely on computation of derivatives. Such algorithms can benefit greatly from Algorithmic (or Automatic) Differentiation (AD), a technology for transforming a computer program for computing a function into a program for computing the function’s derivatives. In this talk, we will give an overview of theory, algorithms and software we have developed over the years to enable efficient computation of derivative matrices – in particular, Jacobians and Hessians – using AD when the matrices are sparse, which is typically the case in large-scale problems. A fundamental technique that has proven effective in exploiting sparsity in large Jacobians and Hessians is computation via compression. Intuitively, the idea here is to reduce computational cost by calculating groups of columns of a derivative matrix at a time instead of calculating a single column at a time. We will discuss graph-theoretic models and algorithms we developed to address the underlying need for partitioning the columns of a matrix into groups that are amenable for such compression-based computation. (Received February 25, 2017)