[redacted]

E HA C[redacted]E GI IC D[redacted]A D EDIC I

ACC AC H GH CLIE E D

H E C EC ED DA A E EMBLE

XINYUE PAN, ASHWIN POTHEN, JANA BOERGER,
HE WANG & BENOIT MONTREUIL

PHYSICAL INTERNET CENTER
SUPPLY CHAIN & LOGISTICS INSTITUTE
SCHOOL OF INDUSTRIAL & SYSTEMS ENGINEERING
GEORGIA INSTITUTE OF TECHNOLOGY

CREATING THE NEXT®

I C I

More **COMPLEX** Current World

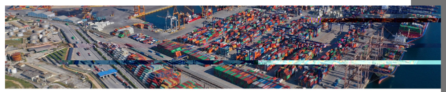
- Volatile
- Uncertain
- Complex
- Ambiguous

Greater **CHALLENGES** in Logistics and Supply Chain Management (LSCM)

- Flexible
- Visible
- Cooperative
- Hyperconnected

Physical Internet Paradigm

- A Metaphor of the Digital Internet
- To solve the inefficient and unsustainable situation in LSCM



I I D

PI

1. Smoothens the installation of necessary **IT system** for data sharing
2. Serves as a key trigger for **digitalizing** LSCM

F

D

1. Supports **seamless interconnection** of LSCM
2. Promotes the **dynamic decision-making**

I D

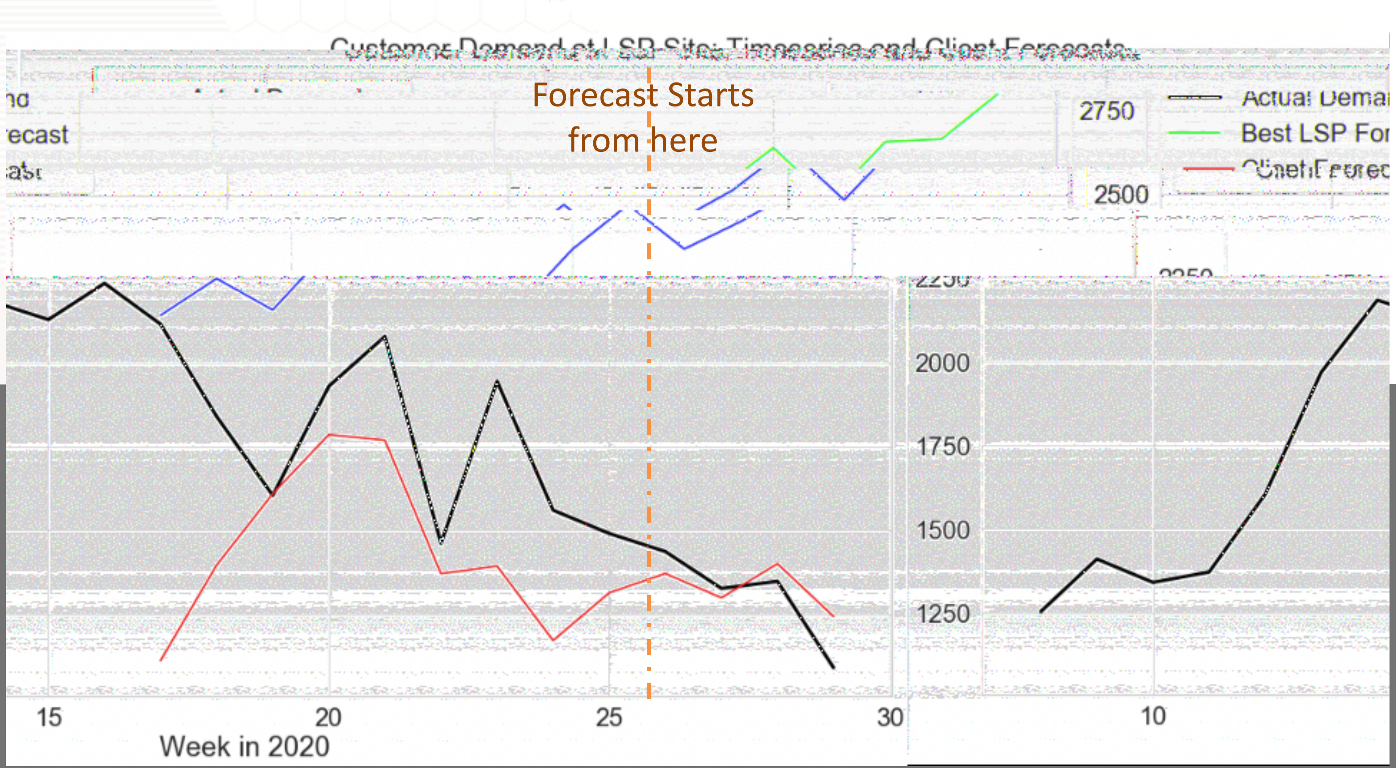


Figure 1: Comparing logistics service provider’s forecasts and client-provided forecasts with real client demand at logistics service provider site

L D

Data Integration

- Integration of databases
- Enterprise Application Integration (EAI)

Sharing Modified Data

- Modifying key contributions of raw data

Federated Learning

- Training global model instead of sharing data

Data Privacy Issue

What if only utilizing insensitive data or the data that clients are willing to provide?

H D E D

- Hyperconnected data ensemble requires as minimal data as possible

- The content of shared data is dependent on clients' choice
  - Aggregated data such as the overall activity forecast
  - ❌ detail information on product or SKU level
  - ❌ Key models
  - ❌ Production plan

