Unconventional data: a new web-based indicator for SMEs Innovativeness

by marco guerzoni | Carlo Bottai | Caterina Liberati | Lisa Crosato | Università Bicocca | Università Bicocca | Università Bicocca | Ca' Foscari University

This work represents a first step towards the creation of an automatic, real-time indicator of innovative activity.

Assessing the presence and intensity of innovative activity within a firm is a complex task, first and foremost because a firm's innovativeness is hidden within its practices and the shared knowledge of its employees. Over the past half-century, the innovation economics literature has made significant progress in measuring this phenomenon. To do so, balance sheets, surveys, patents, and product catalogs have been used as the main data sources. However, each of these sources has limitations in capturing such latent features, particularly when applied to small and medium-sized enterprises, as underlined by the OECD [1992].

This study suggests that SMEs' corporate websites can serve as a new data source to improve indicators of the innovative activity of a firm. Being an outcome of the activity of the firm itself, constantly updated, and publicly accessible, corporate websites appear as good candidates to solve the limitations of the currently available approaches.

Based on the prediction probability of a machine learning algorithm on Italian manufacturing companies, we develop an indicator of a firm's innovativeness, we describe its properties and discuss its robustness.