Canadian Research Software Conference ‘21

Peer-to-Peer Privacy Preserved Linking

Francis Jeanson, PhD.
A growing number of health studies want to **combine data** for big data analytics/ML

**Linking patient data for precision health analytics**
Patient Linking Remains a Big Challenge...

- To protect privacy, studies cannot share PHI
- Trusted third-parties need to hold secret keys
- Most solutions are deterministic with no way to link probabilistically on many fields
- Current solutions require cloud storage ... or
- Local software installation
A linking app that lives on your browser and connects you securely to peers to find exact or probable patient matches.
Fine Controlled Linking

Select any field to link on

Cryptographically process fields

All processing is done in-browser

No data are stored on the cloud
Peer-to-Peer Encrypted Linking

Create a unique peer link and share it
A secure P2P encrypted connection is created
Data are partitioned for shared compute
A probability score is assigned to matching rows
Matching rows are shared with approval

https://link.datadex.net/session/DF45HKL3972

<table>
<thead>
<tr>
<th>Local Table</th>
<th>Remote Table</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Row 1</td>
<td>Row 3</td>
<td>100%</td>
</tr>
<tr>
<td>Row 2</td>
<td>Row 1</td>
<td>75%</td>
</tr>
<tr>
<td>Row 3</td>
<td>Row 2</td>
<td>0%</td>
</tr>
</tbody>
</table>

Peer 2 Peer

<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>ID</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>GH23PO5</td>
<td>0KB4EF2J</td>
<td>GH23PO5</td>
<td>0KB4EF2J</td>
</tr>
<tr>
<td>8FR20DA</td>
<td>LK2E4RS1</td>
<td>8FR20DA</td>
<td>LK2E4RS1</td>
</tr>
<tr>
<td>90S23VX7</td>
<td>PC3D5S4F</td>
<td>90S23VX7</td>
<td>PC3D5S4F</td>
</tr>
</tbody>
</table>
Tangible Benefits

- Complete data
- Improved statistics
- Increased sample sizes
- Greater significance
- New entity features
- Reveal new correlations
- Peer-to-Peer data exchange
- Strong security
- Browser based
- No installation required
Our Partners & Network