H D E F L D

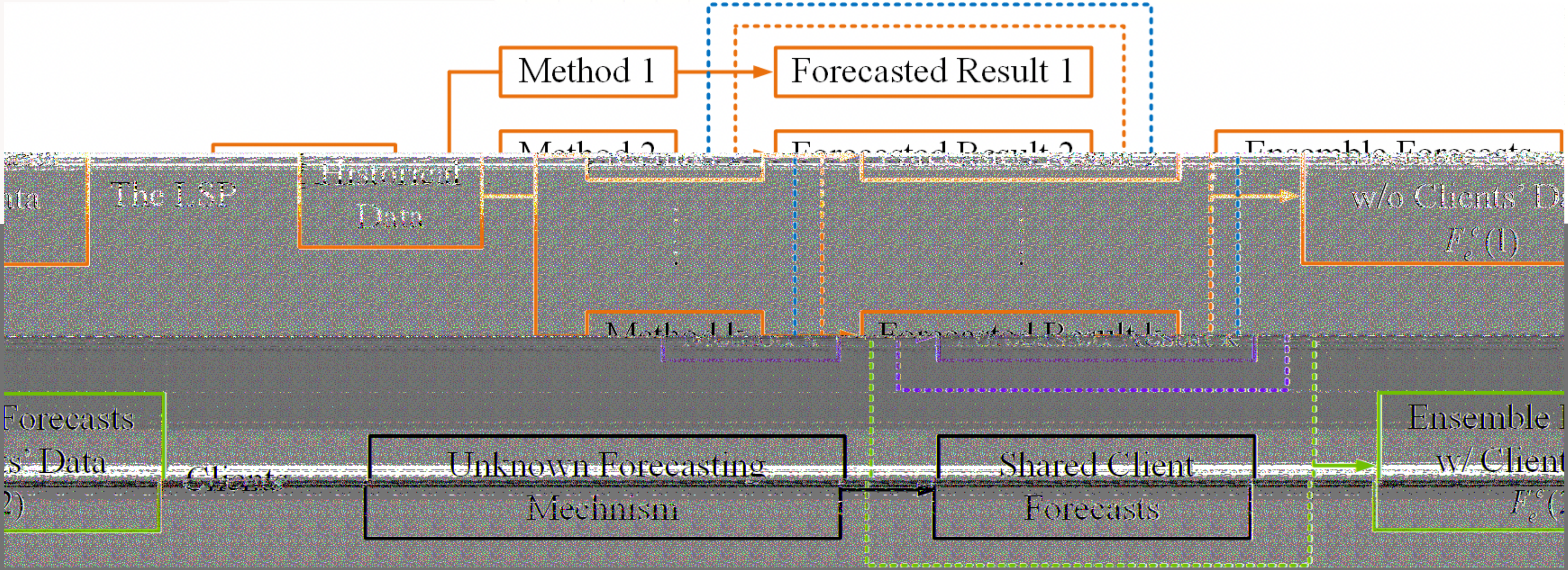
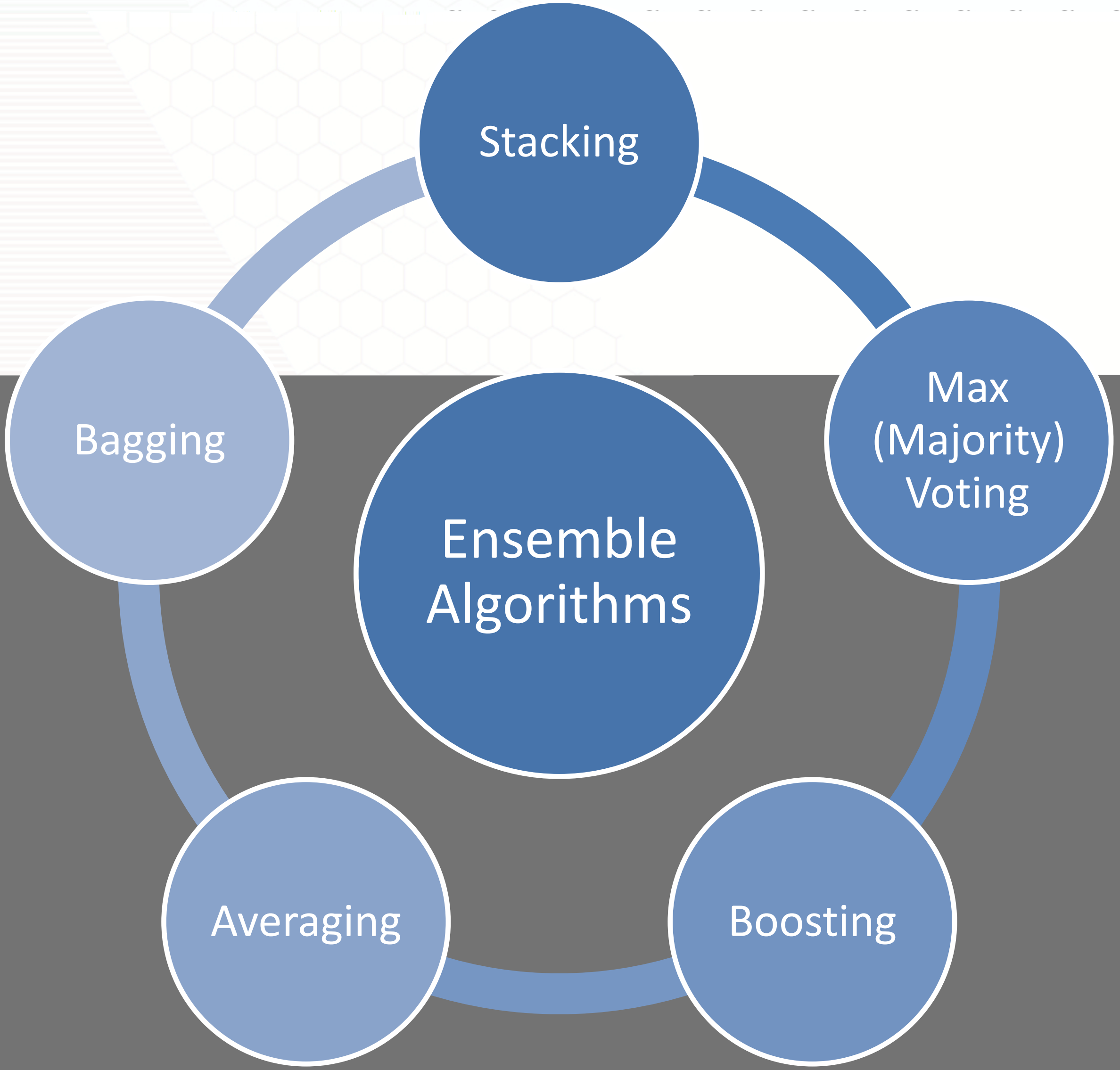


Figure 2: The Framework of logistics demand prediction with hyperconnected data ensembles

