ERRATA SHEET for "SOM−based data visualization methods" in Intelligent Data Analysis 3 (1999), pp. 111−126.
Almost all citations in Tables 1 and 2 were wrong, and some references were completely missing from the bibliography. Below are the corrected Tables and the missing references.

November 22nd 1999
Espoo, Finland
Juha Vesanto
Juha.Vesanto@hut.fi

Table 1
Some vector quantization algorithms [a]

Algorithm	Notes
k−means [1]	Only best matching (closest) cluster center of the sample vector is updated
maximum entropy [29, 30]	All cluster centers are updated according to their distance to the sample vector
neural gas [29]	All cluster centers are updated according to their ranking order in distance to the sample vector
SOM	The cluster centers are updated according to their distance from the BMU of the sample vector on the map grid (h_{ck})

[a] All listed algorithms are iterative and are based on minimizing the reconstruction error of the original data set using a certain number of cluster centers. The difference between the algorithms is the way the cluster centers are updated on each adaptation step.

Table 2
Some vector projection algorithms [a]

Algorithm	Energy function	Notes
Multi−Dimensional Scaling (MDS) [9, 21]	$\Sigma_i \Sigma_{j<i} (X_{ij} - Y_{ij})^2$	Distances in the input space are approximated by distances in output space (typically 2D plane). This is the metric version of MDS
Sammon's projection [31]	$\Sigma_i \Sigma_{j<i} (X_{ij} - Y_{ij})^2 / X_{ij}$	Local distances in the input space are emphasized. Sammon's projection has an inherent instability: the $1/X_{ij}$ term when $X_{ij} \to 0$. This can be corrected by using e.g. $1/max(c, X_{ij})$ in its stead, where c is a suitably selected constant
Curvilinear Component Analysis (CCA) [28]	$\Sigma_i \Sigma_{j<i} (X_{ij} - Y_{ij})^2 f(Y_{ij})$	Local distances in the output space are emphasized. $f(Y_{ij})$ is a function monotonically decreasing with output space distance
SOM	$\Sigma_i \Sigma_k X_{ik}^2 h_{ck}(Y_{ck})$	Input space distances X_{ik} are measured between prototypes (k) and data vectors (i) rather than between all pairs of data vectors. Similarly input space distances are measured between map units (Y_{ck})

[a] In the energy functions X_{ij} denotes the distance between vectors i and j in the input space, and Y_{ij} in the output space. In the case of SOM, X_{ik} denotes the distance between vector i and prototype k. The output space distance Y_{ck} is measured between the BMU of the data sample c and all other map units; h_{ck} is the neighborhood kernel. Note that in the first three algorithms input space distances X_{ij} are fixed, but in the case of SOM they change, since the SOM training algorithm moves the prototype vectors to minimize the quantization error. Optimization of the projection is only secondary goal. Note also that the mathematical treatment of the SOM has proven to be very difficult. The energy function given for the SOM holds only for the case of discrete data set and fixed neighborhood kernel, but it does give a qualitative idea of the situation in the general case.

[28] P. Demartines, J. Hérault, Curvilinear component analysis: a self−organizing neural network for nonlinear mapping of data sets, IEEE Transactions on Neural Networks 8 (1997), pp. 148−154.
[29] T.M. Martinetz, S.G. Berkovich, K.J. Schulten, "Neural−gas" network for vector quantization and its application to time−series prediction, IEEE Transaction on Neural Networks 4 (4) (1993), pp. 558−569.
[30] K. Rose, F. Gurewitz, G. Fox, Statistical mechanics and phase transitions in clustering, Physical Rev. Lett. 65 (8) (1990), pp. 945−948.
[31] J.W. Sammon, Jr., A nonlinear mapping for data structure analysis, IEEE Transactions on Computers C−18 (5) (1969), pp. 401−409.