E A


Stacking
Max (Majority) Voting
Boosting
Averaging
Bagging
Ensemble Algorithms

Ensemble Method for including shared Client Forecasts

Under the general framework, we propose a weighted average ensemble method for including forecasts shared by client, with data-driven time series forecasts generated by the LSP:

$$F_e^c(1) = \sum_{j=1}^{k} F_{LSP_j}^c \alpha_j,\ \sum_{j=1}^{k} \alpha_j = 1,$$

$$F_e^c(2) = \sum_{j=1}^{k} F_{LSP_j}^c \alpha_j + F_c^c \alpha_{k+1},\ \sum_{j=1}^{k+1} \alpha_j = 1,$$

where,

- superscript c: the client level
- $F_e^c(1)$: the ensembled forecast without shared client data
- $F_e^c(2)$: the final forecast with client data ensembles
- $F_{LSP_j}^c$: the forecasted result generated by the j^{th} method from the LSP
- F_c^c: the forecast shared by clients
- α_j: the j^{th} weighting parameter and $0 \le \alpha_j \le 1$

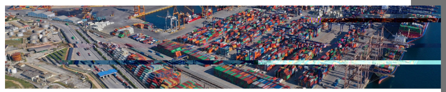
Experiments: Ensemble Methods for major Logistics Service Provider

A major LSP in North America has access to 2.5 years of historical daily demand data for the client's products at 2 sites, i.e. number of pallets shipped from the site to the destination.

The historical data is subject to seasonal fluctuations, and the LSP incorporates 4 timeseries forecasting methods at a daily level:

- Holt-Winters (HW) method (Winters, 1960) - $F_{LSP_1}^{C}$
- Double-Seasonal (DS) method (Taylor, 2003) - $F_{LSP_2}^{C}$
- Multi-Seasonality (MS) method (Gould et al., 2008) - $F_{LSP_3}^{C}$
- Bouchard-Montreuil (BM) method (Bouchard and Montreuil, 2009) - $F_{LSP_4}^{C}$

The client provides weekly aggregated demand forecast for a 13-week horizon for each site. With access to 50 such weekly forecasts (*50 Start Dates*), we conduct experiments to generate logistics demand forecast with data ensembles.



Experiments: Ensemble Methods for major Logistics Service Provider

We employ Monte-Carlo cross validation technique with a fixed number of iterations $N = 30$. For each iteration $i \in \{1, 2, \dots, N\}$, we randomly select 80% of dates as the training set from the total set of start-dates S ($|S| = 50$). Let the training set be $\hat{S} \in S$, and the test set $S' = S \backslash \hat{S}$, then the training error for i-th iteration is calculated using the following error metric:

$$E_i^{TR} = \frac{1}{|\hat{S}|} \sum_{s \in \hat{S}} MAE_s^i = \frac{1}{|\hat{S}|} \sum_{s \in \hat{S}} \frac{1}{13} \sum_{h=1}^{13} \left| A_{s,h}^i - F_{s,h}^i \right|,$$

$$E^{TR} = \frac{1}{N} \sum_{i=1}^{N} E_i^{TR}$$

where

- the superscript/subscript i refers to the number of the current iteration
- $\hat{S} \in S$ refers to the set of start-dates in the training set
- MAE_s^i refers to the Mean Absolute Error of the ensemble forecasts for a given start-date $s \in \hat{S}$
- $A_{s,h}^i$ represents the actual demand starting from date s for h period ahead
- $F_{s,h}^i$ represents the ensembled forecast made from s for h period ahead.

Finally, the optimal weighting parameters are selected by minimizing the final train error E^{TR}.

Results: Visualizing the impact of Ensemble Methods

Site 1: 8% decrease in test error E^{TE} upon integration with the client's forecasts

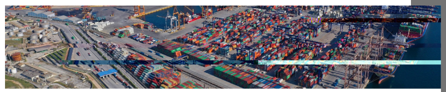
$$F_e^c(1) = 0.382 * F_{LSP_1}^c + 0.403 * F_{LSP_2}^c + 0.183 * F_{LSP_3}^c + 0.032 * F_{LSP_4}^c$$

$$F_e^c(2) = 0.252 * F_{LSP_1}^c + 0.297 * F_{LSP_2}^c + 0.067 * F_{LSP_3}^c + 0.056 * F_{LSP_4}^c + 0.328 * F_c^c$$

Site 2: 30% decrease in test error E^{TE} upon integration with the client's forecasts

$$F_e^c(1) = 0.020 * F_{LSP_1}^c + 0.977 * F_{LSP_2}^c + 0.0032 * F_{LSP_3}^c$$

$$F_e^c(2) = 0.0004 * F_{LSP_1}^c + 0.235 * F_{LSP_2}^c + 0.028 * F_{LSP_3}^c + 0.736 * F_c^c$$



Results: Including client forecasts significantly improves accuracy

Site 1

Demand Forecast Method	CV Error Train	CV Error Test	Improvement Train	Improvement Test
Best LSP Forecast DS	467.27	456.71		
Ensemble of LSP Forecasts $F_e^c(1)$	459.50	449.71	**1.66%**	**1.53%**
Client Forecast $F_{customer}^c$	503.81	482.28	**-7.82%**	**-5.60%**
Ensemble with Client Forecast $F_e^c(2)$	444.15	419.50	**4.95%**	**8.15%**

Site 2

Demand Forecast Method	CV Error Train	CV Error Test	Improvement Train	Improvement Test
Best LSP Forecast DS	450.09	485.31		
Ensemble of LSP Forecasts $F_e^c(1)$	450.67	485.75	**-0.13%**	**-0.09%**
Client Forecast $F_{customer}^c$	342.15	359.69	**23.98%**	**25.88%**
Ensemble with Client Forecast $F_e^c(2)$	324.42	338.09	**27.92%**	**30.34%**

Improvement I is measured as the decrease in CV error $E^{T:TR\ or\ TE}$ of method F versus the best LSP forecast F'

$$I = \frac{E_{F'}^{T} - E_{F}^{T}}{E_{F'}^{T}}$$

The ensemble of the LSP's data with the insensitive client's data improves overall forecasting accuracy

C

- Proposed a hyperconnected data ensembled framework based on data sharing under Physical Internet (PI) paradigm
  - Requires as minimal data as possible
  - The content of shared data is dependent on the clients’ choices in order to avoid the data privacy issue

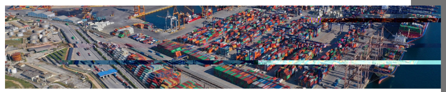
- Conducted computational experiments
  - Demand prediction accuracy increases by integrating the aggregated forecasts from clients

Thanks!

Xinyue Pan, Ashwin Pothen, Jana Boerger,
Prof. He Wang & Prof. Benoit Montreuil
Physical Internet Center
Supply Chain & Logistics Institute
School of Industrial & Systems Engineering
Georgia Institute of Technology, Atlanta, U.S.A.

xpan67@gatech.edu

- ALICE-ETP (2020). Roadmap to the Physical Internet, Executive Version, http://www.etp- logistics.eu/wp-content/uploads/2020/11/Roadmap-to-Physical-Intenet-Executive-Version_Final.pdf
- Ballot, E., Montreuil, B., & Meller, R. (2014). *The Physical Internet*. https://hal-mines- paristech.archives-ouvertes.fr/hal-01113648
- Bennett, N., & Lemoine, G. J. (2014). What a difference a word makes: Understanding threats to performance in a VUCA world. In *Business Horizons* (Vol. 57, Issue 3, pp. 311–317). Elsevier Ltd. https://doi.org/10.1016/j.bushor.2014.01.001
- Boerger, J., & Montreuil, B. (2020). *Data-driven analytics-based capacity management for hyperconnected third-party logistics providers*.
- Bouchard, J., and B. Montreuil (2009). Méthode de prévision journalière de séries temporelles saisonnières. *Congrès international de génie industriel–CIGI-2009*
- Darvish, M., Larrain, H., & Coelho, L. C. (2016). A dynamic multi-plant lot-sizing and distribution problem. *International Journal of Production Research*, *54*(22), 6707–6717. https://doi.org/10.1080/00207543.2016.1154623
- Dwork, C. (2010). Differential privacy in new settings. *Proceedings of the Annual ACM- SIAM Symposium on Discrete Algorithms*, 174–183. https://doi.org/10.1137/1.9781611973075.16
- Gould, P. G., Koehler, A. B., Ord, J. K., Snyder, R. D., Hyndman, R. J., & Vahid-Araghi, F. (2008). Forecasting time series with multiple seasonal patterns. *European Journal of Operational Research*, *191*(1), 207–222. https://doi.org/10.1016/j.ejor.2007.08.024
- Kim, N., B. Montreuil, W. Klibi, N. Kholgade (2020). Hyperconnected Urban Fulfillment and Delivery, *Transportation Research Part E Logistics and Transportation Review*, Vol. 145, 23p. https://doi.org/10.1016/j.tre.2020.102104
- Konečný, J., McMahan, H. B., Yu, F. X., Richtárik, P., Suresh, A. T., & Bacon, D. (2016). *Federated Learning: Strategies for Improving Communication Efficiency*. http://arxiv.org/abs/1610.05492
- Lee, H. L., & Whang, S. (2000). Information sharing in a supply chain. *International Journal of Manufacturing Technology and Management*, *1*(1), 79–93. https://doi.org/10.1504/IJMTM.2000.001329
- Linthicum, D. S. (2000). *Enterprise application integration*. Addison-Wesley Professional,. https://www.worldcat.org/title/enterprise-application-integration/oclc/42397664



- Lu, Y., Huang, X., Dai, Y., Maharjan, S., & Zhang, Y. (2020). Blockchain and Federated Learning for Privacy-Preserved Data Sharing in Industrial IoT. *IEEE Transactions on Industrial Informatics*, *16*(6), 4177–4186. https://doi.org/10.1109/TII.2019.2942190
- Mayr, A., Binder, H., Gefeller, O. and Schmid, M., (2014). The evolution of boosting algorithms-from machine learning to statistical modelling. *arXiv preprint arXiv:1403.1452*.
- Montreuil, B., Ballot, E., & Fontane, F. (2012). An open logistics interconnection model for the physical internet. *IFAC Proceedings Volumes (IFAC-PapersOnline)*, *45*(6 PART 1), 327–332. https://doi.org/10.3182/20120523-3-RO-2023.00385
- Montreuil, Benoit. (2011). Toward a Physical Internet: meeting the global logistics sustainability grand challenge. *Logistics Research*, *3*(2–3), 71–87. https://doi.org/10.1007/s12159-011-0045-x
- Pan, S., Trentesaux, D., McFarlane, D., Montreuil, B., Ballot, E., & Huang, G. Q. (2021). Digital interoperability in logistics and supply chain management: state-of-the-art and research avenues towards Physical Internet. *Computers in Industry*, *128*, 103435. https://doi.org/10.1016/j.compind.2021.103435
- Romera-Paredes, B., Aung, M.S. and Bianchi-Berthouze, N., 2013, April. A One-Vs-One Classifier Ensemble With Majority Voting for Activity Recognition. *In Esann.*
- Sallez, Y., Pan, S., Montreuil, B., Berger, T., & Ballot, E. (2016). On the activeness of intelligent Physical Internet containers. *Computers in Industry*, *81*, 96–104. https://doi.org/10.1016/j.compind.2015.12.006
- Sweeney, L. (2002). L. Sweeney. k-anonymity: a model for k-ANONYMITY: A MODEL FOR PROTECTING PRIVACY 1. In *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems* (Vol. 10, Issue 5).
- Taylor, J. W. (2003). Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of the Operational Research Society*, *54*(8), 799–805. https://doi.org/10.1057/palgrave.jors.2601589
- Themistocleous, M., Irani, Z., & Love, P. E. D. (2004). Evaluating the integration of supply chain information systems: A case study. *European Journal of Operational Research*, *159*(2 SPEC. ISS.), 393–405. https://doi.org/10.1016/j.ejor.2003.08.023

- Ting, K.M. and Witten, I.H., 1997. Stacking bagged and dagged models.
  Tretola, G., Biggi, D., & Verdino, V. (2015). *Enabling Logistics Data Sharing in the Physical Internet*. https://www.researchgate.net/publication/282607787_Enabling_Logistics_Data_Sharing _in_the_Physical_Internet
- Winters, P. R. (1960). Forecasting Sales by Exponentially Weighted Moving Averages. *Management Science*, *6*(3), 324–342. https://doi.org/10.1287/mnsc.6.3.324